Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

INGEGNERIA ELETTRONICA, TELECOMUNICAZIONI E TECNOLOGIE DELL'INFORMAZIONE

Ciclo 34°

Settore Concorsuale: 09/E3 - ELETTRONICA

Settore Scientifico Disciplinare: ING-INF/01 - ELETTRONICA

DESIGN OF LOW-POWER ANALOG CIRCUITS FOR WAKE-UP RADIO IN IOT NODES

Presentata da: Alessia Maria Elgani

Coordinatore Dottorato

Prof. Aldo Romani

Supervisore

Prof.ssa Eleonora Franchi Scarselli

Co-supervisore

Prof. Antonio Gnudi

Esame finale anno 2022

UNIVERSITÀ DI BOLOGNA

Abstract

ARCES-DEI

Doctor of Philosophy

Design of low-power analog circuits for Wake-up Radio in IoT nodes

by Alessia Maria ELGANI

The Wake-Up Radio is an enabling technology for Wireless Sensor and Actuator Networks, an important structure within the Internet-of-Things. It is a minimal receiver integrated in sensor or actuator nodes which allows a reduction in their power consumption while also enabling asynchronous communication without latency between the gateway and sensor nodes themselves.
This Ph.D. thesis proposes three different implementations for the Analog Front-End of a nanoWatt Wake-Up Radio for medium range applications which receives OOK-modulated signals as input. Two of these implementations have been designed using a 90-nm BCD technology, which would allow the integration of the Wake-Up Radio in nodes with actuation capabilities, thus requiring power devices, whereas the third one has been designed using a 90-nm CMOS technology. The whole activity has been performed in the framework of the STMicroelecronics - ARCES joint lab.
The Analog Front-End of a Wake-Up Radio first extracts the envelope of the input OOK-modulated signal through an Envelope Detector, then digitizes it through a decision circuit. The main challenge this topic poses is the sensitivity-power trade-off, that is a longer wake-up distance requires a higher power consumption.
All proposed implementations are clockless and leverage the second-order non-linearities of a MOSFET in subthreshold for envelope extraction. The first implementation features a band-pass active Envelope Detector AC-coupled to a Schmitt trigger. The input DC voltage and two threshold voltages of the trigger are generated in a temperature compensated way by a dedicated block. However, Manchester coding is needed to avoid multiple errors in case of long sequences of equal bits.
The second implementation solves this issue by using an active Envelope Detector with a low-pass response, which outputs a signal with the same shape as the input envelope, DC-coupled to a standard comparator. Moreover, the reference voltage for the comparator is generated within the Envelope Detector itself. However, an external capacitance is needed to realize the low-pass response. The bulk voltages of the comparator input transistors are available on the outside to set its equivalent threshold. The operation of this implementation has been proven to be temperature stable through simulations.

In the third implementation, a passive Envelope Detector is employed to enhance sensitivity without having to drastically increase ED current consumption. Temperature compensation of the Envelope Detector and the external matching network is performed by a block similar to a Brokaw cell with the aim of stabilizing the Envelope Detector input resistance, thus sensitivity, with temperature. The proposed solutions have performances aligned with the state-of-the-art.

Contents

List of Figures

List of Tables

List of Abbreviations

AFE	**A**nalog **F**ront **E**nd
BBL	**B**ase**b**and **L**ogic
BC	**B**ias and **C**alibration
BCD	**B**ipolar **C**MOS **D**MOS
BER	**B**it **E**rror **R**ate
BP	**B**and-**p**ass
CDR	**C**lock and **D**ata **R**ecovery
CL	**C**ontrol **L**ogic
CMOS	**C**omplementary **MOS**
CTAT	**C**omplementary **T**o **A**bsolute **T**emperature
DC	**D**uty **C**ycling
DMOS	**D**ouble-diffused **MOS**
DTMOS	**D**ynamic **T**hreshold-voltage **M**etal **O**xide **S**emiconductor
ED	**E**nvelope **D**etector
FAR	**F**alse **A**larm **R**ate
FoM	**F**igure **o**f **M**erit
GO	**G**ated **O**scillator
HDL	**H**ardware **D**escription **L**anguage
IC	**I**ntegrated **C**ircuit
IoT	**I**nternet-**o**f-**T**hings
ISM	**I**ndustrial **S**cientific and **M**edical
LNA	**L**ow **N**oise **A**mplifier
LP	**L**ow-**p**ass
MDR	**M**issed **D**etection **R**ate
MOSFET	**M**etal **O**xide **S**emiconductor **F**ield **E**ffect **T**ransistor
OOK	**O**n-**O**ff **K**eying
PTAT	**P**roportional **T**o **A**bsolute **T**emperature
PVT	**P**rocess **V**oltage **T**emperature
RF	**R**adio **F**requency
RMS	**R**oot **M**ean **S**quare
RTL	**R**egister **T**ransfer **L**anguage
SIPO	**S**erial-**I**nput **P**arallel-**O**utput
SoC	**S**ystem **o**n **C**hip
WSAN	**W**ireless **S**ensor and **A**ctuator **N**etwork
WUR	**W**ake **U**p **R**adio

Chapter 1

Introduction

The Internet-of-Things (IoT) is notably the main infrastructure of the connected world we are currently experiencing and will likely be experiencing for decades to come. It is a network of nodes with the aim of connecting all sorts of elements, including daily use objects, to drastically enhance their functionalities by adding cooperation between them. The IoT infrastructure permeates basically every aspect of the life of people, so its applications are the most varied: smart homes and cities, smart agriculture and industry, healthcare, logistics and more.

As mentioned before, the cooperation between nodes is the key to the success and usefulness of the IoT. This is the reason why IoT nodes are organized in sub-networks, called Wireless Sensor and Actuator Networks (WSAN), typically composed of a central node, the gateway, and several end-nodes communicating wireless. Therefore, nodes interact both with the environment and with each other, ultimately providing services to the end user. Typically, the gateway manages all communication within the WSAN, receives data from sensor and actuator nodes, issues instructions to them and provides users access to the gathered data. Therefore, a generic node needs to integrate several subsystems with different tasks, that is sensing, elaboration, communication and actuation. This makes it an actual System-on-Chip (SoC). One of the most suitable technologies for SoCs is BCD, which allows the integration of bipolar, CMOS and DMOS transistors on a unique process platform. Bipolars are typically employed for analog functions, CMOS for digital (or analog) functions and DMOS for power functions.

In order to place nodes virtually anywhere, they are usually battery powered, which sets a strict constraint on their power consumption. Among the many strategies and technologies adopted to reduce WSAN node power consumption is the reduction of communication activity by use of a Wake-Up Radio (WUR). As a matter of fact, most of the power consumption of a node is typically due to its wireless transceiver, so shutting it down whenever it is not needed is a solution. The WUR is an ultra-low-power additional receiver with which the node is equipped. It is an always-on device with the main task of continuously monitoring the channel and waking up the rest of the node only when a Wake-Up signal is received. Once the tasks requested by the gateway have been fulfilled, the node goes back to sleep mode

with just the WUR on. Ultimately, the WUR allows purely asynchronous communication while keeping node power consumption low.

Since the WUR is a minimal receiver, the simplest possible type of modulation, On-Off Keying (OOK) modulation, is usually employed, whereas the carrier RF frequency usually belongs to the ISM frequency band. In the WUR scenario, there are two main issues: the sensitivity-power trade-off and addressing capabilities. Sensitivity is directly linked to the maximum distance at which a Wake-Up event can take place. Typically, the lower sensitivity is, the higher is the power consumption. On the other hand, a node with addressing capabilities can discriminate whether the Wake-Up signal is addressed to itself or to another node, thus avoiding false Wake-Ups and providing additional power saving. It is possible to identify three groups of WURs depending on their power consumption and corresponding sensitivity range. In particular, short range WURs have wake-up distances below 1 m and are completely passive. Medium range WURs have wake-up distances in the order of tens or hundreds of meters and a power consumption in the nanoWatt range. Finally, long range WURs have wake-up distances above 1 km and a power consumption in the μW range. This thesis focuses on nanoWatt WURs for medium range applications, that is with a wake-up distance in the order of tens or hundreds of meters.

Three prototypes have been designed and implemented over the course of this Ph.D. research activity using a 90-nm BCD or CMOS technology in the framework of the STMicroelectronics - ARCES joint lab. The first prototype only includes an Analog Front-End (AFE) for use in a WUR, whereas the second and third prototypes also include a Baseband Logic (BBL), developed by another Ph.D. student.

This thesis is organized as follows:

- **Chapter 2** provides more insight on WSANs and explains why the WUR is a useful tool for the reduction of node power consumption.

- **Chapter 3** details the structure of a WUR, its AFE in particular. Then the main Figures of Merit which can be employed to make comparisons between different architectures are presented. Other solutions proposed in literature are presented as well. A thorough discussion of the main aspects of an AFE for a WUR follows. Among the discussed aspects is the demodulation mechanism based on the second-order non-linearities solicited in subthreshold. As a matter of a fact, operation in subthreshold is required since target applications are ultra-low-power. At the end of the chapter, all three integrated prototypes are introduced.

- **Chapter 4** introduces the basic active Envelope Detector (ED) architecture and thoroughly discusses its band-pass (BP) and low-pass (LP) implementations with their advantages and drawbacks. Then it details the first and the second WUR prototypes, which feature an active ED, including the measurement phase. The first prototype implements a band-pass active ED AC-coupled to

a Schmitt trigger, whose input DC voltage and two thresholds are generated by a dedicated biasing block. The main drawbacks of this implementation are the need for Manchester coding and the complexity and current consumption of the biasing block. The second prototype implements an active low-pass ED DC-coupled to a standard comparator. As an advantage, it does not require the use of the Manchester coding and has a much simpler biasing block. However, it necessitates an external capacitance.

- **Chapter 5** presents the features, advantages and drawbacks of passive EDs. Then it details the third WUR prototype, which implements a passive ED in order to improve sensitivity with respect to the previous versions without a drastic current consumption increase. Moreover, an original temperature compensation technique is implemented in this prototype to stabilize sensitivity with temperature.
- **Chapter 6** presents the conclusions.
- **Appendix A** defines a new Figure of Merit for the AFE of a medium range WUR with the aim of performing fairer comparisons.

Chapter 2

Motivation

Some of the material reported in this chapter is reused from [1] and [2] (©2018, 2020 IEEE), in agreement with IEEE copyright policy on theses and dissertations.

As mentioned in the Introduction, WSANs, the structure enabling cooperation between nodes within the IoT, are typically constituted of a gateway and several sensor and actuator nodes, ultimately providing services to the end user, performing distributed sensing and automatic interactions with the environment. All communication within the network happens through the wireless medium and is managed by the gateway. For practical reasons, nodes are typically battery-powered. Limited battery lifetime currently poses a big challenge to researchers as it affects network costs both directly through battery costs and indirectly through maintenance costs. Therefore, it is prime to find effective strategies to minimize the power consumption of nodes, thus increase node battery lifetime and ultimately reduce network costs.

2.1 Current solutions to energy constraints in IoT nodes

Both hardware and software solutions have been proposed in recent years. As mentioned in the Introduction, IoT nodes are actual SoCs, including sections performing sensing, actuation, data elaboration and communication with the other nodes. A power management section may be needed as well in case the node also performs the harvesting of energy from the environment. Such a situation does not change the fact that the transceiver is typically the most power-hungry section of the whole node. This prompted researchers to find ways to shut down the node transceiver whenever its operation is not needed.

A well-known power reduction approach, called Duty Cycling (DC), consists in periodically turning the node off, especially its main radio and its microcontroller, ("sleep mode") and back on ("transmitting mode") according to a fixed or variable time schedule [3]. Some of this activity is required in order to keep network synchronization. This method aims at reducing the idle-listening power consumption when the radio is only listening to the medium without the need to transmit any data. It is very effective for power consumption reduction but it introduces significant latency in communication. As a matter of fact, communication can only be synchronous,

whereas an event-driven approach cannot be adopted. Moreover, it does not completely eliminate the idle-listening consumption as the radio is periodically switched on even when there is no need for communication [3].

The WUR is a valid alternative to DC, as it enables asynchronous communication and reduces overall power consumption without introducing significant latency [4] [5]. WURs are always-on devices with the main task of continuously monitoring the channel and waking up the rest of the node only when a Wake-Up signal is received. As a matter of fact, the activity required for network synchronization is not needed and idle-listening power consumption is completely saved. Therefore, WURs are particularly suitable for event-driven applications.

2.2 Wake-Up Radio

As just mentioned, the purpose of a WUR is to monitor the channel instead of the main node transceiver and wake the rest of the node up whenever a Wake-Up signal is received. This is beneficial in terms of the overall power consumption of the node because the WUR is a minimal receiver, which trades performance, often in terms of bandwidth, for a much lower consumption than that of the main transceiver.

2.2.1 Wake-Up Radio operation

Its operation is presented in the following.

- A general node equipped with a WUR is normally in sleep mode. This implies the WUR is the only portion of the node which is on, as shown in Figure 2.1. As a result, the overall node power consumption corresponds to the consumption of the WUR itself.
- Whenever the WUR receives a Wake-Up message, it outputs an impulse, which acts as an interrupt for the rest of the node and wakes it up. Figure 2.2 shows this wake-up mode.
- The node is then pushed to its main mode of operation and can fulfil the requests sent by the gateway. As Figure 2.3 suggests, the overall power consumption of the node in this mode corresponds to that of the main radio and the microcontroller, since the power consumption of the WUR is largely negligible.
- Once the tasks requested by the gateway have been completed, the node goes back to sleep mode, as shown in Figure 2.4. Again, the overall power consumption corresponds to that of the WUR.

Therefore, the working principle of the WUR allows the power consumption of the main radio to be replaced by that of the WUR, which is much lower, for most of the time. This gives nodes mounting a WUR a large advantage in terms of energy consumption over those performing asynchronous communication without a WUR.

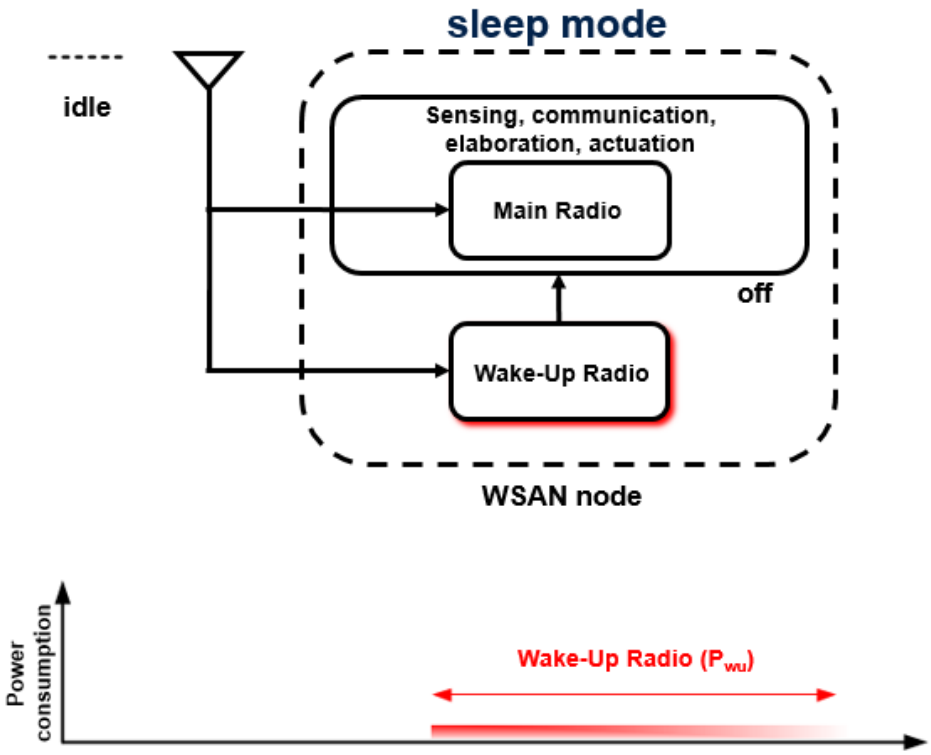

Figure 2.1: A WSAN node equipped with a WUR in sleep mode.

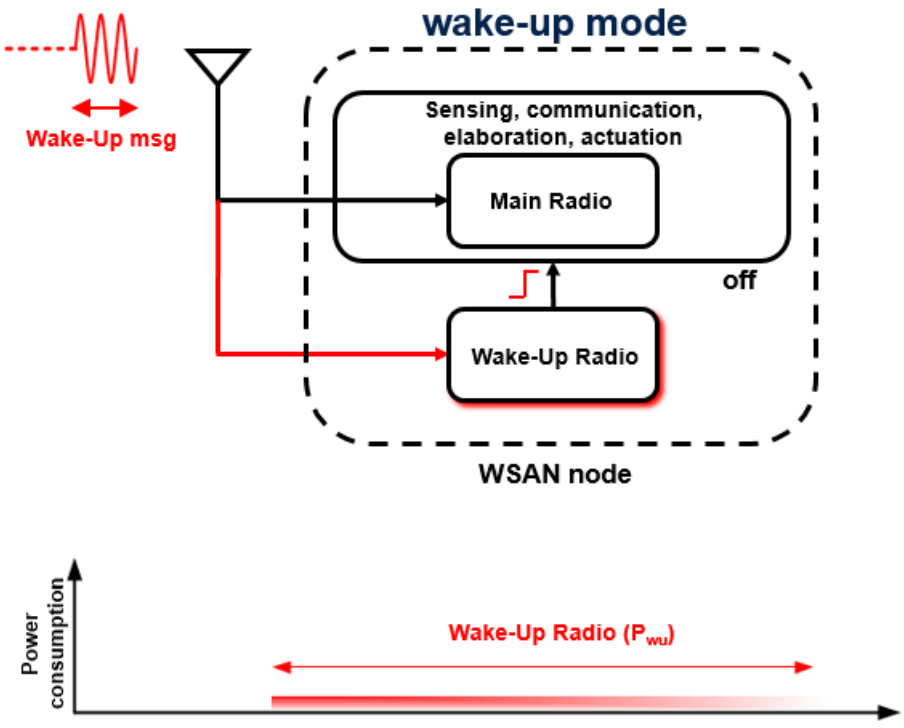

Figure 2.2: A Wake-Up message is received by a WSAN node equipped with a WUR.

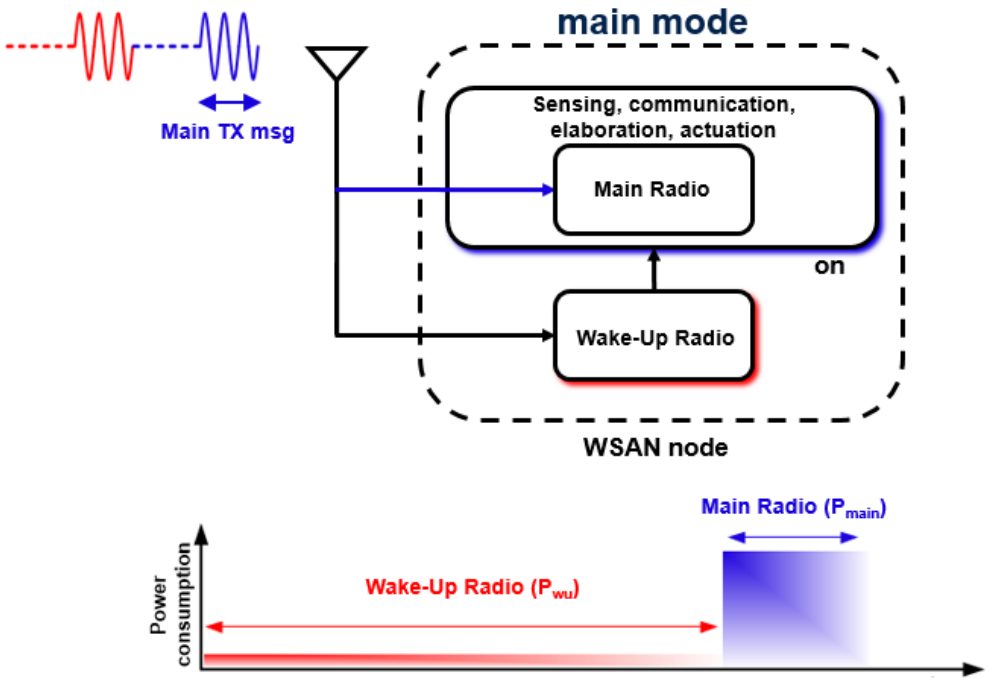

Figure 2.3: A WSAN node equipped with a WUR in its main mode of operation.

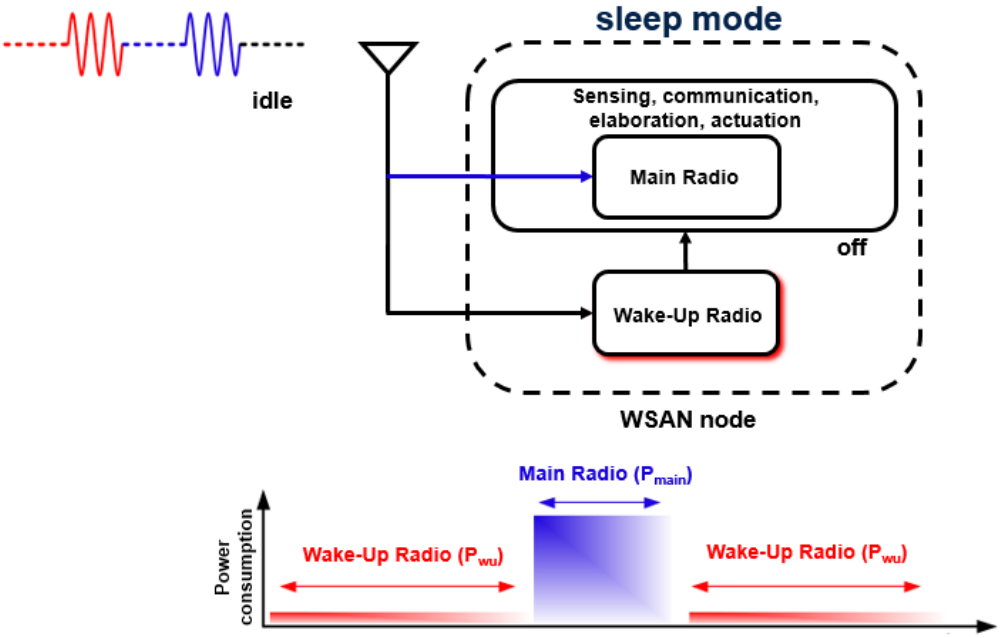

Figure 2.4: A WSAN node equipped with a WUR back to sleep mode after the tasks requested by the gateway have been completed.

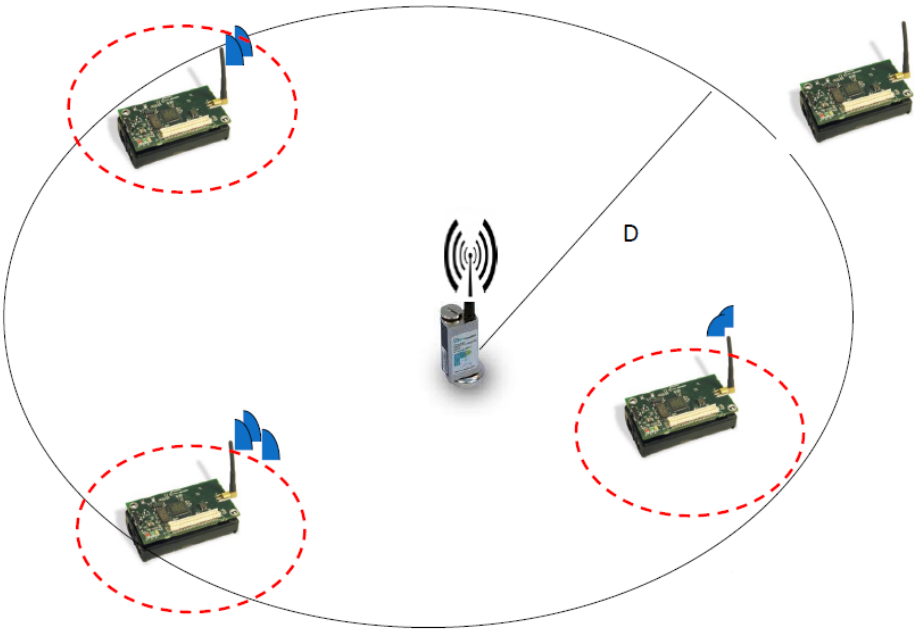

Figure 2.5: A WSAN composed of a gateway and several sensor and actuator nodes, some of which are within the maximum wake-up distance and can thus be activated by the gateway through a wake-up message.

2.2.2 Wake-Up Radio scenario

OOK modulation

Two aspects contributing to the efficiency and convenience of WURs are the choices of OOK modulation and carrier frequencies in the ISM band. Since power minimization is one of the main targets when designing WURs, the simplest possible modulation, OOK modulation [6], is usually employed. Moreover, the carrier frequency for the OOK-modulated signal is usually chosen within the ISM band, such as 433 MHz or 868 MHz, as this frequency band is free for use in IoT applications.

Wake-up distance ranges

As long as WURs are concerned, two important parameters are sensitivity and addressing capabilities.
Let's consider a generic WSAN, shown in Figure 2.5. The sensitivity of the WUR of each sensor or actuator node is defined as the minimum input amplitude or power which can be correctly received by the WUR. This directly determines the maximum distance between the gateway and the sensor node at which a Wake-Up operation can occur, D in Figure 2.5. Any node placed further from the gateway than distance D cannot be activated by the gateway itself through a Wake-Up message. As will be clearer in the following, there is a direct trade-off between sensitivity and power consumption, that is the higher is the power consumption the lower the sensitivity

and the higher the wake-up distance. Actually, WURs usually belong to one of three categories according to their power consumption and maximum wake-up distance:

- *short range WURs*: their wake-up distance is below 1 m and they are completely passive.
- *medium range WURs*: their wake-up distance is in the order of tens or hundreds of meters and their power consumption in the nanoWatt range.
- *long range WURs*: their wake-up distance is above 1 km and their power consumption is in the μW range.

Addressing

Addressing is another key aspect for WURs. A node with addressing capabilities can detect whether a Wake-Up message is directed to itself or to another node within the network. If a network was composed of nodes with no addressing capabilities, a Wake-Up message directed to any of them would cause all nodes within the maximum wake-up distance to be activated. With reference to Figure 2.5, all nodes circled in red would wake up as a result. This would be very inconvenient, making addressing capabilities prime.

Chapter 3

Analog Front-End for a medium range Wake-Up Radio

Some of the material reported in this chapter is reused from [1] and [2] (©2018, 2020 IEEE), in agreement with IEEE copyright policy on theses and dissertations.

This chapter shows the structure of a generic WUR, divided into AFE and BBL, and presents the specific structure of the AFE within a WUR targeting medium range applications, as defined in Section 2.2.2. Previous literature is then discussed, both regarding discrete component and integrated architectures, starting from the Figures of Merit (FoM) used to assess their performance. Finally, the three prototypes we designed are presented.

3.1 Structure of a medium range WUR AFE

A generic WUR is usually composed of two subsystems: the AFE and the BBL. The AFE has the purpose of turning the input OOK-modulated signal into a stream of bits, whereas the BBL compares the received sequence to the actual address of the specific node, previously stored. In the event of a received sequence matching the address, the BBL generates a Wake-Up impulse. Therefore, it is the BBL which implements the addressing capabilities of the WUR.

The categories mentioned in Section 2.2.2 are becoming more and more blurry. Yet, the choice of what wake-up distance range to design for is critical for the structure of the WUR itself. For instance, an amplification of the incoming signal at Radio Frequency (RF), i.e. by the use of a Low Noise Amplifier (LNA) tuned to the carrier RF frequency, which is typical in receivers, entails a higher power consumption but makes long wake-up distances possible. Therefore, it is recommended if the target is the design of a long range WUR. On the other hand, simpler demodulation methods, such as by the use of EDs, are usually adopted when targeting medium ranges [7]. Moreover, circuits most likely need to be designed in subthreshold to comply with nanoWatt power requirements. This is also useful for demodulation, which usually leverages the second-order non-linearities in the current of a MOSFET in subthreshold. The scope of this thesis is exclusively medium range WURs, leading

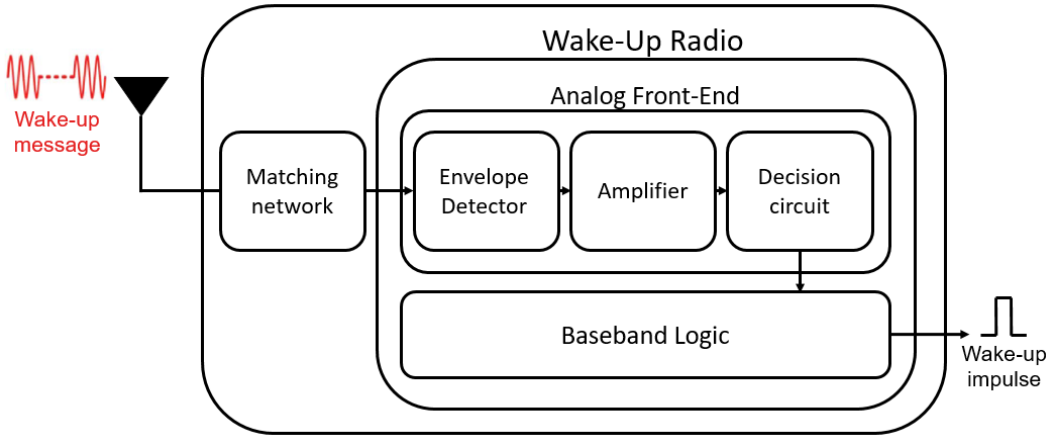

Figure 3.1: Architecture of a generic medium range WUR, with focus on the AFE.

to the absence of an LNA in favor of EDs operating in subthreshold, as will be further discussed later.
Figure 3.1 shows the main building blocks of a medium range WUR, with focus on the AFE. First, though not strictly included in the WUR, an external matching network is usually employed with the aim of matching the input impedance of the chip to the impedance of the antenna. As a beneficial side effect, the incoming OOK-modulated signal is also amplified. Then, as just mentioned, an LNA providing further amplification to the RF signal before demodulation is not present. The subsequent block, strictly necessary for the operation of the WUR, is a rectifier, an ED, which extracts the envelope of the incoming OOK-modulated signal. It is followed by a block performing amplification at baseband. Finally, a decision circuit digitizes the extracted envelope, i.e. turns it into a stream of bits. This represents the output of the AFE, which is then fed to the BBL.

3.2 Figures of Merit

The performances of a WUR can be assessed making use of FoMs. Obviously, the BBL may also contribute to the performances of a WUR by increasing the tolerance to errors on single bits, thus allowing a sensitivity improvement. Even though the scope of this thesis only includes the AFE, results taking into account the contribution of the BBL will also be included.
The performances of a WUR are defined by a few parameters:

- sensitivity, i.e. the minimum input power which is correctly detected by the WUR,
- bitrate, i.e. the number of symbols which can be received in a unit time,
- DC power consumption.

FoMs are a useful tool to assess and compare the performances of different WUR architectures taking these parameters into account. An additional parameter may be die area, but it usually does not appear in FoMs as it is not related to circuit operation.
The most widely used FoM was defined in [8]:

$$FoM_1(dB) = -P_{SEN}(dBm) + 5\log BW - 10\log\frac{P_{dc}}{1mW} \tag{3.1}$$

where $P_{SEN}(dBm)$ is the sensitivity in dBm, BW is the bitrate and P_{dc} is the overall DC power consumption and assuming demodulation to occur through a non-linear process by the use of an ED. If demodulation was performed through a linear operation, that is with a mixer, $10\log BW$ would appear instead of $5\log BW$. Additional details are provided in Section 4.2. As mentioned in Section 3.1, demodulation is usually performed leveraging second-order non-linearities, and is thus a non-linear process, when dealing with medium range applications, leading to the use of Equation 3.1. Though defined for EDs, this FoM is commonly used to assess the performances of the whole AFE under consideration.
Actually, FoM_1 may not be considered a completely fair FoM since it is not an invariant for a given ED architecture, which makes it not entirely suitable to compare the performances of different architectures, as shown in Appendix A. Moreover, it indirectly takes into account the supply voltage by considering the power consumption, which gives architectures with low supply voltages an advantage. This is the reason why another FoM was defined in [2] as follows:

$$FoM_2(dB) = -P_{SEN}(dBm) + 5\log BW - 5\log I_{bias} \tag{3.2}$$

Again, more details are provided in Appendix A.

3.3 Previous literature dealing with the WUR AFE

Due to the interest in WUR and asynchronous protocols, several WUR hardware designs have recently been presented both using discrete components and integrated solutions. They will be detailed in the following.

3.3.1 Discrete component architecture

The current state-of-the-art discrete component WUR architecture, shown in Figure 3.2, is the one presented in [6] and [1]. Its AFE receives OOK-modulated 16-bit messages and is preceded by a matching network composed of a commercial lumped inductance and capacitance (C = 7 pF, L = 22 nH). Rectification, i.e. envelope extraction, is passive and performed by a voltage doubler composed of Schottky diodes. The threshold voltage for the subsequent comparator is set by low-pass filtering the envelope itself to extract its average. Finally, the only active section of the circuit is

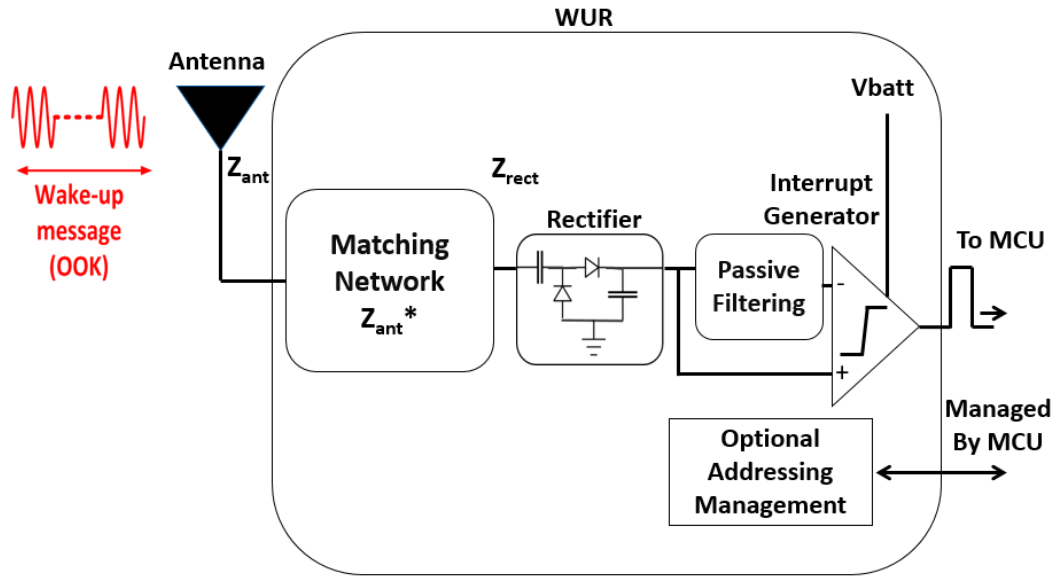

Figure 3.2: State-of-the-art discrete component WUR [6].

the comparator, chosen among three different commercial discrete components.
The threshold generation process is effective because it provides a threshold which adapts to the input signal amplitude but it is functional only if no long sequences of equal bits are received, i.e. a long sequence of "ones" would push the threshold dangerously close to the input amplitude whereas a long sequence of "zeros" would push it towards ground. As described in Section 4.6.2, a similar effect occurs in the second proposed prototype, leading to the presence of a maximum packet length. Another important drawback in the threshold generation process of the architecture from [6] is the fact that the values of the R and C performing passive filtering are heavily dependent on the operating bitrate.
As for sensitivity, the offset of the comparator is actually one of its two main limiting factors, to the point that in the event of using a comparator with the wrong offset polarity the system cannot work, also posing a yield issue. The other important limiting factor for sensitivity is comparator noise. Both can lead to errors in the comparison, especially when the adaptive threshold gets close to ground or the input amplitude. This system achieves a -55-dBm sensitivity with a 1.2-μW power consumption, corresponding to a 99.2-dB FoM.

3.3.2 Integrated architectures

This section focuses on the main integrated solutions that have been presented recently in literature, with which the prototypes presented in this thesis can be put in direct comparison. Different aspects will be highlighted, both on the block level and on the circuit level. Tables 3.1 and 3.2 summarize the performances of such solutions in chronological order, also including the results of the experimental characterization of the first two proposed prototypes along with the simulated results of

Table 3.1: Performance comparison between the proposed prototypes and several prototypes in literature, part 1.

	Proposed prototype 1	Proposed prototype 2	Proposed prototype 3 (simulated)	[9], 2012	[10], 2013	[11], 2016	[6], 2016	[12], 2017
Temperature range [°C]	-20 : 60	unknown	-40 : 120	unknown	unknown	unknown	unknown	unknown
Carrier frequency [MHz]	771	433	433	915	915	2400	868	2400
Matching gain [dB]	12	17.3*	14.7	12	5	5	unknown	12.5
Bitrate [$kbit/s$]	2	1	0.5	100	12.5	8.192	1	200
Sensitivity at 27°C [dBm]	-46[a]	-52.3*,[b]	-63.1[b]	-41[a]	-43[b]	-56.5[b]	-55[b]	-50[a]
Supply voltage [V]	1.2	0.6	0.6	1.2	1.2/0.5	1/0.5	1.2	0.8
Power at 27°C [nW]	36	12.9**	16.8	98	116	236	1200	4500
Correlator	no	yes	yes	no	yes	yes	yes	no
Code length	N.A.	63	256	N.A.	31	31	16	N.A.
Technological node [nm]	90	90	90	130	130	65 low-power	N.A.	180
Die area [mm^2]	0.7	0.37	1	0.03	0.35	2.25	N.A.	0.456
External components	LC	LC*	LC	LC	LC	LC	N.A.	LC
FoM_1 [dB]	106.9	116.2	124.3	106.1	102.8	112.3	99.2	100
FoM_2 [dB]	100.1	105.6	114.3	101.4	unknown	unknown	100	102.8

* LC matching network simulated: sensitivity projected as sum of measured AFE sensitivity and simulated matching network gain
** Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time
[a] Sensitivity defined through 10^{-3} Bit Error Rate
[b] Sensitivity defined through 10^{-3} Missed Detection Rate

Table 3.2: Performance comparison between the proposed prototypes and several prototypes in literature, part 2.

	[13], 2017	[8], 2018	[14], 2019	[15], 2019	[16], 2019	[17], 2020	[18], 2020	[19], 2020	[20], 2021
Temperature range [$^\circ C$]	unknown	unknown	unknown	unknown	unknown	unknown	-10 : 40	unknown	-40 : 85
Carrier frequency [MHz]	403	113.5	434	750	152	450.8	9000	430	915
Matching gain [dB]	18.5	25	26	13	27	26	13.5	unknown	unknown
Bitrate [$kbit/s$]	0.3	0.3	0.1	200	0.2	0.1	0.033	42	16.4
Sensitivity at 27$^\circ C$ [dBm]	-63.8[b]	-69[b]	-79.2[b]	-50[b]	-76[b]	-80.9[b]	-69.5[b]	-108[b]	-70.2[c]
Supply voltage [V]	0.4	0.4	0.4	0.4	1/0.6	0.54	0.4	unknown	unknown
Power at 27$^\circ C$ [nW]	4.5	4.5	0.42	1486[**]	7.4	40	7.3	41000	2700
Correlator	yes	yes	yes	yes	yes	yes, analog	yes	yes	yes
Code length	32	16	11	40	8	11	18	31	unknown
Technological node [nm]	180	180	65 low-power	65	130	65 low-power	65/180[***]	65 low-power	65
Die area [mm^2]	4.5 (estimated)	4.5 (estimated)	0.05 (estimated)	0.75	1.95	0.15	0.14/ 4.5 cm^2[***]	3	0.33
External components	transformer	transformer	LC	L	LC	LC	transformer	LC	LC
FoM$_1$ [dB]	129.7	134.9	153	104.8	138.8	134.9	128.5	145	117
FoM$_2$ [dB]	115.9	121.1	134.1	103.7	unknown	126.6	115.8	unknown	unknown

[**] Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time
[***] Antenna included
[b] Sensitivity defined through 10^{-3} Missed Detection Rate
[c] Sensitivity defined through 10^{-3} Packet Error Rate

the third one for easier comparison. This is an anticipation of the results achieved by all three proposed prototypes, which will be further discussed in Sections 4.5 and 4.8 and in Chapter 5.
As Tables 3.1 and 3.2 show, it is possible to define sensitivity in different ways. In particular, when the BBL is not present, i.e. no correlation between the incoming bitstream and the node address occurs, sensitivity can only be defined through the Bit Error Rate (BER), that is the percentage of wrong bits the AFE outputs over the total number of bits the AFE receives.
When a BBL is present and correlation occurs, sensitivity can be defined through the Missed Detection Rate (MDR), i.e. the percentage of missed Wake-Up impulses over the total number of packets corresponding to the node address the WUR receives. In this case, it is possible to tolerate errors in the received packet and still have the BBL recognize the packet as correct.
Alternatively, sensitivity may also be defined through the Packet Error Rate (PER), i.e. the percentage of correct packets the WUR recognizes over the total number of packets, whether corresponding to the node address or not. Finally, another metric which can be evaluated once a BBL is present is the False Alarm Rate (FAR), which corresponds to the number of Wake-Up impulses caused by noise occurring in a preset time interval.

Clocked and clockless architectures

Most integrated architectures known from literature [8], [10]-[14], [16], [18]-[20] include an oscillator which is constantly on, as shown in Figure 3.3. The clock is essential in the BBL for data extraction, but it has strong implications on the design of the AFE as well. As a matter of fact, it may be beneficial for the AFE, where it allows the use of latched comparators or the processing of signals in the discrete-time charge domain, which would not be feasible otherwise. Indeed, latched comparators are typically faster than conventional ones under equal conditions in terms of current thanks to positive feedback and also allow the implementation of offset cancellation techniques [21]. However, the clock consumes power, even if calibrated low-power RC oscillator designs are adopted.
Conversely, in clockless architectures, shown in Figure 3.4, the oscillator is turned on only upon reception of the first bit of an incoming message and turned off once the whole message has been received, as recently proposed also in [15]. The circuit operates in two phases. During the first phase, the AFE is the only active section. When recognition of the first bit of the message takes place, which occurs at the first transition of the AFE output signal, the second phase starts: the oscillator and the BBL are turned on and the incoming bitstream is compared with the stored address. If the transition causing the second phase to start is spurious (solely due to noise), the oscillator and the BBL turn back off after a predefined time interval, pushing the circuit back into the first phase.
Adopting the clockless approach is beneficial in terms of power. As a matter of fact,

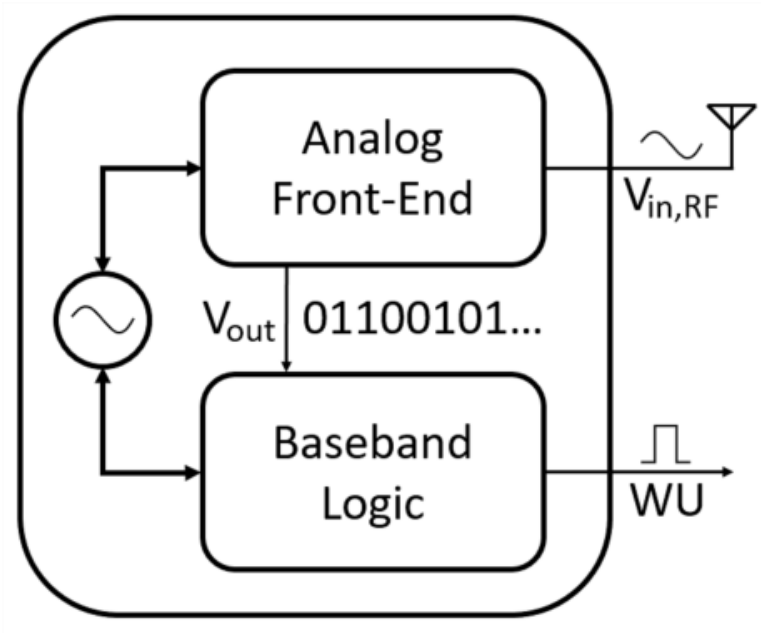

Figure 3.3: Generic structure of a clocked WUR. Picture reused from [2].

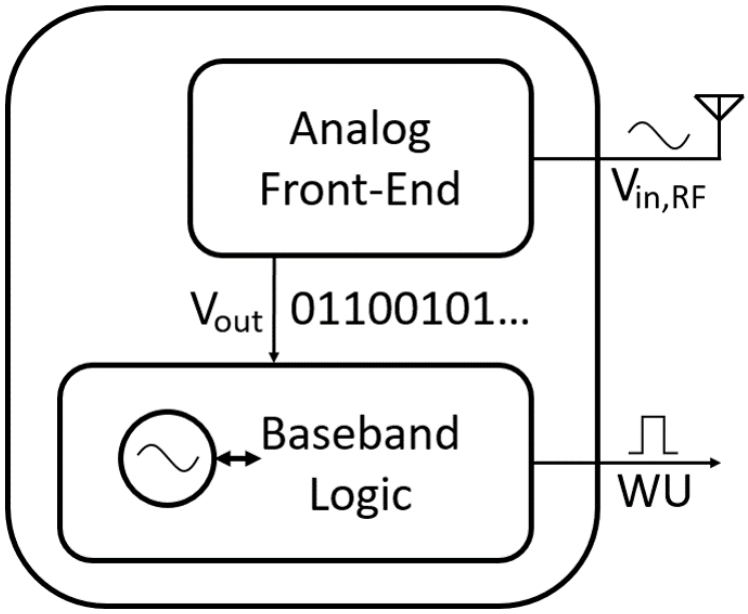

Figure 3.4: Generic structure of a clockless WUR. Picture reused from [2].

the average power consumption for clocked architectures corresponds to the sum of the consumption of the different blocks. On the other hand, for clockless architectures, the average power consumption can be calculated as [15]:

$$P_{avg} = \frac{P_1 t_1 + P_2 t_2}{t_1 + t_2}, \tag{3.3}$$

where P_1, t_1, P_2, t_2 are the power and the ON time of the two phases. This implies that if the specific application is characterized by long idle periods, i.e. $\frac{t_2}{t_1} \ll 1$, in a clockless AFE architecture $P_{avg} \sim P_1$. This means the power consumption of the AFE strongly dominates the overall power consumption, whereas the BBL does not count in terms of power, but its functionalities are available nonetheless.

Demodulation mechanism based on second-order non-linearities

As mentioned in Section 3.1, in medium range WURs demodulation cannot be carried out by using conventional receiver architectures, characterized by mixers, due to power constraints. Due to the same constraints, circuits must be biased with nanoAmps, thus operating in subthreshold. Therefore, most architectures exploit the non-linearities generated in a MOSFET in subthreshold [8] - [20]. The basic working principle is presented in the following. Let's consider the basic circuit in Figure 3.5 as an example, where I_{bias} is such that the MOSFET is operating in the subthreshold region. The input signal is OOK-modulated:

$$v_{RF}(t) = v_M(t) \cos(\omega t) \tag{3.4}$$

Since the MOSFET is in subthreshold, its drain current is an exponential function of the input voltage:

$$I = I_{bias}\, e^{-\frac{v_{RF}}{nV_T}}. \tag{3.5}$$

where n is the subthreshold ideality factor and V_T is the thermal voltage. The exponential can be written as a Taylor series and truncating at the second order has proven enough to accurately describe the behavior of the circuit with simple math. This yields

$$I \approx I_{bias}\left[1 - \frac{v_{RF}}{nV_T} + \frac{1}{2}\left(\frac{v_{RF}}{nV_T}\right)^2\right]. \tag{3.6}$$

Substituting Equation 3.4 in Equation 3.6 and assuming the RF components, i.e. components at the carrier frequency and twice the carrier frequency, to be way out-of-band and get filtered out immediately, the expression of the drain current only includes the low frequency components of Equation 3.6:

$$I = I_{bias}\left[1 + \frac{v_M^2(t)}{4n^2V_T^2}\right] = I_{bias} + \Delta I(t), \tag{3.7}$$

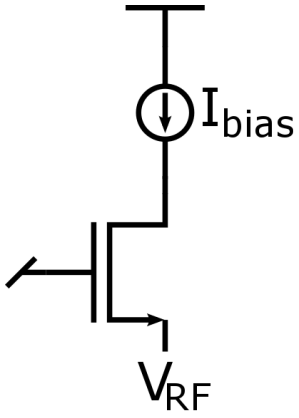

Figure 3.5: Simplified schematic to illustrate the demodulation mechanism based on the second-order non-linearities which characterize subthreshold currents, usually employed in medium range WURs.

where ΔI is an envelope dependent part and can be equivalently attributed to an effective low frequency input voltage source $v_{IN}(t)$ replacing $v_{RF}(t)$):

$$v_{IN}(t) \equiv \frac{\Delta I(t)}{g_m} = \frac{v_M^2(t)}{4nV_T}. \tag{3.8}$$

where the transconductance of the MOSFET in subthreshold $g_m = I_{bias}/4nV_T$ has been used.

In this example a common-gate circuit was employed, but it works exactly the same if the input signal is applied to the gate, leaving the circuit in a common source configuration. Also, it works for MOSFETs biased at no static current other than their leakage.

Sensitivity-bitrate trade-off

As already explained, most medium range architectures feature no amplification at RF, that is a power consuming LNA is not present and the first stage within the IC is the ED. Since the matching network is passive, it generates virtually no noise. Therefore, most of the noise affecting the AFE output signal is due to the ED, unless the ED provides very little gain to the rectified signal. This situation is usually avoided because it would pose stringent constraints on the following stages, i.e. the decision circuit.

Therefore, WUR sensitivity is ultimately limited by the noise which the ED adds to the signal. The bandwidth of the ED is the frequency range over which ED noise power spectral density gets integrated to determine the integrated noise it adds to the signal:

$$v_{RMS,out}^2 = \int_0^{BW} \frac{4k_BT\gamma}{g_m} df = \frac{4k_BT\gamma}{g_m} BW. \tag{3.9}$$

Equation 3.9 holds assuming all noise in the ED to be white, which is usually not completely true but a useful approximation.

This poses a direct trade-off between sensitivity and the bandwidth of the ED itself, which also determines the maximum bitrate at which the system can operate. As a matter of fact, many solutions proposed in literature have extremely low bitrates, lower than 0.5 kbit/s and at times even lower than 0.1 kbit/s, in order to have very low integrated noise added to the signal by the ED [8], [13], [14], [16], [17]. Of course, a low bitrate results in a long latency, at times even in the order of 500 ms [18], which is not acceptable for all applications. Sometimes, the reduction of the code length is taken as a countermeasure [14], [16], [17] but latency may still be an issue if bitrates below 1 kbit/s are employed.

Active and passive EDs

As explained in Section 3.1, the demodulation mechanism of an incoming OOK-modulated signal leverages the second-order non-linearities of MOSFETs in sub-threshold. They can either be biased at very low currents, such as nanoAmpères, or biased with no static current at all. This happens in active and passive EDs, respectively.

Both approaches have recently been widely used. For instance, [10], [11], [14]-[17], [19], [20] employ passive EDs. They are typically Dickson charge pumps composed of diode-connected MOSFETs and capacitances, like the one shown in Figure 3.6. The main advantage given by such a choice is the inherent absence of flicker noise, which, unlike the channel thermal noise, is related to the transconductance of the MOSFET used for demodulation. No static current entails virtually a null transconductance, leaving the ED noise power spectral density with zero flicker noise.

Dickson charge pumps have also appeared in literature in the pseudo-balun configuration, shown in Figure 3.7. This allows the conversion gain of the ED to be multiplied by two, as the output signal becomes differential, without any additional power consumption [18]. However, this ED is followed by a fully-differential baseband amplifier, thus requiring a common-mode feedback loop.

Yet, passive EDs typically have two major drawbacks. As will be shown in more detail in Chapter 5, they may cause a significant propagation delay. This is related to the fact that the ED output signal amplitude is directly proportional to the number of diode stages, which makes it common to use several tens of stages in order to maximize the output signal amplitude. In turn, this increases the propagation delay and directly limits the achievable ED bandwidth, thus the bitrate. Actually, in [20] this issue has been circumvented by using two Dickson charge pumps in parallel instead of one with an equivalent number of stages to speed up the operation of the ED. The output signals of the chains in parallel are then fed to a charge summation and amplification circuit operating in the discrete-time domain, ultimately saving even more static power by leveraging the clocked nature of the system.

The second major drawback of passive EDs is related to capacitances C_C and C_L in

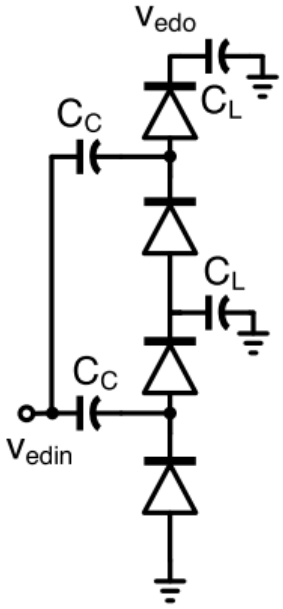

Figure 3.6: A Dickson charge pump, a typical passive ED [14].

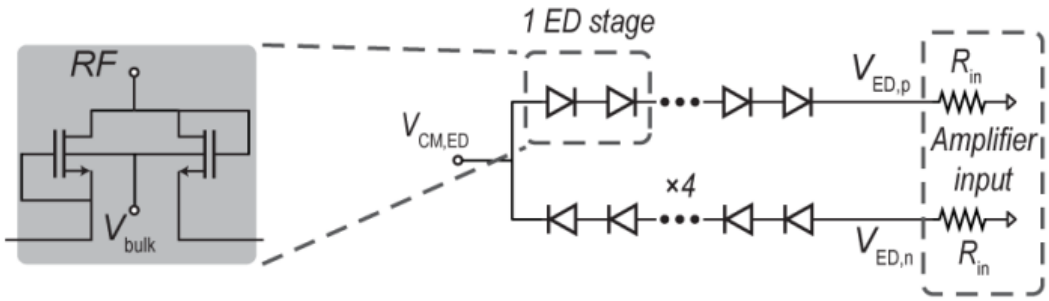

Figure 3.7: A pseudo-balun Dickson charge pump [18].

Figure 3.6. Their value is strictly related to the input RF frequency, since they act as coupling capacitors and RF filtering capacitors, respectively. This means the operation of the ED itself is dependent on the input RF frequency. The consequences include the fact that simulations during the design phase must be carried out using the exact RF frequency for which the design is being performed, resulting in long and difficult simulations, and the RF frequency may not be changed easily once the design is finished. Even more importantly, the second drawback is a much higher ED input capacitance, caused by the fact that capacitances C_C directly load the input. This may cause a dramatic drop in the gain of the matching network, which sees the ED as a load as well.

On the other hand, active EDs are biased at a non-zero current, which makes it possible to perform both demodulation and amplification of the extracted envelope in a single block, as will be shown in Chapter 4. Their main disadvantage is indeed that of consuming static power, which in turn causes them to inherently have flicker noise, unless countermeasures are taken. As an advantage, they typically allow

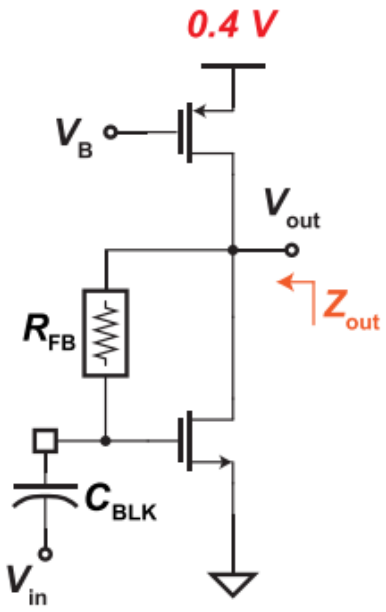

Figure 3.8: An example of active ED [8].

higher bitrates and have superior flexibility. For instance it may be possible to make their bitrate programmable by changing their bias current. Still, the relationship between sensitivity and current is quite unfavorable, i.e. sensitivity can be improved by increasing bias current by it takes a lot of current. This idea is explained in Section 4.2. An example of active ED is shown in Figure 3.8. Moreover, simulations can be carried out using any RF frequency, as long as it is high enough for non-linearities to be solicited. As a result, simulations may be much faster and lighter than those for designing passive EDs. Finally, the input capacitance of a passive ED only corresponds to a gate or a drain capacitance, depending on which terminal the input signal is applied to, which typically results in a higher matching gain.

Decision circuits

As for decision circuits, latched comparators are the most common choice within clocked solutions [8], [10]-[14], [16]-[18]. As mentioned in Section 3.3.2, the presence of an always-on clock often makes latched comparators the preferred choice for their higher speed and absence of static power consumption. Some solutions even employ Successive Approximation Register Analog-to-Digital Converters [19], [20] or a complex structure composed of a couple of oscillating Voltage-Controlled Delay Lines, one of which acts as a reference, to realize a voltage-to-time conversion and compare two phases instead of two voltages and ultimately perform time-encoded integration [14].
Clocked systems often implement offset control as well. For instance, [14] uses a Delay-Locked Loop to set the phase difference to zero when no input signal is applied. Also, [16] implements an automatic offset control algorithm, which not only

corrects mismatches in the circuit but also compensates for environmental noise. It is based on the detection of RF "ones" followed by "zeros", which suggests the presence of noise, and results in an increase in the comparator threshold. It acts on the preamplifier of the comparator by removing or adding MOSFETs in parallel to each of the preamplifier branches. A similar approach is also adopted in [10]. Actually, different solutions adopt the same approach but externally by hand [8], [9], [13]. Finally, [18] even exploits its clocked nature to implement an auto-zeroing amplifier before the comparator.

Robustness

Robustness against Process, Voltage and Temperture (PVT) variations is often considered when designing the AFE of a WUR. As a matter of fact, circuits may be dealing with extremely low signal amplitudes and even small errors in the setting of DC operating points may lead to significant gain loss. This would ultimately degrade sensitivity. Additional branches mentioned in Section 3.3.2 and used for offset control may also be employed to compensate for PVT variations. Some solutions also feature dedicated circuits or techniques to automatically compensate for temperature change. For instance, [18] uses a Complementary To Absolute Temperature (CTAT) voltage reference, whose DC output voltage is then amplified and inverted, to ultimately bias the bulk of the diode-connected MOSFETs in its passive ED. This way, the output impedance of the ED, linked to the threshold voltage of the diodes, can be kept roughly constant and the ED bandwidth remains constant as well.

Interference rejection

Interference rejection is an issue when dealing with all WURs. Indeed, medium range WURs may be particularly subject to desensitization due to the reception of interferences because of the lack of selectivity of the IC itself. As a matter of fact, the matching network is the only element which filters unwanted frequencies.
This is not always the case though. An example of an effective interference rejection mechanism is [19], which is based on the fact that the intermodulation of two OOK-modulated signals offset by Δf produces energy at an intermediate frequency Δf, which can be isolated through a band-pass (BP) filter. Instead, both tones along with the interferer get filtered out. This idea is shown in Figure 3.9. Still, this solution requires precisely two-tone OOK signalling, which is a more complicated signalling method than the common 1-tone OOK modulation.

Other examples of interference rejection approaches in WURs include the analog correlator, instead of the typical digital one in [17] and the automatic control of the comparator threshold in [10] and [16] presented in Section 3.3.2, which may increase the threshold of the comparator in the presence of an interferer, trading sensitivity for interference rejection.

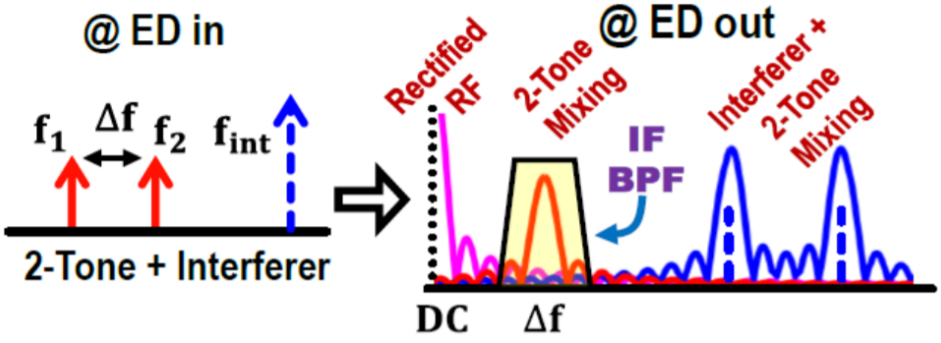

Figure 3.9: Advantage of two-tone RF OOK for interference rejection [19].

Technologies

Choosing the suitable technology may significantly help reach design targets when dealing with WURs. For instance, the use of less scaled technological nodes, such as the 180 or 130 nm one, typically allows a reduction in leakage currents thanks to thicker gate oxide layers [8]-[10], [12], [13]. On the other hand, more scaled technological nodes, such as the 65 nm one, make the reduction of die area possible [15], [20]. Some solutions take advantage of both aspects by choosing a more scaled technological node with a low-power (LP) option [11], [14], [17], [19]. These solutions usually include circuits biased with currents as low as a few picoAmpères. Another solution even uses two different technologies to optimize separately the RF part, which includes the ED and baseband amplifier, for low ED input parasitics and the baseband part, including the comparator and the BBL, for leakage minimization [18].

Matching networks

The matching network is an essential part of a WUR. Its main purpose is to match the output impedance of the antenna to the input impedance of the chip. This avoids reflections of the incoming signal at the input of the chip itself, which would obviously cause loss of signal and sensitivity degradation. Moreover, it provides gain without consuming any power, which is particularly important when active EDs are used. Finally, if no dedicated circuits are present, it is also the only portion filtering out interferences.
Unfortunately, typically the higher is the carrier frequency, the lower is the maximum achievable gain, unless special countermeasures are taken and careful matching network-ED codesign is carried out [18]. Most solutions employing sub-GHz carrier frequencies use lumped component matching networks, composed of off-chip inductors and capacitors [9]-[12], [14]-[17], [20]. Typically, the higher is the carrier frequency, the smaller the inductances and capacitances needed, also leading to smaller area occupation. Still, some solutions with sub-GHz carrier frequencies use customized transformers for higher gain [8], [13].

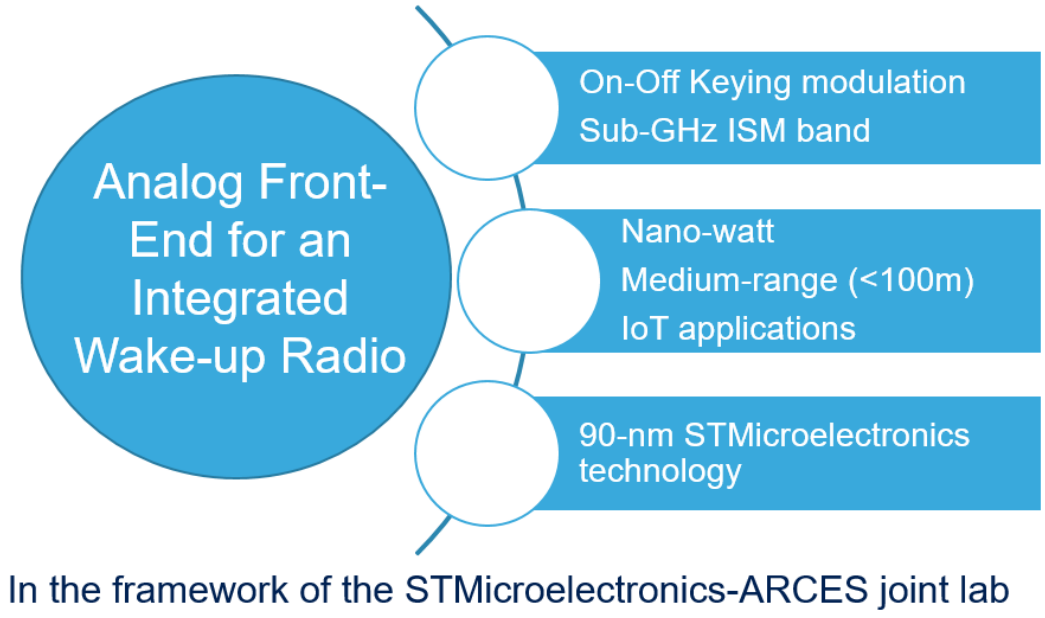

Figure 3.10: Specifications for the Ph.D. activity presented in this thesis.

3.4 Proposed architectures

The target of the Ph.D. research activity presented in this thesis was to design and implement the AFE of an integrated WUR with the following features:

- OOK-modulated input signals
- carrier RF frequency in the Sub-GHz ISM band
- medium range WUR, i.e. suitable for IoT applications with a wake-up distance within 100 meters
- nanoWatt power consumption
- 90-nm STMicroelectronics technology

As an additional specification, a 1-kbit/s bitrate was chosen. Finally, two of the protoypes were implemented using a BCD technology to make integration in a node with actuation capabilities possible. Instead, the third prototype was implemented in a CMOS technology.
The design of the proposed prototypes presented in this thesis followed a detailed study of the discrete component architecture presented in Figure 3.2 and ultimately aimed at addressing its yield and flexibility issues and improving its sensitivity and power consumption.
Since all prototypes have a power consumption in the order of nWs, all circuits operate in subthreshold, which also allowed the rectification process to leverage the second-order non-linearities of a subthreshold current. The clockless approach was

chosen to avoid dealing with an always-on power consuming clock and another Ph.D. student took care of the study and design of the BBL. As for the ED, both the active and passive approaches were explored. In particular, the active one is expected to be more robust thanks to the self-biasing of the circuit, whereas the passive one is expected to yield a better sensitivity. Due to the clockless nature of the design, a latched comparator was not feasible. PVT variations were considered during the design and lead to the implementation of blocks or techniques dedicated to reducing performance degradation. Instead, no particular interference rejection techniques were implemented. The matching network remained the only portion introducing selectivity.

Three prototypes have been designed and implemented:

- *The first prototype*, taped out and measured in 2018 and 2019, only includes the AFE. Its supply voltage is 1.2 V. Its ED and baseband amplifier are implemented as a single active block, which has a BP response. The decision circuit is a Schmitt trigger. A dedicated block generates a temperature independent DC input voltage for the Schmitt trigger and its two thresholds. This prototype requires the use of the Manchester coding to avoid multiple errors but the capacitance in the ED is integrated. This prototype is presented in [2] and [1] and its ED has been filed as a patent [22]. More details are provided in Chapter 4.

- *The second prototype*, taped out in 2019 and measured in 2020, includes both the AFE and the BBL. Its supply voltage is 0.6 V. Its ED and baseband amplifier are implemented as a single active block, which has a low-pass (LP) response. The decision circuit is a conventional comparator. Its effective threshold is set by forcing the bulk voltages of the two transistors in its input pair. This prototype does not require the use of the Manchester coding but the capacitance in the ED is way too big to be integrated. This prototype is presented in [23] and [24]. More details are provided in Chapter 4.

- *The third prototype*, taped out in 2021, includes both the AFE and the BBL as well. Its supply voltage is 0.6 V. Its ED and baseband amplifier are implemented as two separate blocks and are differential. The ED is passive and includes an additional circuit to keep sensitivity constant with temperature. The decision circuit is a conventional comparator. Its effective threshold is determined by filtering the output of the amplifier through an RC and an offset cancellation circuit. Moreover, it does not require off-chip components and its sensitivity is expected to improve with respect to the previous prototypes. This prototype is currently under fabrication. More details are provided in Chapter 5

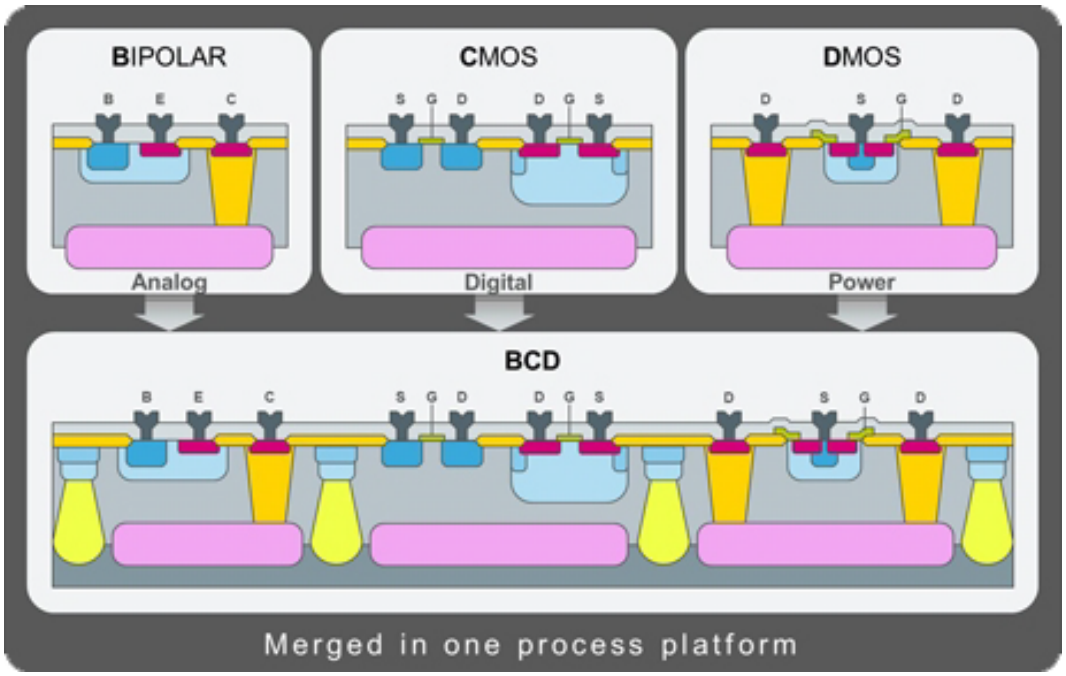

Figure 3.11: BCD technology [25].

3.4.1 Smart power technology

As mentioned in Chapter 1, a generic IoT node is usually composed of different subsystems, each performing a precise task, which makes it an actual SoC. Namely, the sensing subsystem gathers data about the external world, whereas the elaboration subsystem processes the data coming from the sensor. The node is also equipped with a communication subsystem able to connect to other nodes, with which to share its data and receive instructions. Finally, it includes an actuator to modify the external world in a useful way. A power management section may be present as well. Examples of sensors are those measuring environmental parameters such as temperature and moisture and chemical parameters such as the concentration of a gas. Examples of actuators are a valve to open or close or a motor to turn on or off. Smart power technology, namely BCD, is especially convenient for SoCs [26] since it allows the integration of bipolar, CMOS and DMOS devices in one process platform, as shown in Figure 3.11. This in turn allows the whole node to be integrated on a single die. In particular, bipolars may be used for analog functions, especially precision ones, CMOS for digital and possibly also analog functions and DMOS for power functions [27], typically found in the actuation and possibly power management sections. This approach produces desirable improvements in terms of area, complexity and reliability at the cost of a non-optimized CMOS section in terms of scaling and a low number of metal layers available for routing. However, the features of a CMOS subsystem implemented in a BCD technology are currently suitable to efficiently integrate digital circuits using finite state machines as a computational base model. Since digital circuits of such a low complexity are the only ones required in such a

SoC, BCD is ultimately the ideal technology for sensor and actuator IoT nodes. This is the reason why two of the WUR prototypes described in this thesis have been implemented in a 90-nm STMicroelectronics BCD technology, available thanks to the collaboration in the framework of the STMicroelectronics - ARCES joint lab.

Chapter 4

Active Envelope Detector Implementations

Some of the material reported in this chapter is reused from [1], [2] (©2018, 2020 IEEE), in agreement with IEEE copyright policy on theses and dissertations, and [23].

This chapter presents the features and advantages of active EDs, which are employed in the first two implemented prototypes presented in this thesis.
The first prototype was taped out in 2018 and features an active band-pass (BP) ED. Due to the peaks appearing at the output of the ED, a hysteretic comparator is needed. Therefore, the ED is AC-coupled to a Schmitt trigger, whose input DC voltage and two threshold voltages are generated by a dedicated block. This prototype only includes the AFE, partly designed by another Ph.D. student [28]. The BBL was not implemented, so the prototype has no addressing capabilities. The operation of this prototype is thus limited to the digitization of the incoming OOK-modulated signal.
The second prototype was taped out in 2019 and features an active low-pass (LP) ED, which results in a bigger capacitance that cannot be integrated, but also in the generation of a threshold voltage within the ED itself. This leads to the use of a standard comparator instead of a trigger and a simpler biasing block. This prototype includes both the AFE and the BBL, designed by another Ph.D. student. The BBL employs Clock and Data Recovery (CDR) circuits with the purpose of correctly synchronizing the bitstream which the AFE outputs to the internal oscillator. This allows to perform a digital correlation between the received bitstream and the address of the specific node, the codeword. If they match, a Wake-Up impulse is generated to turn on the rest of the node. Moreover, it allows to receive packets with a longer length than just the codeword.

4.1 Introduction

As anticipated in Section 3.4, the design of the prototypes presented in this thesis was preceded by a detailed study of the discrete-component architecture in [6], Figure 3.2, to address its issues. For instance, its yield issues are mainly due to the offset

of its discrete off-the-shelf comparator. As a matter of fact, in the event of using a comparator featuring an offset with the wrong polarity, the whole system cannot work properly. Moreover, its flexibility is limited due to the fact that the values for the RC filter are strictly related to the bitrate. The time constant characterizing the RC filter needs to correspond to a few bit-times, so that at the end of the preamble the threshold is ready but it does not get reset to either the voltage corresponding to a "0" or that corresponding to a "1" until a few equal consecutive bits have been received. Actually, this sets the maximum number of equal consecutive bits which can be correctly received. Additionally, sensitivity is ultimately limited by the smaller between the noise and the offset of the comparator in [6], making offset reduction prime.
The integration of this architecture as is, that is with no baseband amplification before the comparator, was considered first. As just mentioned, comparator offset reduction is prime for sensitivity optimization when using this apporach. In ICs this can be accomplished by enlarging MOSFETs, especially the ones belonging to the first comparator stage. As a result, the speed of the circuit can be impaired, actually posing an offset-speed trade-off. Adding baseband amplification of the extracted envelope before the comparator helps reduce offset constraints on the comparator itself. Finally, the choice was to implement both envelope extraction and baseband amplification of the extracted envelope in a single block for better flexibility and robustness, as will be clearer in Section 4.2.

4.2 Active Envelope Detectors

As just mentioned, both the first and the second prototype feature an active ED, a BP and a LP one, respectively. The implications of using each of the two have been thoroughly investigated.
Active EDs both perform OOK demodulation and the baseband amplification of the extracted envelope. An example of active ED is presented in [8], which is reported in Figure 4.1 for convenience.

The main advantage of this circuit is its simplicity and low current, as it is composed of only one branch. As previously discussed, all transistors operate in the subthreshold region and in particular transistor M_1 operates in order to exploit the exponential $I - V_{GS}$ characteristic. Provided capacitor C is large enough to effectively ground the gate for all frequencies within the signal band, M_1 works as a common gate amplifier as well as a rectifier. Resistance R provides the bias to the gate of M_1, which is therefore self-biased. However, R would need to be prohibitively large, in the order of hundreds of MΩ, in order to provide enough voltage gain, given the small transconductance value which can be achieved with currents in the order of nanoAmpères. In [8] this problem is circumvented by replacing R with MOSFETs in the OFF state, which in turn requires 8 bits for calibration against process variations. Conversely, in this prototype follower M_3, as shown in Figure 4.2, was added

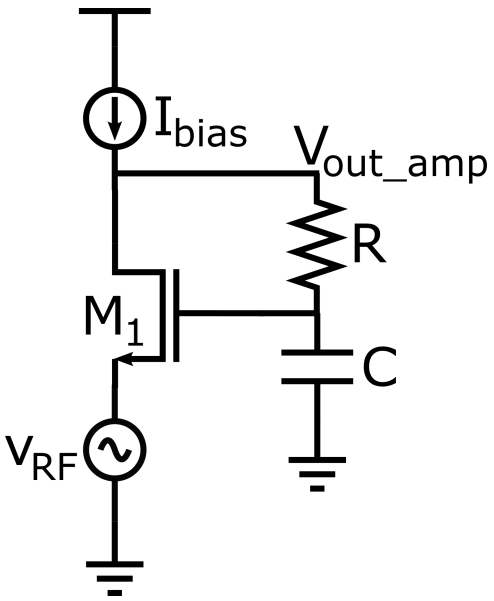

Figure 4.1: Basic topology of an active ED [8]. Picture reused from [2].

in order to be able to implement R as an actual resistor without compromising the baseband gain and reducing the need for calibration against process variations. This comes at the expense of an additional current consuming branch. The architecture of the ED in Figure 4.2 has been filed as a patent [22].

The demodulation principle is the one described in Section 3.3.2. The MOSFET performing the actual demodulation is M_1, therefore I_{bias} and g_m in Equations 3.5-3.8 refer to M_1. Referring to Equation 3.8, $v_{IN} = v_M^2(t)/4nV_T$ is the effective low frequency input voltage source which replaces v_{RF}. The expression of the approximated small-signal ED response to $v_{IN}(t)$ can be derived by circuit inspection. Assuming unity gain for the follower and the two poles to be well separated for simplicity,

$$\frac{v_{out\_amp}}{v_{IN}}(s) = \frac{1 + sR^*C}{\left(1 + s\frac{R^*C}{g_{m1}R_o}\right)(1 + sR_oC_o)}, \tag{4.1}$$

where $R^* = R + R_{OUTFOLL}$. $R_{OUTFOLL}$ is the output resistance of the follower stage, whereas $R_o = r_{out1}$ is the output resistance at the drain of M_1, where $r_{out1} = 1/\lambda I_{bias}$ and assuming the output resistance of current generator I_{bias} to be much higher than r_{out1}. Finally, C_o is the total capacitance seen between the drain of M_1 and ground, due to the current generator, M_3 and parasitics. In particular, the location of the two poles in Equation 4.1 can be found using the open circuit time constant method as follows:

- *The dominant pole* is due to capacitance C and its time constant is $\tau = R_1C$. Resistance R_1 can be calculated as

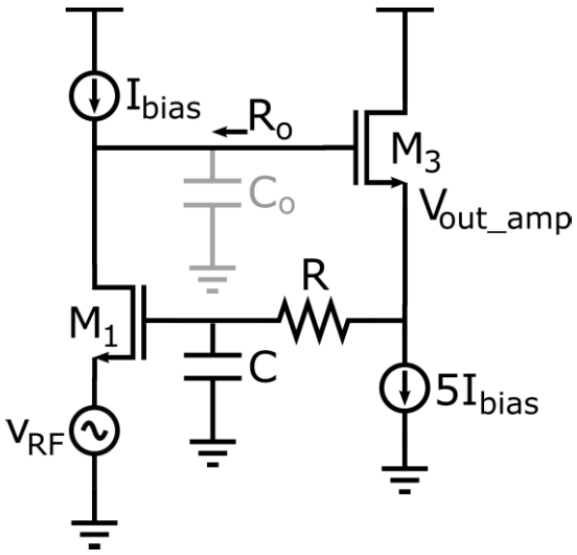

Figure 4.2: Schematic of the proposed BP ED. C_o accounts for all parasitic capacitances at the drain of M_1. Picture reused from [2].

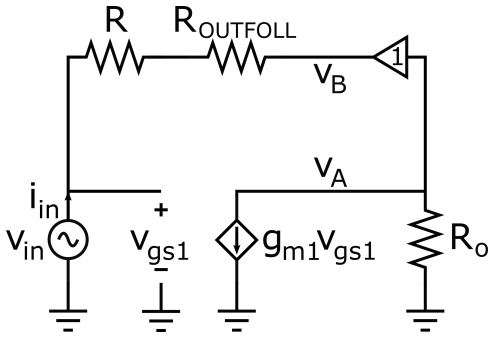

Figure 4.3: Small-signal equivalent circuit of the proposed ED, in Figure 4.2, to determine the location of the dominant pole, due to capacitance C, using the open circuit time constant method. The follower stage is assumed to have unity gain for simplicity.

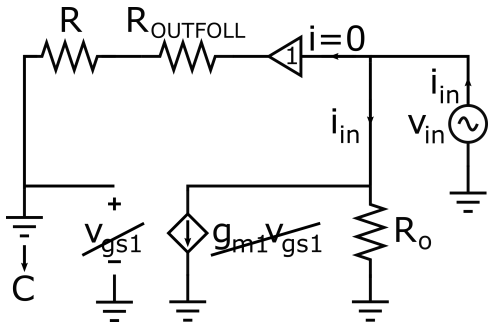

Figure 4.4: Small-signal equivalent circuit of the proposed ED, in Figure 4.2, to determine the location of the second pole, due to capacitance C_o, using the open circuit time constant method. Assuming the two poles to be well separated, capacitance C is considered a short.

$$R_1 = \frac{v_{in}}{i_{in}} \tag{4.2}$$

using the small-signal equivalent circuit in Figure 4.3. Since $v_{gs1} = v_{in}$ and the gate of the follower is supposed to absorb no current,

$$v_A = -g_{m1} R_o v_{in}. \tag{4.3}$$

Since unity gain is assumed for the follower, $v_B = v_A$. This leads to

$$i_{in} = \frac{v_{in} + g_{m1} R_o v_{in}}{R^*}. \tag{4.4}$$

Assuming $g_{m1} R_o \gg 1$ and substituting Equation 4.4 into Equation 4.2,

$$R_1 \sim \frac{v_{in}}{\frac{g_{m1} R_o v_{in}}{R^*}} = \frac{R^*}{g_{m1} R_o}, \tag{4.5}$$

leading to

$$\tau = \frac{R^* C}{g_{m1} R_o}. \tag{4.6}$$

- *The second pole* is due to capacitance C_o and its time constant is $\tau = R_2 C_o$. Resistance R_2 can be calculated as

$$R_2 = \frac{v_{in}}{i_{in}} \tag{4.7}$$

using the small-signal equivalent circuit in Figure 4.4. Capacitance C is considered a short circuit thanks to the assumption of well separated poles, which leads to $v_{gs1} = 0$ and $g_{m1} v_{gs1} = 0$. Again the gate of the follower is supposed to absorb no current, which yields

$$v_{in} = R_o i_{in}. \tag{4.8}$$

Substituting Equation 4.8 into Equation 4.7,

$$R_2 = R_o, \tag{4.9}$$

leading to

$$\tau = R_o C_o. \tag{4.10}$$

Equation 4.1, which describes the response shown in Figure 4.5, shows that the low-frequency gain is close to one as a result of the negative feedback. For angular

frequencies above $1/R^*C$, capacitance C progressively shorts the gate of M_1 to signal ground and the gain increases up to its peak value, equal to $g_{m1}R_o$ for the case of well separated poles, when M_1 behaves as a common gate stage. The second pole is due to the capacitance and the resistance at the drain of M_1, R_o and C_o. AC noise simulations show the main noise contributors are M_1 and current generator I_{bias} and the output noise shaping function is similar to Equation 4.1.

Comparison between low-pass and band-pass ED

Most active EDs presented in literature have a response that can be classified somewhat improperly as LP, such as that in [8] and that in the second prototype presented in this thesis. Such a response can be obtained by fixing the R^*C time constant to a large enough value so as to place the first pole $g_{m1}R_o/R^*C$ at a frequency lower than the inverse of the total transmission time, in order to have an actual LP time response across all received bits. This type of frequency response is shown in Figure 4.5 in orange, while the corresponding time domain response is reported again in orange in Figure 4.7, assuming the input RF signal is the one portrayed in Figure 4.6.
On the other hand, the second pole $1/R_oC_o$ must be placed at a frequency close to the maximum bitrate that the system is supposed to handle. In general, higher bitrates require larger drain currents through M_1, given the dependence of R_o on the bias current. In such LP solutions, a limitation may be represented by the low-frequency flicker noise, which undergoes no filtering. As a matter of fact, its reduction may require an increase in the area of M_3 and especially M_1, which may in turn result in the need for an increase in the bias current in order to maintain operation at the chosen bitrate despite higher junction capacitances.

On the other hand, in the first prototype presented in this thesis an ED with a truly BP response has been implemented. This is obtained by placing the first pole close to the second pole, as shown in Figure 4.5 in blue. The BP time response, shown in Figure 4.7 again in blue, is characterized by a positive/negative output pulse corresponding to each falling/rising edge transition in the input envelope, which means that envelope transitions are recognized rather than the envelope itself. Actually, in this case the ED is an edge detector. For the correct operation of the BP edge detector, the first pole must be high enough to guarantee the relaxation of the output signal to its quiescent value within one bit-time, so as to avoid inter-bit interference. This also improves flicker noise suppression without the need to increase transistor sizes, circumventing any possible speed issues, and more importantly leads to a smaller value of capacitance C, making it eligible for integration. On the other hand, too small values for C need to be avoided as too small of an R^*C constant would reduce the amplitude of the output pulse, as shown in Figure 4.10.

Assuming the use of an ideal OOK modulation, the pulse peak amplitude V_{OUTM}, calculated from Equations 4.1 and 3.8,

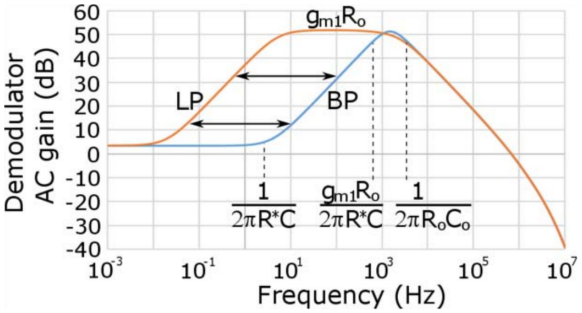

Figure 4.5: Simulated frequency response of the ED v_{out_amp}/v_{IN} for $C = 100\ nF$ (orange) and $C = 500\ pF$ (blue). Picture reused from [2].

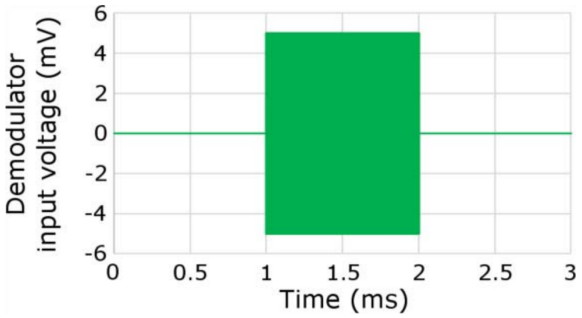

Figure 4.6: Input RF signal. Picture reused from [2].

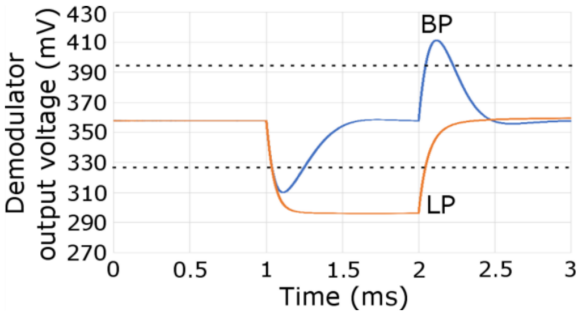

Figure 4.7: Simulated time-domain response of the ED v_{out_amp} for $C = 100\ nF$ (orange) and $C = 500\ pF$ (blue). The thresholds are also visible in dotted line. Picture reused from [2].

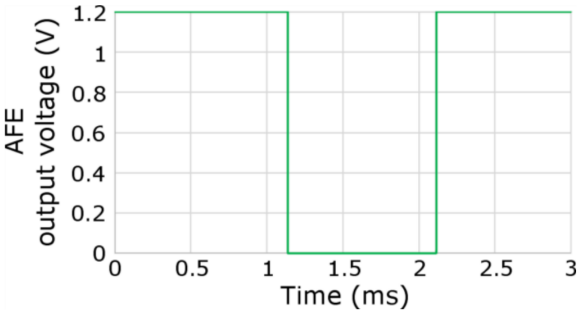

Figure 4.8: Simulated time-domain response of the decision circuit in both cases. Picture reused from [2].

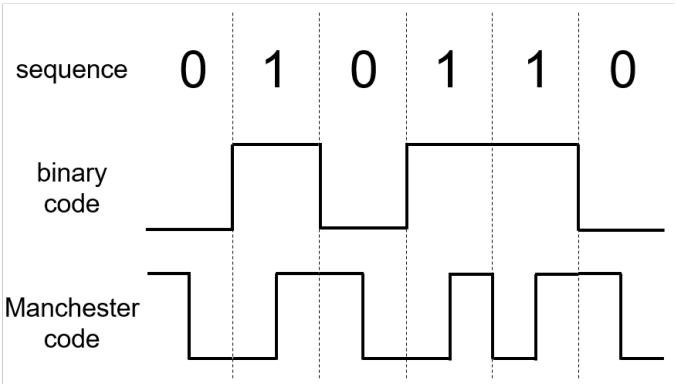

Figure 4.9: Binary vs. Manchester coded sequence. Picture reused from [2].

$$V_{OUTM} \propto \frac{g_{m1} R_o}{4nV_T} v_M^2, \tag{4.11}$$

where v_M is the amplitude of the input envelope. An ED rectification gain can also be defined directly from Equation 4.11 as

$$G_{RECT} \equiv \frac{V_{OUTM}}{v_M} = \frac{g_{m1} R_o v_M}{4nV_T} \propto v_M. \tag{4.12}$$

G_{RECT} is itself proportional to v_M because second-order non-linearities are being leveraged.

Since in the BP case the time-domain response of the ED is the one shown in blue in Figure 4.7, a comparator with hysteresis, i.e. a trigger with two thresholds, is well needed to recover an envelope signal with full digital level. Its behavior is illustrated by the waveforms in Figures 4.7 and 4.8. The two thresholds can be adjusted depending on signal amplitude. A possible drawback of the trigger is that, in case of sequences of identical transmitted bits, the occurrence of a detection error due to noise at the beginning of the sequence causes the misdetection of the entire sequence. This problem can be minimized by using codes with frequent transitions, such as Manchester coding, where a logic "0" is represented by a high-to-low transition in the bit-time while a logic "1" by a low-to-high transition (see Figure 4.9). On the other hand, for the same bitrate, this requires to double the ED bandwidth, with a penalty in terms of current.

A BP and a LP version of the ED in Figure 4.2 were designed for comparison, assuming the same bias current for the two designs. Sensitivity P_{SEN} and bitrate BW are taken into account to calculate the normalized sensitivity $P_{SEN,norm}$ [15], which is hereby used to assess the performances of the EDs only,

$$P_{SEN,norm} = -P_{SEN}\,(dBm) + 5 \log BW, \tag{4.13}$$

where P_{SEN} is defined as the value of the voltage v_M, measured in dBm on a reference 50-Ω resistance) corresponding to 11-dB SNR at the ED output [29, p. 500].

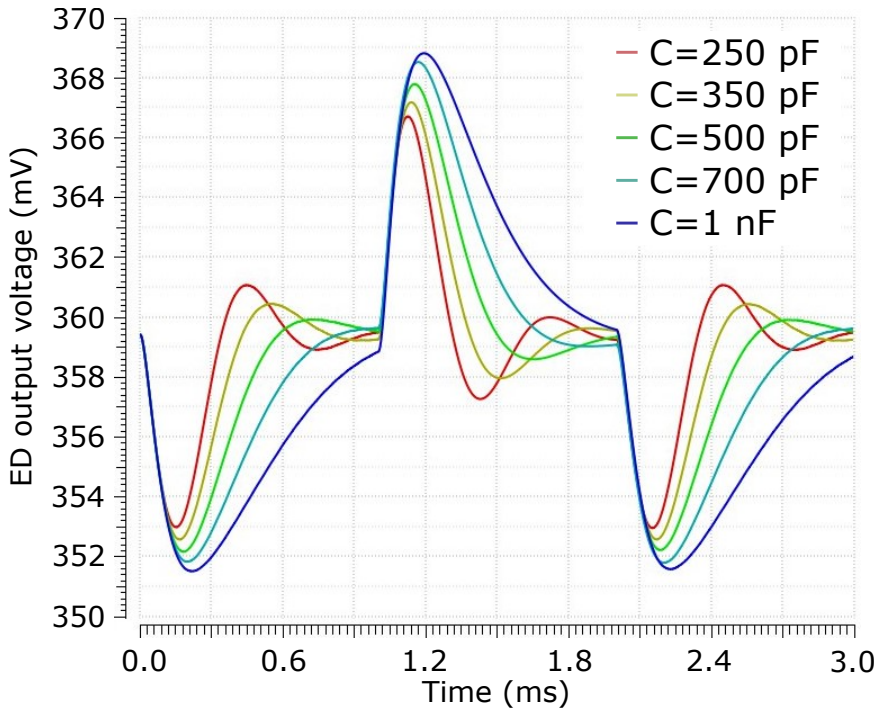

Figure 4.10: ED output signal of the circuit in Figure 4.2 for different values of C with a 5-mV RF input. Picture reused from [2].

More information on the operational approach used to make these estimations later. Results are based on simulations. Since both designs have the same total current (6 nA, of which $I_{bias} = 1$ nA and 5 nA in the follower, with a 1.2-V supply voltage), the second pole is at the same frequency, which has been fixed at roughly 1 kHz, that is the bitrate of the LP ED, while the BP detector uses a Manchester coded signal with 0.5 kbit/s rate. In both designs R = 75 MΩ. The choice of $C = 500\ pF$ for the BP version was driven by the results in Figure 4.10, which shows the ED output signal of the circuit in Figure 4.2 for different values of C with a 5-mV RF input.

As stated earlier in this section, Equation 4.1 was derived modeling the follower as a unity gain block. Actually, it is apparent from Figure 4.10 that this is far from true, since a ringing response appears when the pole due to C is too close to the second one, which occurs when C is too small. The analysis of the ratio between the output peak amplitude and the total Root Mean Square (RMS) output noise voltage indicates it is advantageous to maximize C, in the range between 250 pF and 1 nF, for maximum Signal-to-Noise Ratio (SNR). Yet, too large values for C must be avoided to allow the output voltage to relax to its quiescent value within one bit-time, preventing the risk of inter-bit interference. On the other hand, in the LP case $C = 100\ nF$ was chosen. A 500-pF capacitance could most likely be integrated. For instance, with the 90-nm BCD technology employed, an integrated capacitor of this value has an estimated area of $190\ \mu m \times 190\ \mu m$ (1.2-V poly-on-pplus capacitor). On the other hand, a 100-nF capacitance would certainly have to be off-chip.

The results of the comparison between the BP and the LP active ED are shown in

Table 4.1: Performance comparison between band-pass and low-pass active EDs. Reused from [2]

Response shape	Capacitance	Power (P_{dc}) [nW]	Bitrate (BW) [kbit/s]	Estimated sensitivity (P_{SEN}) [dBm]	Estimated normalized sensitivity ($P_{SEN,norm}$) [dB]
BP	500 pF	7.2	0.5	-34.9	-48.4
LP	100 nF	7.2	1	-34.2	-49.2

Table 4.2: Performance comparison between low and high throughput band-pass ED [2].

Response shape	Capacitance	Power (P_{dc}) [nW]	Bitrate (BW) [kbit/s]	Estimated sensitivity (P_{sen}) [dBm]	FoM_1 [dB]	FoM_2 [dB]
Low thr.	500 pF	7.2	0.5	-34.9	99.8	89.5
High thr.	5 pF	720	50	-34.9	89.8	89.5

Table 4.1. They indicate that there is no significant difference between the two cases in terms of estimated normalized sensitivity, the halving of the bitrate in the BP solution being on average compensated by its better sensitivity, mainly due to a more efficient suppression of flicker noise.

Bitrate programmability

The topology in Figure 4.2 can also be easily customized according to the throughput required for the chosen application, both for a BP and for a LP ED. For instance, the frequency of the second pole is multiplied by a factor k whenever I_{bias} is multiplied by k, whereas the frequencies of the zero and the first pole are multiplied by the same factor k when capacitance C is divided by k. So this topology can be implemented for both low and high throughput applications, where the high throughput requires a higher current and a smaller capacitance than the low throughput one, making the integration of C even more practical. A comparison between the performances of a low and a high throughput BP active ED is shown in Table 4.2. As explained in Section 3.2, FoM_1 changes although the topology of the circuit is the same in the two cases and is therefore not suitable to make perfectly fair comparisons, whereas FoM_2 does not.

Sensitivity estimation through simulations and sensitivity-current relationship

It is very convenient to be able to easily estimate sensitivity based on simple simulations during the design phase. Actually, this task can be performed thanks to the formula drawn from System 4.14. The first equation is Equation 4.12, defining the rectification gain of the ED. This gain also allows to link the ED RMS output noise

voltage $v_{RMS,out}$ to the ED RMS output noise voltage referred to the input $v_{RMSeq,in}$, resulting in the second equation of the system. Finally, the input SNR may be defined as in the third equation.

$$\begin{cases} G_{RECT} \equiv \frac{V_{OUTM}}{v_M} = \frac{g_{m1} R_o v_M}{4nV_T} \\ \frac{v^2_{RMS,out}}{G^2_{RECT}(v_M)} = v^2_{RMSeq,in} \\ SNR_{in} = \frac{v_M}{v_{RMSeq,in}} \end{cases} \tag{4.14}$$

Solving the system yields

$$v_M = \sqrt{\frac{SNR_{in} \times v_{RMS,out} \times 4nV_T}{g_m R_o}}, \tag{4.15}$$

which allows to easily estimate sensitivity. As a matter of fact, it is enough to set a desired input SNR and run an AC noise simulation to obtain the values of $g_{m1}R_o$ and $v_{RMS,out}$. The fact that $v_M \propto \sqrt{v_{RMS,out}}$ results in $5 \log BW$ appearing in Equation 3.1. If a mixer was used, i.e. demodulation was a linear operation, it would occur that $v_M \propto v_{RMS,out}$, which would result in $10 \log BW$ instead.

Equation 4.15 also gives more insight on the change of performances of the ED, Figure 4.2, due to programmability. As described in Section 4.2, current I_{bias} may be multiplied by a factor k to increase the operating bitrate of the ED. The value of capacitance may be multiplied by k accordingly to keep the shape of the response. In this case, g_{m1} gets multiplied by k as well, whereas AC gain $g_{m1}R_o$ remains constant, since $R_o \propto 1/I_{bias}$. This results in $v_{RMS,out}$ to stay the same, yielding no change in the minimum detectable value for v_M. On the other hand, a current increase may be used to improve this. In this case, C_o has to be increased accordingly through the addition of a parallel integrated capacitance in order to keep the same bitrate, leaving the same value for C as well. This again causes a multiplication of g_{m1} without any changes in $g_{m1}R_o$, but $v_{RMS,out}$ gets divided by $\sqrt{k}$. Finally, this results in a division of the minimum detectable value for v_M by $\sqrt[4]{k}$, yielding

$$v_{M,min} \propto \sqrt[4]{I_{bias}}, \tag{4.16}$$

where $v_{M,min}$ corresponds to the voltage sensitivity.
Equation 4.15 has been used to make all sensitivity estimations reported in Chapter 4, considering an 11-dB SNR at the ED output [29, p. 500].

4.3 First prototype: active band-pass Envelope Detector

As previously mentioned, the first prototype presented in this thesis features an active BP ED. The structure of the overall AFE implemented in the first prototype is shown in Figure 4.11. First of all, a clockless approach was chosen for the reasons explained in Section 3.3.2. As mentioned, the use of an active ED allows demodulation and baseband amplification of the extracted envelope to be implemented in a

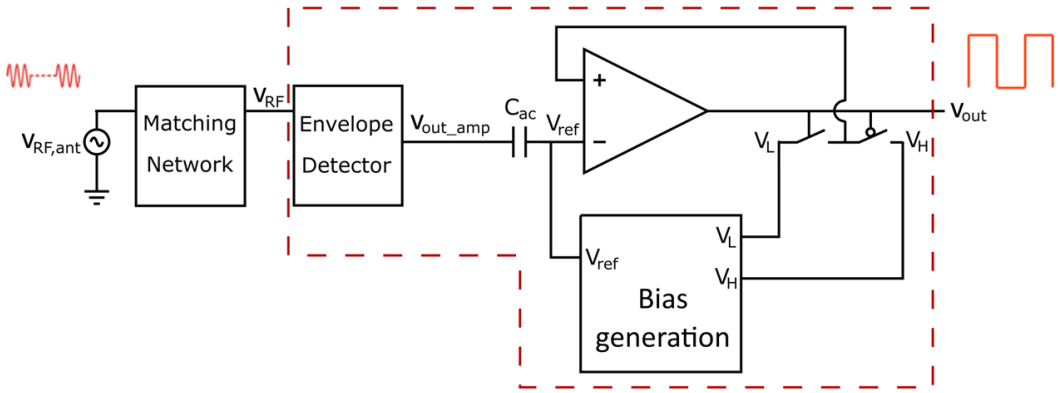

Figure 4.11: High level block diagram of the first prototype. Picture reused from [2].

single stage, thus yielding better flexibility and robustness. The baseband amplification added before the comparator allows to break its offset-speed trade-off: since the amplification preceding the comparator is high enough, comparator noise and offset become negligible and transistor sizes in the comparator itself can become smaller, thus causing lower capacitive loading of the output of the first stage of the comparator. Active EDs also make bitrate programmability possible as shown in the previous section. Finally, a BP ED was implemented to avoid the need for off-chip capacitors, at the cost of using a Schmitt trigger. As matter of fact, this kind of decision circuit requires two thresholds due to its hysteresis and the use of the Manchester coding to prevent errors in the event of long sequences of equal consecutive bits. Finally, in this particular chip, the BBL is not present but the addition of one was planned for a following prototype.

4.3.1 Envelope Detector

The schematic of the actual implemented ED is shown in Figure 4.12. It differs from the one already discussed in Figure 4.2 essentially for the addition of cascode transistor M_2 to achieve higher gain. Consistently, diode D_3, implemented though a diode-connected MOSFET, was also added in the feedback branch to properly bias M_1 and M_2. The frequency response is still given by Figure 4.1, provided the resistance of D_3 is added in series to the output resistance of the follower $R_{OUTFOLL}$ and $R_o = r_{out1} g_{m2} r_{out2}$, where g_{m2} and r_{out2} are the transconductance and the output resistance of M_2, respectively. To properly bias the gate of M_2, the chain of diodes $D_1 - D_2$, implemented through diode connected MOSFETs, was employed. It should be noticed that the addition of cascode transistor M_2 is made possible by the choice of keeping the supply voltage at 1.2 V. For lower supply voltages the topology in Figure 4.2 is more suitable.

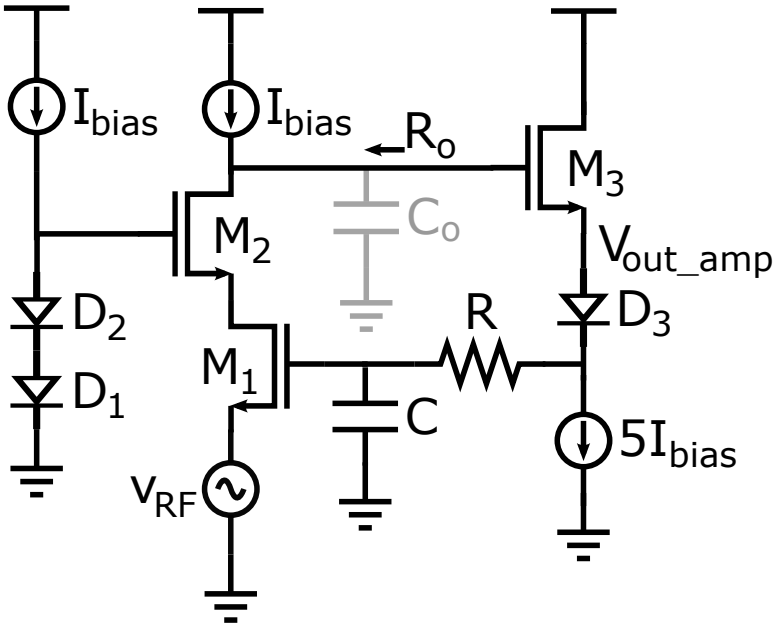

Figure 4.12: Schematic of the proposed ED. C_o accounts for all parasitic capacitances at the drain of M_2. I_{bias} is copied from M_1 in Figure 4.14. Diodes are implemented through diode-connected MOSFETs. Picture reused from [2].

4.3.2 Schmitt Trigger

As shown in Figure 4.5, the ED generates pulses at each change in the input bit, which makes a hysteretic comparator necessary, i.e. a Schmitt trigger.

AC coupling between BP ED and trigger

Due to the BP response of the ED, it is natural to employ an AC-coupling between the ED itself and the trigger, as it involves no loss of information. This leaves the possibility of biasing the trigger independently of the output DC voltage of the ED. In Section 4.3.3, a circuit is proposed for the generation of the two threshold voltages as well as the input DC voltage for the trigger independently of the detector output and robustly against PVT variations, without the need for additional circuits for calibration. Thresholds only need to be trimmed according to the signal amplitude at the ED input, which in turn depends on the distance between nodes and the application.

Trigger structure

The complete transistor-level schematic of the trigger, designed by another Ph.D. student, is shown in Figure 4.13. It consists of a differential amplifier and two switches (M_H and M_L) driven by the amplifier output voltage v_{out}, which feed either a low (V_L) or a high (V_H) voltage level back to the amplifier input within a bistable loop with positive feedback. The V_H and V_L levels are programmable in order to

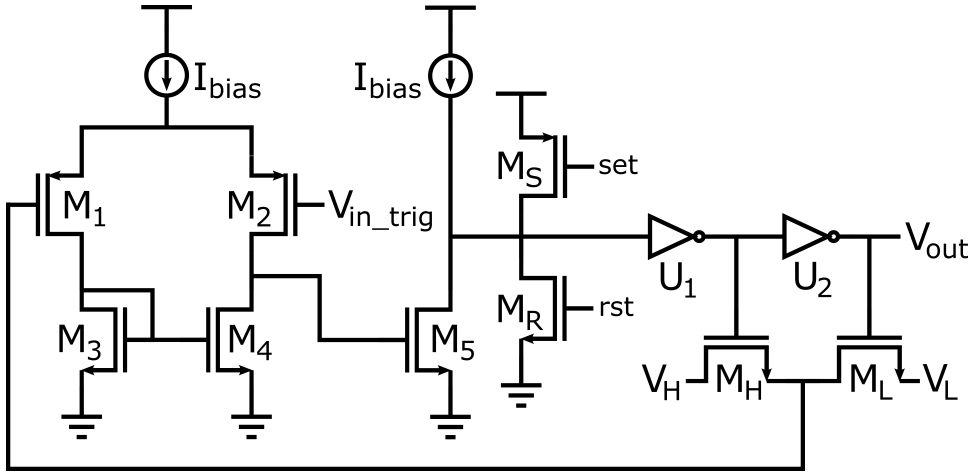

Figure 4.13: Schematic of the trigger including the two-stage differential amplifier and the switches for the threshold voltage feedback. I_{bias} is copied from M_1 in Figure 4.14. Picture reused from [2].

tailor noise and disturbance rejection according to the input signal amplitude. M_S and M_R allow to preset the trigger to the desired state at the end of each received message and prepare it for the reception of the following one, i.e. initialize the circuit. A calibration routine is needed upon node deployment (that is, once the input power has been set by the transmitter output power and the distance between the transmitter and the WUR) and/or at scheduled times to set the proper values for V_H and V_L. In particular, the calibration criterion may be the maximization of the Bit Error Rate (BER) or the Packet Error Rate. As mentioned in Section 3.3.2, [16] implements a similar threshold calibration loop, which sets a comparator offset to a level which provides a desired False Positive Rate.
The main target of the trigger design was to keep its input offset and noise voltage at values that do not limit the overall receiver sensitivity, which is ultimately determined only by the noise of the first stage, i.e. the ED. This goal is not difficult to achieve due to the large gain obtained with the ED, as shown in Section 4.5.1. Positive feedback guarantees fast enough switching despite transistors in the first stage having a relatively large area for low offset.

4.3.3 Biasing Circuit

As mentioned in Section 4.3.2, the aim of the biasing block is to provide the trigger DC input voltage and thresholds as robustly as possible against process and temperature variations. As shown hereafter, this block requires a non-negligible bias current. In an ultra-low power design, it may be reduced to a current generator with a programmable resistance ladder: while this would certainly reduce the number of branches, thus the consumed current, it would not provide any temperature compensation. If such an approach was chosen, the programmable resistance ladder would have to be recalibrated at every change in the operating temperature. Instead,

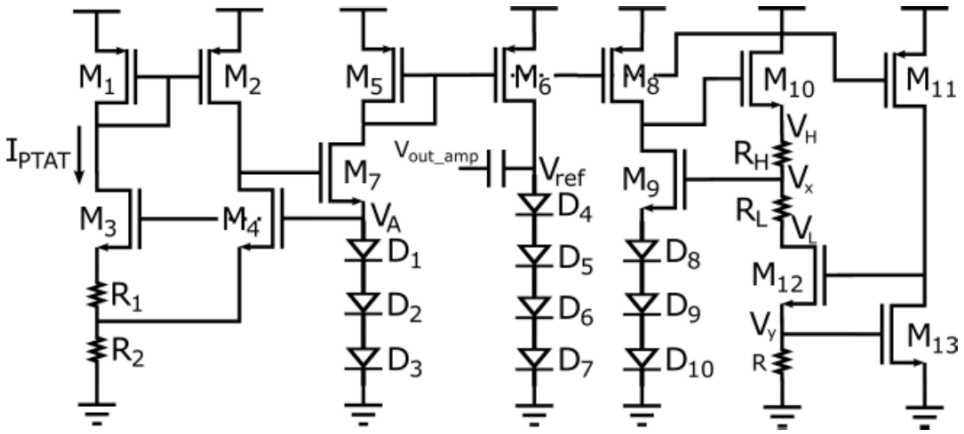

Figure 4.14: Schematic of the temperature independent bias generation circuit. All mirrors are 1:1. Diodes are implemented through diode-connected MOSFETs. M_9 and M_{13} have the same dimensions as D_1-D_{10}. Picture reused from [2].

the architectural choice here is to limit the need for calibration only to the start-up phase.

The biasing block schematic, reported in Figure 4.14, aims at providing V_{ref}, V_H and V_L. As usual, all transistors are operated in the subthreshold regime, thus exhibiting an exponential characteristic. The reference voltage generator cell composed of M_1-M_4 and R_1-R_2 [30] is the core of the block, whose output is a fairly temperature-stable V_A. V_A is the gate voltage of M_3 and M_4 and is given by the sum of the Proportional To Absolute Temperature (PTAT) voltage drop across resistance R_2 and of V_{GS4}, which has a negative temperature coefficient instead. V_A is then applied to a chain of diode-connected transistors and the resulting current is copied by M_5-M_6 into the chain of MOSFET diodes D_4-D_7, which have the same dimensions as D_1-D_3, thus the same V_{GS}. Therefore, $V_{ref} \propto V_A$. This allows the generation of a temperature-compensated input bias voltage V_{ref} for the trigger, which is AC-coupled to the ED output. The same current is also copied by M_8 into M_9 and D_8-D_{10}. Under the assumption that V_{GS9} is roughly equal to the voltage across D_1-D_{10}, voltage $V_x \cong V_{ref}$ is generated as well. This is the case since V_{DS9} is large enough to force M_9 to work in saturation, therefore resulting in the current of M_9 being a function of V_{GS9} only. This is proven right by simulations, as the simulated V_{GS} of the diodes is roughly 97 mV while the simulated V_{GS9} is roughly 95.5 mV at room temperature. Finally, M_{11} forces into M_{13} the same current which flows in D_1-D_{10}, which, under the assumption that M_{13} is in saturation as well, sets $V_{GS13} = V_y \cong V_{ref}/4$, making it a temperature independent voltage. The local feedback between M_9 and M_{10} and between M_{12} and M_{13} further enforces this condition. V_y is then applied to resistance R, so as to obtain

$$V_H - V_{ref} \cong V_H - V_x = (R_H/R)\, V_y \tag{4.17}$$

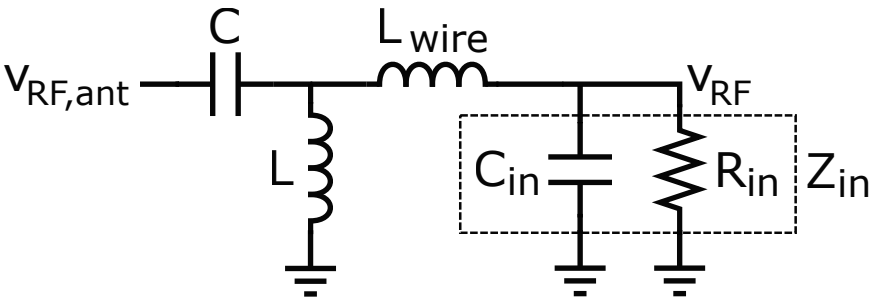

Figure 4.15: Schematic of the proposed matching network. Z_{in} indicates the input impedance of the ED. Picture reused from [2].

$$V_{ref} - V_L \cong V_x - V_L = (R_L/R)\, V_y \tag{4.18}$$

Resistor values, in particular $R_1 = 10\ M\Omega$, $R_2 = 17\ M\Omega$, and transistor sizes, in particular $(W/L)_3 = 104.2$ and $(W/L)_4 = 4.2$ and $(W/L)_1 = (W/L)_2$, were chosen so as to minimize the dependence of V_A, thus V_{ref}, on temperature as well as to keep the current consumption of the block in the nA range at room temperature, while the number of diodes was selected for $V_{ref} = V_{DD}/2$.
This approach has two main drawbacks, i.e. rising consumption with temperature and large resistor area. Consumption increases with temperature because currents rise in most of the branches in the block, in particular currents through M_1 and M_2 rise linearly, through M_5, M_6, M_8, M_{11} rise exponentially and the current though M_{10} follows the dependence on temperature of resistors, i.e. is roughly constant. It should be noticed that the exponential increase of current with temperature is due to the application of a temperature independent voltage to a chain of diode-connected MOSFETs. This problem could be avoided by replacing the diodes with integrated resistors, which however should be very large for nA-power consumption and would require the availability of large resistance-per-square layers, which is not the case in our technology. In our design, resistors with a resistance-per-square $R_{sq} = 1.3\ k\Omega$ are available and resistors R_1 and R_2 alone occupy most of the area of the chip.

4.3.4 Matching Network

The matching network has been implemented with a simple off-chip L-shaped LC stage, as shown in Figure 4.15, targeting an 868-MHz carrier frequency. Inductance L short-circuits to ground the source of transistor M_1 of Figure 4.12 in DC to allow the flow of current I_{bias} to ground. The L and C component values have been chosen based on an estimate of the chip input capacitance, including pads, and inductance, mainly due to the bonding wire. The matching network provides additional voltage gain, which in the ideal case (lossless network) would be equal to the square root of the ratio between the ED input impedance, Z_{in}, and 50 Ω. Since the ED input

impedance is very high, the actual voltage gain is limited by the network Q, mainly determined by the inductor. For instance, assuming $Q = 80$, it is possible to achieve 18-dB gain.

4.4 AFE Implementation

The described prototype was designed using an STMicroelectronics 90-nm BCD technology. The standard supply voltage $V_{DD} = 1.2\ V$ has been used for all blocks. Both the ED and comparator are biased with 1-nA currents. The follower stage within the ED is biased with 5 nA. The total currents for the ED, the trigger and the biasing block at room temperature are 8.7 nA, about 2.5 nA and 38 nA, respectively. Conversely, at 85°C the ED consumes 12 nA, the trigger 3.5 nA and the biasing block 360 nA. Bias generators I_{bias} in Figures 4.12 and 4.13 are copied from the PTAT current flowing through M_1 in Figure 4.14.

In order to make the thresholds adaptable to the signal level, R_H and R_L in Figure 4.14 are implemented through two resistor chains whose lengths are made programmable with a 5-bit control signal each, designed by another Ph.D. student and shown in Figure 4.16 [28]. The programming is carried out through a Serial-Input Parallel-Output (SIPO) Shift Register which accepts a 10-bit sequence and feeds the bits to the two decoders controlling R_H and R_L in a parallel fashion. This results in the use of only 3 pins, i.e. one for the clock signal, one for the input and one for the output for measurement sessions.

The layout of the implemented circuit is at the bottom of Figure 4.17. Within the actual circuit, the biasing block is shown in blue. As mentioned in Section 4.3.3, resistors make up a large amount of the overall circuit area, which is due to the lack of high resistivity components in the chosen technology. The circuit area, excluding capacitance C in the ED, is 0.7 mm^2, $\sim 60\%$ of which is occupied by the biasing block. A direct chip-on-board bonding technique has been adopted to minimize parasitics at the RF input. Capacitance C was chosen to be off-chip solely for testing purposes. Figure 4.18 shows a photograph of the chip, which is mounted on the board shown in Figure 4.17, designed by another Ph.D. student.

4.5 Results

This section presents the results concerning the first WUR prototype, both simulated and measured. Some differences between the two were found during measurement sessions, as will be described in Section 4.5.2. Four chips were measured.

4.5.1 Simulated results

The post-layout simulated AC response of the ED is shown in Figure 4.19, with the output taken both before (v_{out_amp}/v_{IN}, dashed line) and after (v_{ref}/v_{IN}, solid

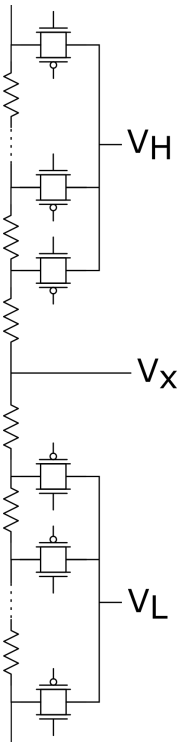

Figure 4.16: Digital potentiometer structures used to generate the Schmitt trigger voltage references [28].

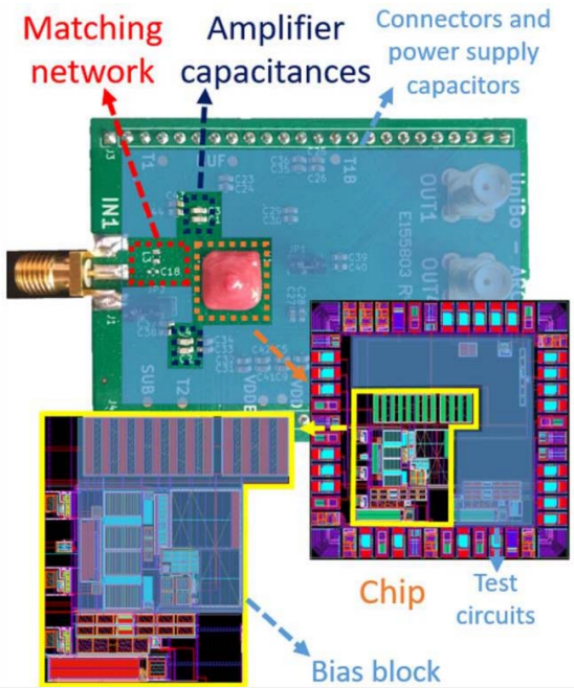

Figure 4.17: WUR testing board and die layout. Picture reused from [2].

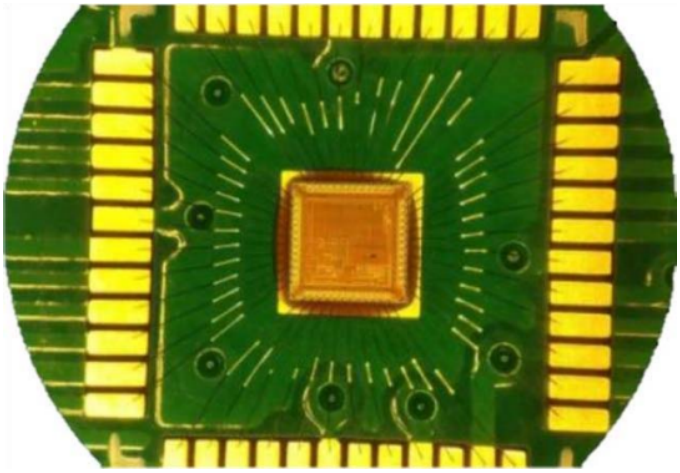

Figure 4.18: Micrograph of the circuit prototype. Picture reused from [2].

line) the AC coupling capacitance, where v_{out_amp} and v_{ref} are defined in Figure 4.14. The simulated integrated output noise power proves to be roughly 200 μV^2. With $v_M = 5\ mV$, an ED output pulse amplitude $V_{OUTM} \cong 50\ mV$ is calculated. The corresponding simulated waveforms are reported in Figure 4.20, which shows the input RF signal (top), the trigger input voltage (that is, the ED output signal after the coupling capacitor) with and without noise together with the set thresholds (center) and the trigger output (bottom). It is apparent that also in the presence of noise the input signal is correctly detected, which gives an estimated sensitivity of 5 mV at the ED input, corresponding to -36 dBm. As an example of the functionality of the ED in process corners, the ED output before the coupling capacitor at -20°C in the MIN corner with low supply voltage and at 85°C in the MAX corner with high supply voltage is shown in Figure 4.21.

The simulated ED rectification gain is plotted in Figure 4.22. The expected linear dependence of G_{RECT} is verified but gain compression is apparent once the input amplitude is high enough. As for the matching network, simulations employing lumped-parameter models of commercial high-Q components at 868 MHz indicate that 18-dB gain can be reached, which leads to a total simulated sensitivity of -54 dBm. With reference to Figure 4.15, the estimated values of C_{in}, including parasitics and pads, is roughly 1.6 pF, while the one of the bonding wire inductance is roughly 4 nH. This results in L = 8.7 nH and C = 400 fF. The reference and threshold voltages generated by the bias block are fairly process stable, thanks to proper transistor sizing: Figure 4.23 and Figure 4.24 show $V_H - V_{ref}$ and $V_{ref} - V_L$, respectively, corresponding to the maximum programmable values for R_H and R_L obtained with a Monte Carlo simulation with 1000 runs at room temperature. $V_H - V_{ref}$ has a mean value of 139.6 mV and a standard deviation of 5.6 mV, while $V_{ref} - V_L$ has a mean value of 143.2 mV and a standard deviation of 5.4 mV.

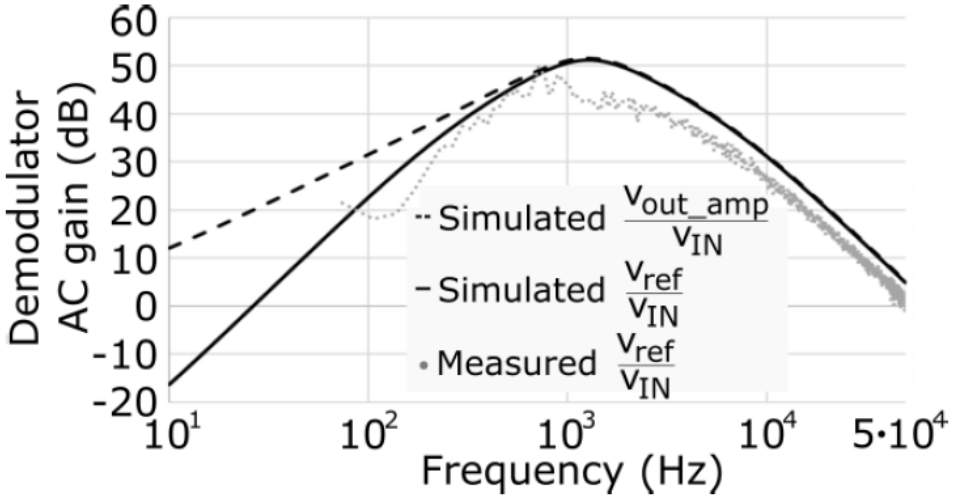

Figure 4.19: Simulated frequency response of the ED: v_{out_amp}/v_{IN} (dashed line) and v_{ref}/v_{IN} (solid line). Measured v_{ref}/v_{IN} with nominal ED bias current (grey dots). Picture reused from [2].

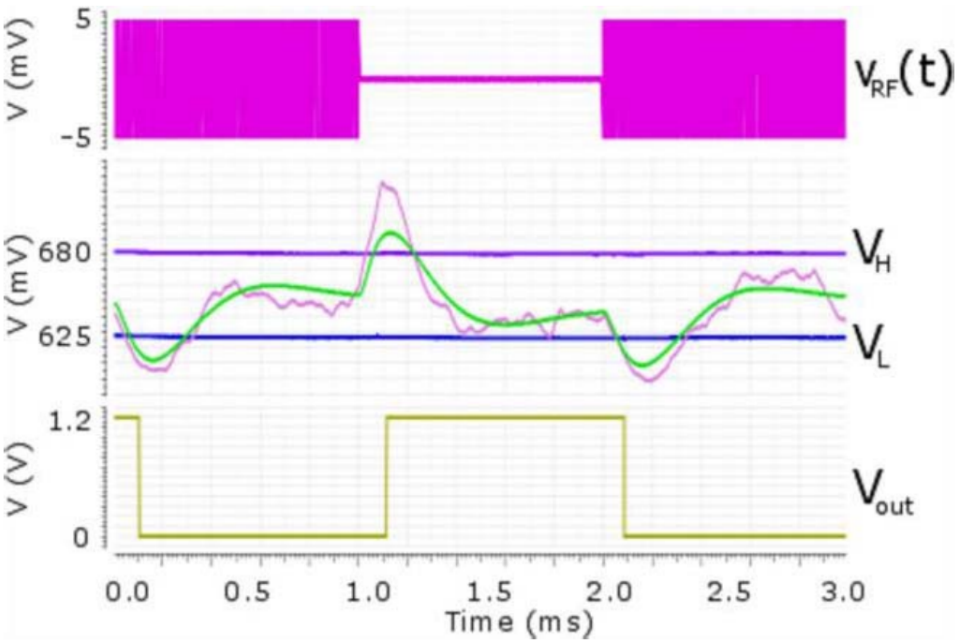

Figure 4.20: Transient noise simulation showing a 1-0-1 RF input voltage with a 5-mV amplitude. From top to bottom: 1) input RF signal; 2) ED output after the coupling capacitor with and without noise (pink and green traces, respectively); 3) output of the inverting trigger. The horizontal violet and blue lines are trigger thresholds V_H and V_L. Picture reused from [2].

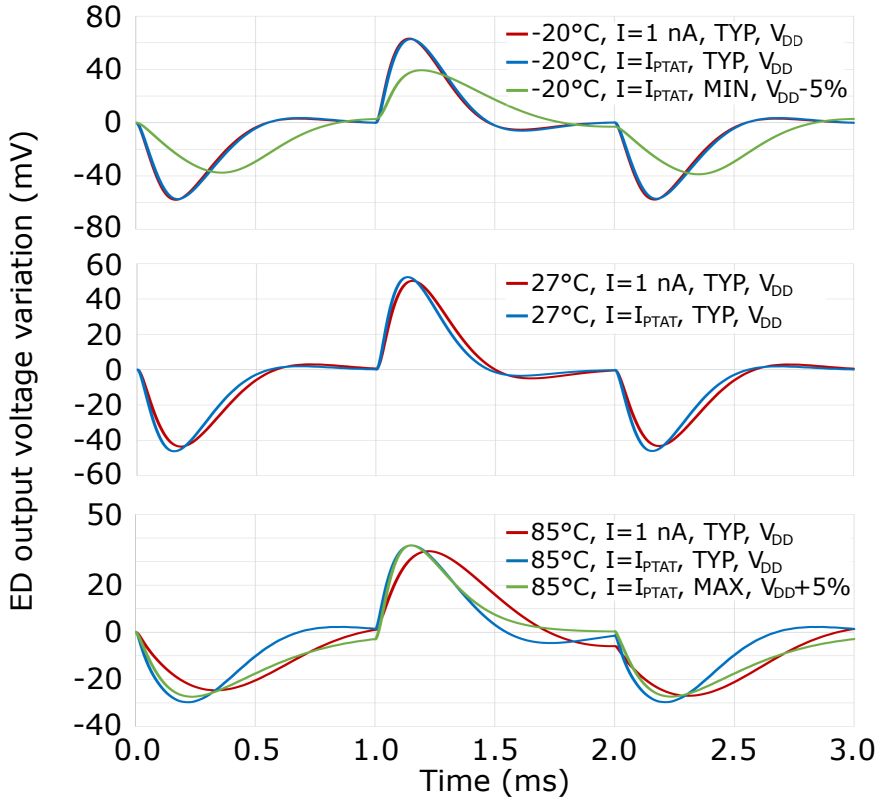

Figure 4.21: Time variation of the simulated ED output signal, before the coupling capacitor, with respect to its quiescent value, at different temperatures and using either a constant 1-nA biasing current or a current copied from the PTAT block of the biasing circuit in Figure 4.14. The two extreme corners, -20°C/MIN with low supply voltage and 85°C/MAX with high supply voltage, are also included. PTAT and constant biasing currents give almost the same results at room temperature and at -20°C. At 85°C the PTAT is preferable. Picture reused from [2].

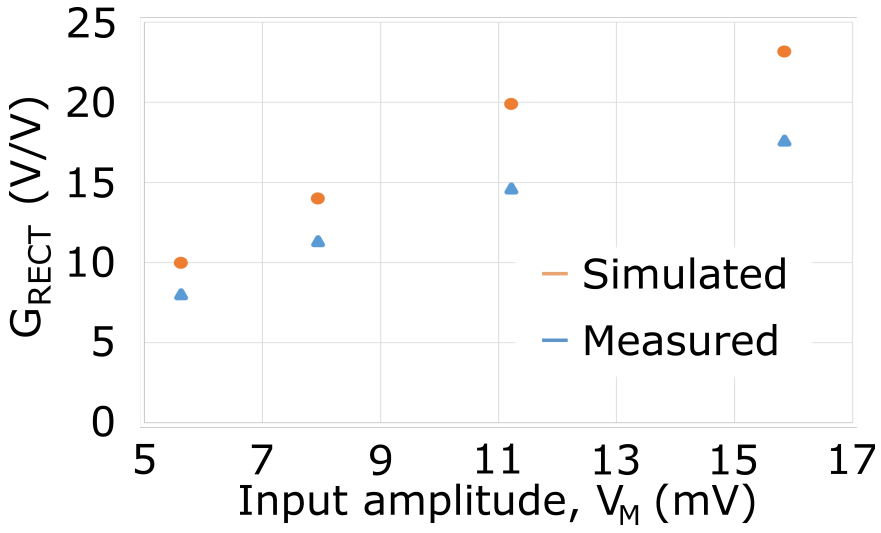

Figure 4.22: Simulated and measured ED rectification gain in a prototype without matching network. Picture reused from [2].

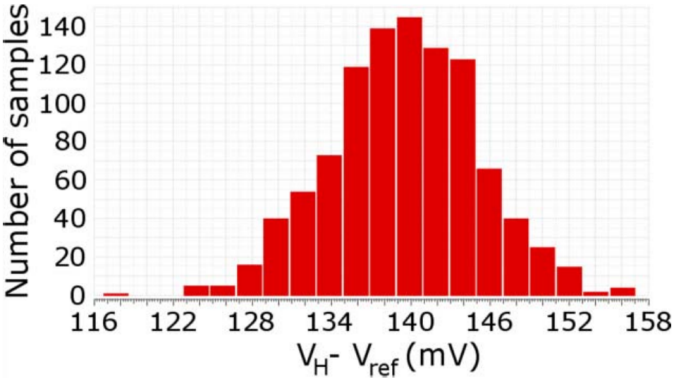

Figure 4.23: $V_H - V_{ref}$ in a 1000-run Monte Carlo simulation with the maximum programmable value for R_H. Picture reused from [2].

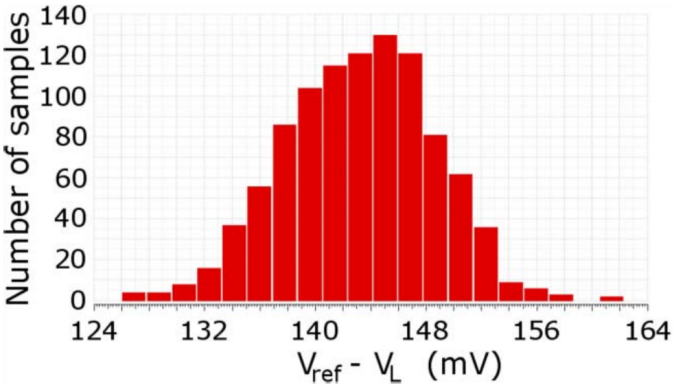

Figure 4.24: $V_{ref} - V_L$ in a 1000-run Monte Carlo simulation with the maximum programmable value for R_L. Picture reused from [2].

4.5.2 Measured results

As stated in Section 4.3.4, the matching network was designed according to simulations rather than measurements, resulting in inaccuracies in the estimation of the chip input impedance Z_{in} and parasitics. In turn, this caused the gain and center frequency of the LC network not to match the simulated ones, likely along with variability issues. The actual frequency corresponding to the matching gain peak, as well as the minimum of the measured reflection coefficient, has proved to be around 771 MHz. The matching gain is calculated as

$$A_V = \frac{v_{RF,1}}{v_{RF,2}}, \tag{4.19}$$

where $v_{RF,1}$ is the RF input amplitude with a $50-\Omega$ resistance in the place of the matching network, while $v_{RF,2}$ is the RF input amplitude with the matching network between the antenna and the ED input, assuming the signal amplitude at the ED output V_{OUTM} to be equal in the two cases. With this process, a matching gain of 12 dB was found instead of 18 dB.

For easier testing, in the actual prototype the biasing currents for the ED and the

Figure 4.25: Measured waveforms corresponding to an input sequence of equal (all 0s or 1s) Manchester-coded bits at 0.5 kbit/s rate. From top to bottom: 1) modulating input signal gating the RF generator; 2) output of the ED; 3) output of the Schmitt trigger. The horizontal blue lines are trigger thresholds V_H and V_L. RF signal: 771 MHz frequency, -44 dBm amplitude. Picture reused from [2].

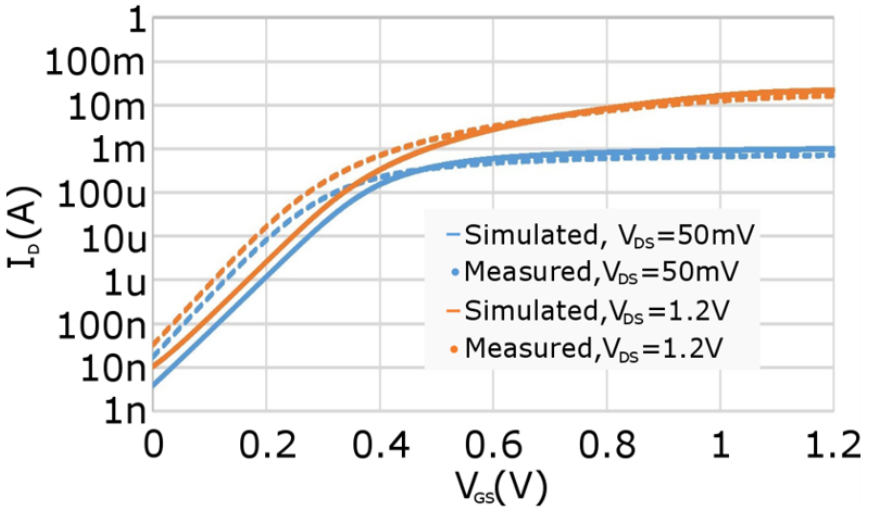

Figure 4.26: Simulated and measured characteristics of a magnified version of MOSFET M_1 in Figure 4.12. It is apparent how simulations overestimate threshold voltage both with a high and a low V_{DS}.

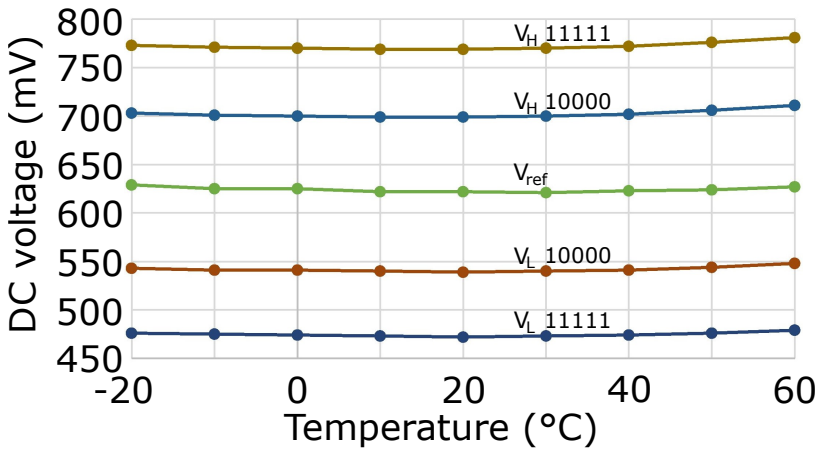

Figure 4.27: Measured trigger threshold voltages V_H and V_L for two control bit configurations and input reference voltage V_{ref} generated by the bias circuit of Figure 4.14 vs. temperature. Picture reused from [2].

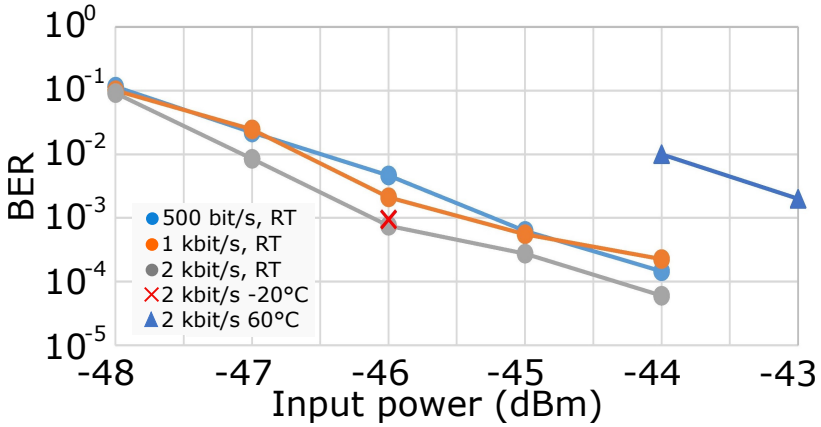

Figure 4.28: BER measured for three sequences of randomly generated Manchester coded bits at 0.5, 1 and 2 kbit/s rate at room temperature (RT). Each sequence is composed of 2.5×10^5 bits. The 2 kbit/s BER at -20°C and +60 °C is also plotted. Picture reused from [2].

trigger, i.e. I_{bias} in Figure 4.12 and Figure 4.13, are not copied from M_1 of the biasing circuit but provided to the chip through an external generator. Therefore, the testing process was carried out using constant ED and trigger biasing currents instead of the planned PTAT one. As shown in Figure 4.21, this leads to results at high temperature slightly worse than expected from simulations.

Sample measured waveforms are reported in Figure 4.25. They refer to a sequence of identical Manchester-coded bits at 0.5 kbit/s rate with a 771 MHz, –44 dBm RF signal at room temperature. The curves demonstrate that the input bit sequence is correctly recognized. The reported measurements have been carried out with bias currents in the ED about four times larger than the nominal value, resulting in a total ED current of 28 nA. This current increase has proved necessary to obtain transient responses similar to the ones found with postlayout simulations during the

design phase. On the other hand, trigger operation was effective also with a current of 2 nA, therefore no current increase was required. Some discrepancies have also emerged between the expected frequency response and the one measured in nominal bias conditions. This is due to an early access to the 90-nm BCD technology node, under development at the time of the design phase of the chip, which resulted in a threshold voltage overestimation in the subthreshold region, as shown in Figure 4.26, and a capacitive effect underestimation. Figure 4.19 shows the AC responses measured by applying an AC signal to the ED input, compared to the simulated ones. The measured ED gain is shown in Figure 4.22 along with the simulated one, displaying the same behavior as simulations.

Figure 4.27 shows the measured V_H, V_L and V_{ref} over the -20°C / +60°C temperature range for two configurations of the control bits. This shows that the threshold margins are well controlled over the entire range. Measurements over 60°C were not possible due to the lack of heat resistant cables. The BER vs. input RF power measured for 0.5-, 1- and 2-kbit/s Manchester-coded bit streams, each of which is composed of 2.5×10^5 randomly generated bits, is plotted in Figure 4.28. An ED bias current equal to roughly 4x the nominal value has been used. A 10^{-3} BER is achieved with a -46 dBm RF input power at room temperature and at -20 °C, almost -43 dBm at 60 °C.

The measured performances of this prototype are summarized in Table 4.3. A 106.9-dB FoM shows these performances are aligned with the performances of some of the prior art, especially at the time when this prototype was designed and implemented. A significant trend towards the reduction of supply voltages in order to reduce power consumption started in 2018. This lead to a reduction to 0.6 V in the supply voltage of the second prototype presented in this thesis as well. Additionally and even more importantly, a LP ED response was chosen to avoid the need for Manchester coding and simplify the architecture by using a standard comparator instead of the trigger.

4.6 Second prototype: active low-pass Envelope Detector

As previously mentioned, the second prototype presented in this thesis features an active LP ED. The structure of the overall AFE is shown in Figure 4.29. First of all, again a clockless approach was chosen for the reasons explained in the previous Section. The AFE in this prototype is an evolution of that presented in Section 4.3. The ED is very similar to the one implemented in the first chip, the main difference being that this time the frequency response is the so-called LP one, which outputs a signal with the same shape as the extracted envelope and allows the use of the voltage across C as a threshold for the subsequent decision circuit. This implies the decision circuit can be implemented as a standard comparator instead of a trigger and there is no need for threshold voltage generation within the biasing block, which only provides biasing currents for the other blocks. Moreover, there is no need for

Table 4.3: Performance comparison between the first proposed prototype and several prototypes in literature which had already been presented at the time of design of the first prototype.

	Proposed prototype 1	[9], 2012	[10], 2013	[11], 2016	[6], 2016	[12], 2017	[13], 2017	[8], 2018
Temperature range [°C]	-20 : 60	unknown	unknown	unknown	unknown	unknown	unknown	unknown
Carrier frequency [MHz]	771	915	915	2400	868	2400	403	113.5
Matching gain [dB]	12	12	5	5	unknown	12.5	18.5	25
Bitrate [$kbit/s$]	2	100	12.5	8.192	1	200	0.3	0.3
Sensitivity at 27°C [dBm]	-46[a]	-41[a]	-43[b]	-56.5[b]	-55[b]	-50[a]	-63.8[b]	-69[b]
Supply voltage [V]	1.2	1.2	1.2/0.5	1/0.5	1.2	0.8	0.4	0.4
Power at 27°C [nW]	36	98	116	236	1200	4500	4.5	4.5
Correlator	no	no	yes	yes	yes	no	yes	yes
Code length	N.A.	N.A.	31	31	16	N.A.	32	16
Technological node [nm]	90	130	130	65 low-power	N.A.	180	180	180
Die area [mm^2]	0.7	0.03	0.35	2.25	N.A.	0.456	4.5 (estimated)	4.5 (estimated)
External components	LC	LC	LC	LC	N.A.	LC	transformer	transformer
$FoM_1[dB]$	106.9	106.1	102.8	112.3	99.2	100	129.7	134.9
$FoM_2[dB]$	100.1	101.4	unknown	unknown	100	102.8	115.9	121.1

[a] Sensitivity defined through 10^{-3} Bit Error Rate
[b] Sensitivity defined through 10^{-3} Missed Detection Rate

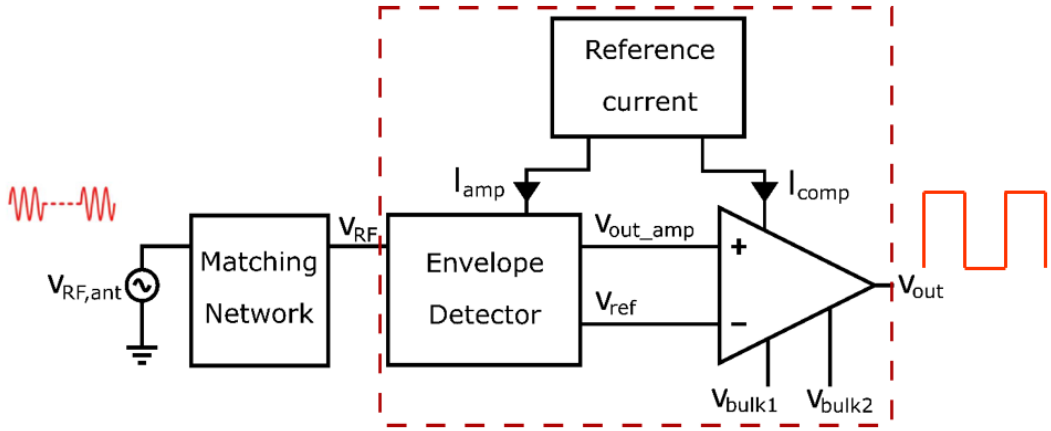

Figure 4.29: High level block diagram of the AFE within the second prototype.

Manchester coding, resulting in the number of equal consecutive bits (N_m) being virtually unlimited, as will be shown in Section 4.6.1. On the other hand, there is a maximum packet length, again due to the LP response. In this prototype, proper comparator offset control is important since it may affect the very functionality of the system, as will be explained in Section 4.6.2, and is performed by providing externally the bulk voltages of the MOSFETs belonging to the input differential pair within the comparator itself. The main drawback of this architecture is the need for an off-chip capacitor.
This prototype includes both the AFE and the BBL, designed by another Ph.D. student. The BBL employs Clock and Data Recovery (CDR) circuits with the purpose of correctly synchronizing the bitstream which the AFE outputs to the internal oscillator. This allows to perform a digital correlation between the received bitstream and the address of the specific node, the codeword. If they match, a Wake-Up impulse is generated to turn on the rest of the node. Moreover, it allows to receive packets with a longer length than the codeword.

4.6.1 Envelope Detector

As just mentioned, an active ED was implemented in this prototype as well, both performing OOK demodulation, i.e. envelope extraction, and baseband amplification of the extracted envelope. It is shown in Figure 4.30 and is ultimately very similar to the sample circuit used for the performance comparison between the BP and the LP approach reported in Table 4.1. It is a direct evolution of the ED implemented in the first prototype and shown in Figure 4.12.

In order to reduce power consumption, the supply voltage was lowered, leading to the removal of cascode transistor M_2. The behavior of the circuit is still described by Equation 4.1 but $R_o = r_{out1}$, again assuming the output resistance of current generator I_{bias} to be much higher than r_{out1}. This obviously results in a reduction of

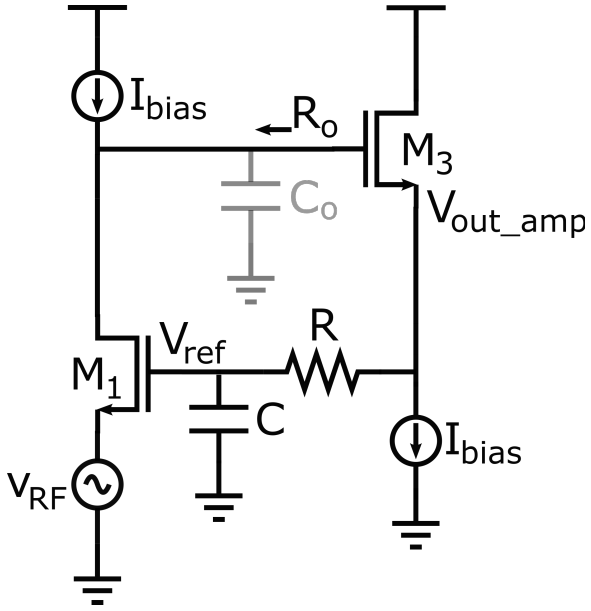

Figure 4.30: Schematic of the ED proposed within the second prototype. Picture reused from [2].

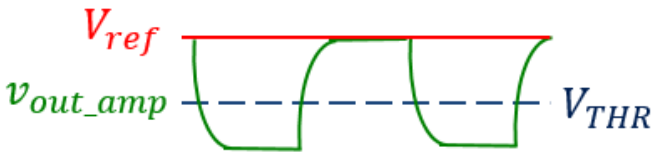

Figure 4.31: Sketch showing the behavior of the ED output voltage v_{out_amp} in green, the voltage across capacitance C V_{ref} in red and the effective threshold of the comparator V_{THR} in blue. Picture reused from [23].

the in-band gain, which in turn causes no changes in the expected sensitivity since integrated noise is reduced as well. Even more importantly, this ED has a LP pass frequency response. As mentioned in Section 4.2, this is achieved by making time constant R^*C bigger, thus moving the first pole to lower frequencies, as shown in orange in Figure 4.5. The most straightforward way to do so is increasing the value of C. The time-domain response, shown in orange in Figure 4.7, is unipolar: when a "0" is received there is no change in the ED output voltage v_{out_amp}, whereas when a "1" is received v_{out_amp} settles to a lower value, as shown in the sketch in Figure 4.31. Therefore, the output signal corresponds to the LP filtered version of the input envelope. This modification has both positive and negative implications. First of all, the drastic increase in the value of C entails the impossibility to integrate it. Therefore, unlike in the BP case, an off-chip capacitor is needed and a maximum packet length exists (more on this in Section 4.6.2). On the other hand, a major advantage lies in the fact that voltage V_{ref} stays almost equal to its quiescent value, corresponding to zero RF input signal, also during the reception of an entire packet. This allows the use

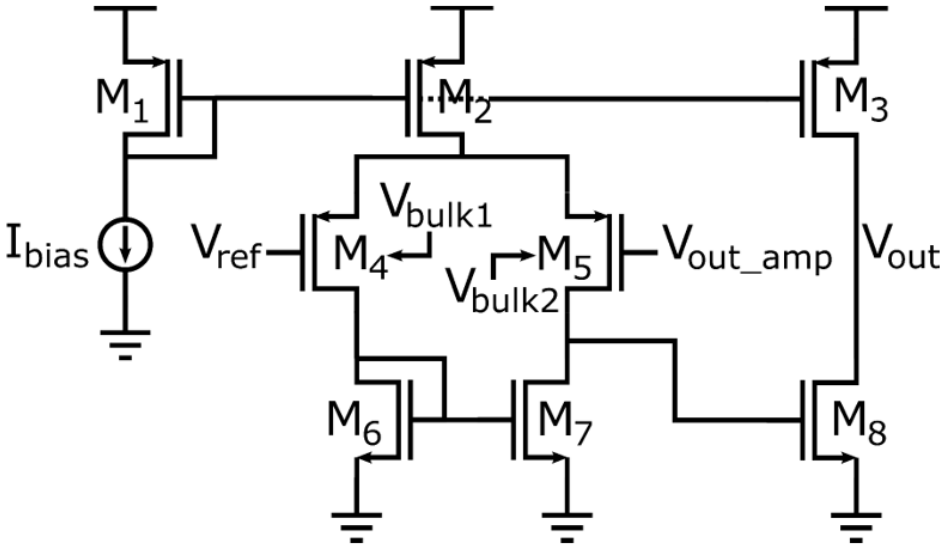

Figure 4.32: Schematic of the comparator proposed within the second prototype and presented in [23].

of V_{ref} as a threshold for the subsequent comparator, avoiding the need for complex threshold generation with additional blocks.

4.6.2 Comparator

Figure 4.32 shows the schematic of the implemented comparator, a standard one. Just as in the previous prototype, the main design target was to keep its input offset and noise voltage at values that do not limit the overall receiver sensitivity and again this task was not difficult thanks to a large enough ED gain, though smaller than in the previous prototype. This time an additional issue is speed, due to the absence of a positive feedback loop, but proper sizing allows fast enough switching. As explained in Section 4.6.1, V_{ref} can be used as a reference for the comparator, which also receives the output of the ED v_{out_amp} as an input. When a "0" is being received V_{ref} and v_{out_amp} have the same value, so in the absence of offset in the input differential pair the value of comparator output voltage v_{out} would be unpredictable. Therefore, introducing such an offset is essential for the very functionality of the comparator. However, if the polarity of this offset was wrong, the comparator could not work either. Therefore, it is important to set an effective comparator threshold which lies exactly at a voltage between the two possible values which v_{out_amp} can have, as shown in Figure 4.31. This task is performed as explained in the following.

Comparator offset control

Let's consider the differential pair in Figure 4.33, where only one of the two input transistors, which are matched, has a source-to-body voltage different from zero. Its offset, V_{off} in the Figure, can be calculated by connecting the two inputs to the

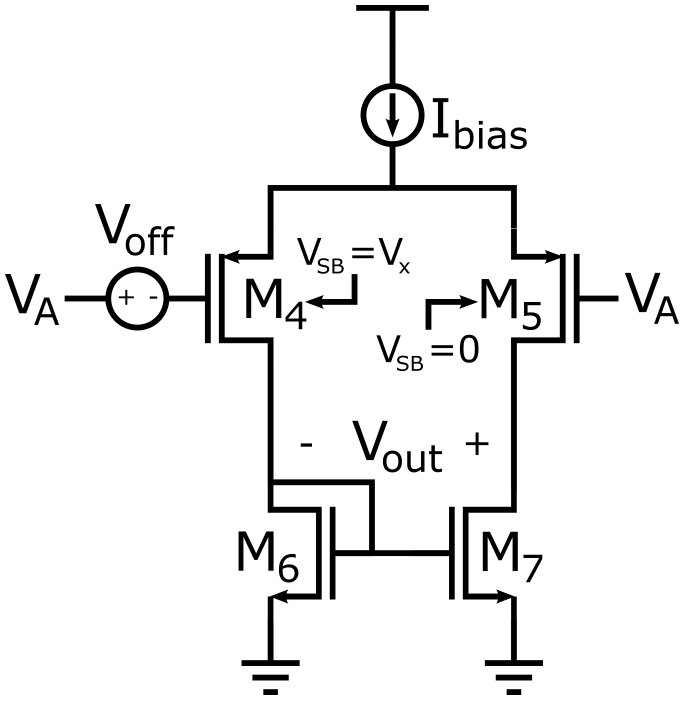

Figure 4.33: Differential pair with an offset introduced by leveraging body effect in one of the two input transistors.

same voltage V_A and setting the output differential voltage, v_{out}, to 0. This way it is possible to establish the relationship in the following equation:

$$\frac{W}{L} e^{\frac{V_{GS}+V_{off}-V_{TH}}{nV_T}} = \frac{W}{L} e^{\frac{V_{GS}-V_{TH0}}{nV_T}}, \tag{4.20}$$

where V_{TH0} is the threshold voltage of M_5, which does not have body effect, and V_{TH} is that of M_4, which does have body effect. Applying logarithms on both sides, we obtain

$$V_{off} = V_{TH} - V_{TH0}. \tag{4.21}$$

From theory,

$$V_{TH} = V_{TH0} + \gamma \left(\sqrt{|2V_f + V_{SB}|} - \sqrt{2V_f} \right) = V_{TH0} + \gamma \left(\sqrt{|2V_f + V_x|} - \sqrt{2V_f} \right), \tag{4.22}$$

where γ is the bulk threshold parameter, related to technology, and V_f is the electrostatic potential of the substrate, which is linked to its doping. Since both are technological parameters, they can be estimated through simulations. The substitution of Equation 4.22 into 4.21 yields

$$V_{off} = \gamma \left(\sqrt{|2V_f|} - \sqrt{2V_f + V_x} \right) = \left(\sqrt{|2V_f|} - \sqrt{2V_f + (V_{bulk1} - V_{bulk2})} \right). \tag{4.23}$$

Therefore, it is ultimately the difference between V_{bulk1} and V_{bulk2} which determines V_{off}.

Maximum packet length

Unfortunately, V_{ref} is not perfectly stuck to its quiescent value throughout the reception of the whole packet. Figure 4.34 shows this phenomenon. Whenever a "1" is received, v_{out_amp} lowers and compresses the I_{bias} generator which biases M_3. Therefore, some current flows from V_{ref} to v_{out_amp}, thus slightly discharging C. Since v_{out_amp} is unipolar, i.e. it cannot have higher values than its quiescent value, there is no way for C to regain the lost charge. Moreover, since there is AC gain from V_{ref} to v_{out_amp} right after the first pole, the slight decrease in V_{ref} translates into a much higher increase in v_{out_amp}. Therefore, after many received "1"s the comparator is no more able to correctly discriminate between a “0” and a “1”, as shown in Figure 4.35. As a matter of fact, this sets a maximum number of received "1"s, thus a maximum packet length assuming packets to include the same number of "1"s as the number of "0"s.

4.6.3 Current reference generator

Since the biasing of the ED with a PTAT current has been proven effective for a more stable ED gain in Section 4.5.1, this approach was adopted in this prototype as well. The reference current generator implemented in the second prototype, shown in Figure 4.36. It is very similar to the cell generating temperature-stable voltage V_A in Figure 4.14, with the only difference that the negative feedback loop setting the voltage at the gate of M_4 is achieved by just connecting its gate to its drain. Again, a temperature-stable voltage, V_{BG}, is generated as the sum of PTAT component $2I_0R_2$ and CTAT component V_{GS4}. Due to the reduced supply voltage, the operation of this block is less precise than that in Figure 4.14, in particular in mirroring the current of M_1 into M_2 at high temperatures. Still, it is sufficient to provide a proper PTAT current.

4.6.4 Matching Network

Just like in Section 4.3.4, the matching network has been designed as a simple off-chip LC stage with multiple purposes, i.e. matching the input impedance of the chip to the antenna impedance, providing passive gain to the incoming signal and short-circuiting to ground the source of transistor M_1 in Figure 4.30 at DC to allow the flow of current I_{bias} to ground. Again the values for components L and C were chosen based on an estimate of the chip input impedance. This time it was decided

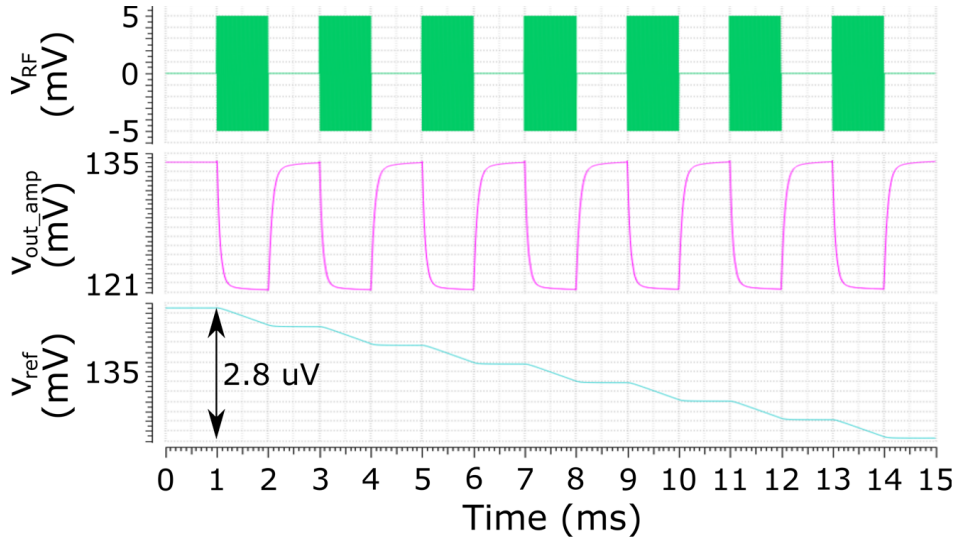

Figure 4.34: Transient simulation involving a stream of alternate bits. The RF input voltage has a 5-mV amplitude (top, green trace). The ED output voltage, v_{out_amp} (middle, pink trace), is a low-pass filtered version of the envelope of the input OOK-modulated signal. The voltage across capacitor C V_{ref} (bottom, blue trace), undergoes a slight decrease whenever a "1" is received.

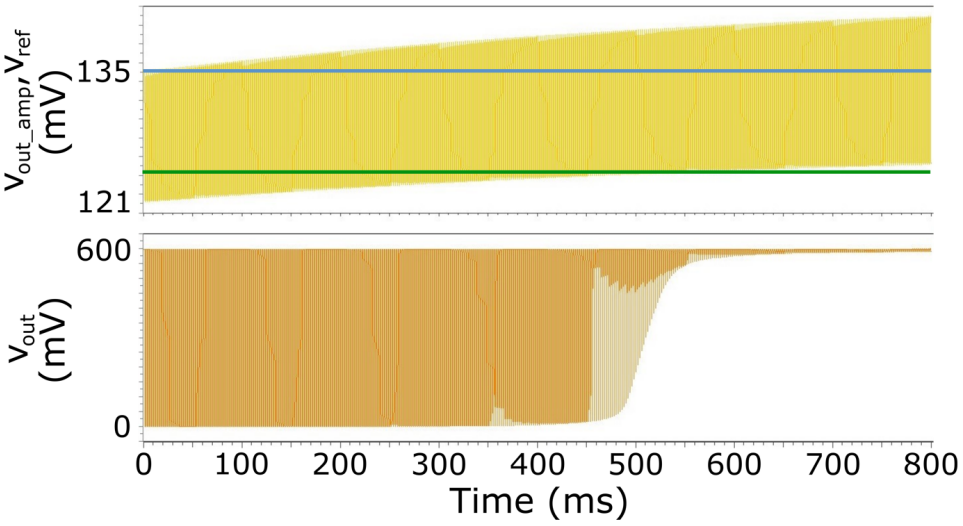

Figure 4.35: Simulated time-domain response of the ED output v_{out_amp} (on top in yellow) and of the comparator output v_{out} (at the bottom in orange) to an 800-bit 1-0-1-0 packet at 1 kbit/s with a 5-mV input amplitude. The supply voltage is 600 mV for the entire AFE. V_{ref} and V_{THR} (on top in blue and green, respectively) are shown as well. V_{ref} decreases by around 100 μV, which is clearly not visible. On the other hand, the ED output rises significantly after a few hundred bits as a result of the lower pole, preventing the correct operation of the comparator.

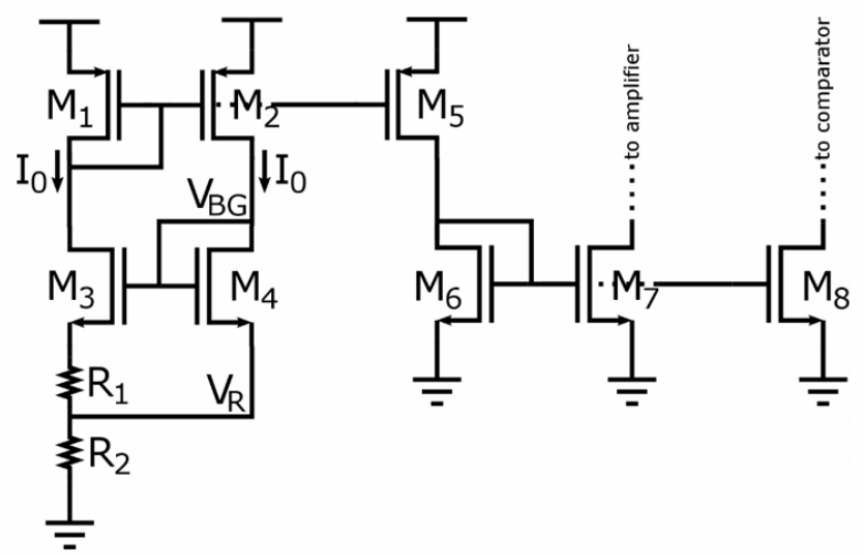

Figure 4.36: Reference current generator implemented in the second prototype. It is very similar to the cell generating temperature-stable voltage V_A in Figure 4.14. Picture reused from [23].

to recalculate them based on the actual measured input impedance though. More details on this are provided in Section 4.8.2.

4.7 Implementation

The circuit was designed using the same STMicroelectronics 90-nm BCD technology which was employed for the previous prototype, presented in Section 4.3. This time the supply voltage was reduced in order to allow a reduction in power consumption, from $V_{DD} = 1.2\ V$ to $V_{DD} = 0.6\ V$. Obviously, this implies that a regulator would be needed in a complete implementation. As already mentioned, this prototype also includes a BBL, which has been designed by another Ph.D. student and will be presented in section 4.7.2. A Bias and Calibration circuit (BC) was added as well.

The layout of the implemented circuit is shown in Figure 4.37. The AFE, the BBL and the BC circuit are circled in yellow. The overall circuit area is around 0.37 mm^2, 0.2 of which due to the AFE, 0.13 due to the BBL and 0.04 due to the BC circuit. Most of the AFE area is due to passives, in particular resistance R in the ED and resistances R_1 and R_2 in the reference current generator. Moreover, unlike in the previous prototype, in this case capacitor C within the ED is way too big to be integrated, so it needs to be mounted on the board, also shown in Figure 4.37 and designed by another Ph.D. student.

4.7.1 AFE implementation

Both the ED and comparator are biased with 1-nA currents. The follower stage within the ED block is biased with roughly 1 nA as well. The total currents for the

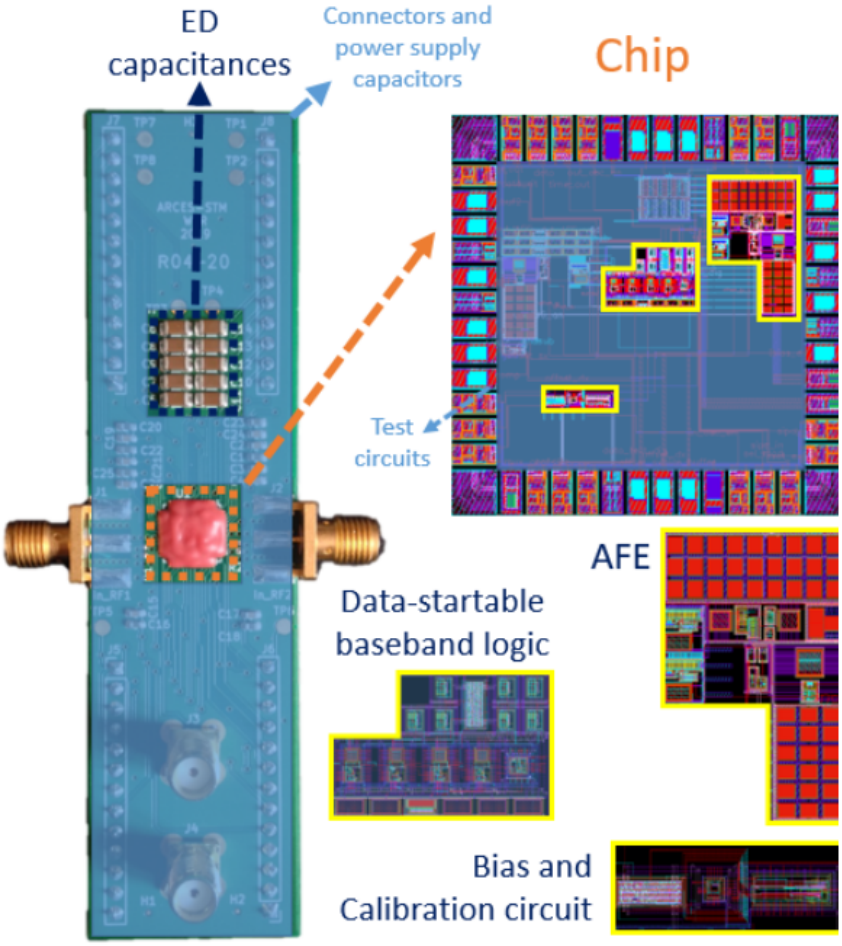

Figure 4.37: WUR testing board and die layout.

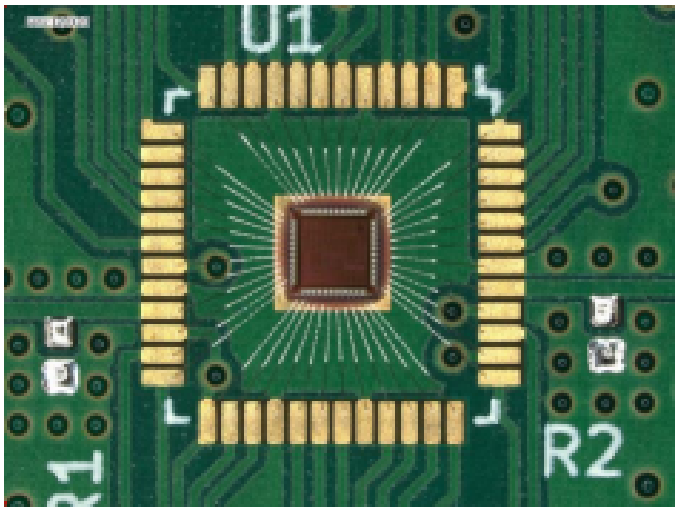

Figure 4.38: Micrograph of the circuit prototype. Picture reused from [23].

ED, the comparator and the reference current generator at room temperature are 4.5 nA, about 2 nA and 6 nA, respectively, resulting in roughly an 8-nW consumption. Conversely, at 85°C the ED consumes 10 nA, the comparator 4.3 nA and the biasing block 8.2 nA. The currents in bias generators I_{bias} in Figures 4.30 and 4.32 come from PTAT current I_0 through transistors M_7 and M_8 in Figure 4.36. Capacitance C in Figure 4.30 is an off-chip 500-nF capacitor, whereas R is a 75-MΩ integrated resistor. Resistors R_1 and R_2 in Figure 4.36 are integrated 13 and 113-MΩ resistors.

4.7.2 BBL implementation

As already mentioned, this prototype also includes a BBL, making the block diagram of the overall chip the one shown in Figure 4.39. The combination of a data-startable BBL and a clockless AFE allows system operation in two phases as described in Section 3.3.2. In particular, during phase 1, only the AFE is active, whereas phase 2 starts upon recognition of the first 0-to-1 transition of the message, occurring at the first transition of the AFE output signal, and the BBL is turned on to compare the incoming bitstream to the stored codeword, i.e. the node address. This allows the reduction of the overall power consumption, especially if the specific application is characterized by long idle periods.

The architecture of the proposed data-startable BBL, which has been designed by another Ph.D. student, is summarized in the following. It is composed of a Gated-Oscillator Clock and Data Recovery (GO-CDR) circuit, a Control Logic (CL) with addressing capabilities to generate the Wake-Up signal and the BC circuit for the GO-CDR. The purpose of the GO-CDR is to provide a clock to the CL to correctly sample on the positive edges a delayed version of the bitstream at the output of the AFE, D_{in} in Figure 4.39. Ideally, the sampling edges are placed at the center of each bit-time (T_b), as shown in Figure 4.41.

Figure 4.40 shows the block diagram of the GO-CDR, which is composed of a delay block, an edge detector implemented through an EXNOR and the GO. The EXNOR gate is fed with D_{in} and DD_{in}, its delayed version, resulting in a $Gate$ signal downwards pulse at each D_{in} transition. When $Gate = 1$ the GO is in free-running mode with frequency $f_{ck} = 1/T_{ck}$, whereas when $Gate = 0$ it is blocked in a predefined state. When $Gate$ switches from 0 to 1, a positive edge of the clock occurs, ideally after $T_{ck}/2$. Therefore, any possible phase error accumulated up to this moment gets cleared, even if the free-running frequency is not precisely matched to the data rate, i.e. $f_{ck} \neq 1/T_{ck}$. Therefore, the only constraint posed by the BBL is the maximum number of equal consecutive bits, N_m, making it possible to receive packets of a longer length than just the codeword. N_m can be calculated as a function of the relative error in the free-running frequency $\alpha = |T_{ck} - T_b| / T_b$:

$$N_m < \frac{1-\alpha}{2\alpha}, \tag{4.24}$$

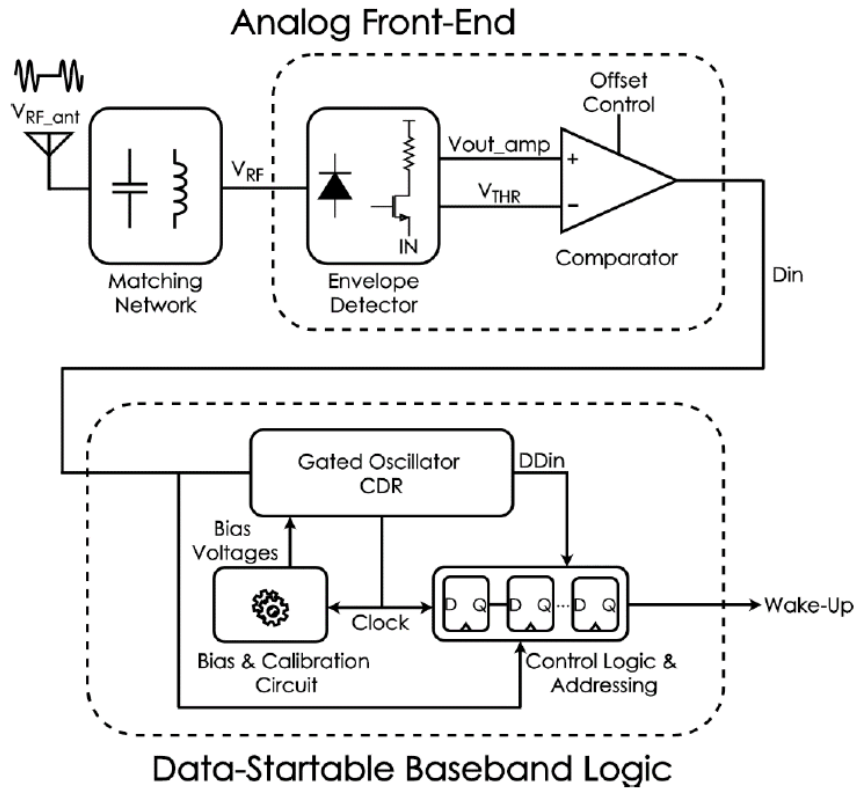

Figure 4.39: Block diagram of the second WUR prototype presented in this thesis. Picture reused from [23].

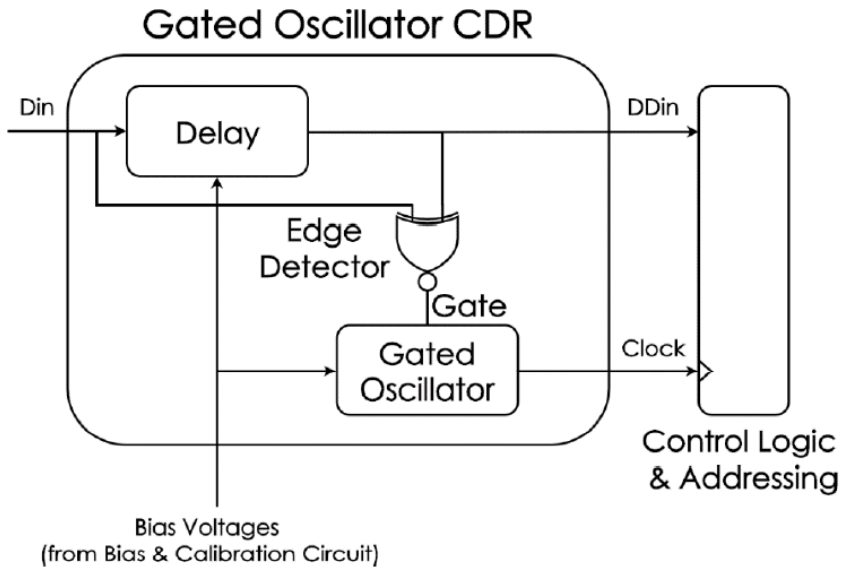

Figure 4.40: Block diagram of the GO-CDR implemented in the second WUR prototype. Picture reused from [23].

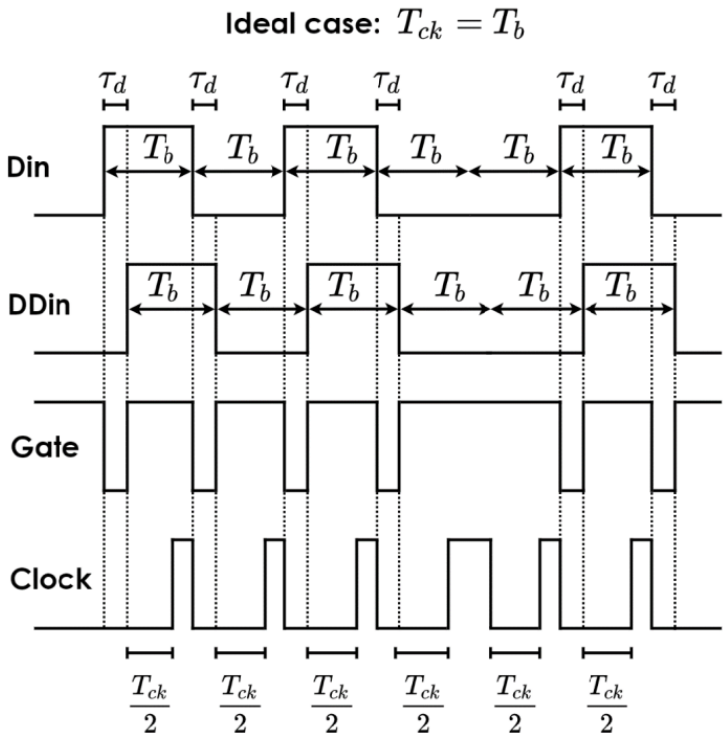

Figure 4.41: GO-CDR circuit behavior. T_{ck} is the clock period, T_b is the bit-time and τ_d is the delay between the bitstream at the output of the AFE, D_{in}, and its internally delayed version, DD_{in}. Picture reused from [23].

assuming no bits are sampled twice or not sampled at all. As mentioned, the proposed BBL also includes a BC for the GO-CDR in order to reduce α to negligible values and allow the WUR to process data containing long sequences of equal consecutive bits.

The GO, shown in Figure 4.42, consists of a three-stage ring oscillator. Each stage is composed of a current-starved inverter, a capacitor and two additional MOSFETs driven by the *Gate* signal to reset the output of each current-starved inverter (O1, O2 and O3) to a predefined state at each pulse in the *Gate* signal. The oscillation frequency is $1/2N\tau_p$, where $N = 3$ is the number of stages and τ_p is the propagation delay of each stage, yielding $\tau_p = T_{ck}/6$. Bias voltages *vbias_p* and *vbias_n* control the charging and discharging currents for capacitors $C1$, $C2$ and $C3$, thus the value of τ_p and ultimately the oscillator frequency. The delay block consists of a stage equal to the ones used in the GO and biased by the same control voltages *vbias_p* and *vbias_n*. Finally, the CL is composed of a SIPO register, in which the codeword is stored along with other configuration parameters, a correlator with a programmable codeword and threshold to compare the incoming bitstream to the stored codeword and additional circuitry. A time-out mechanism has also been implemented which turns off the GO-CDR either right after a Wake-Up impulse or after a predefined amount of time in case of no Wake-Up, in order to avoid wasting energy in the event

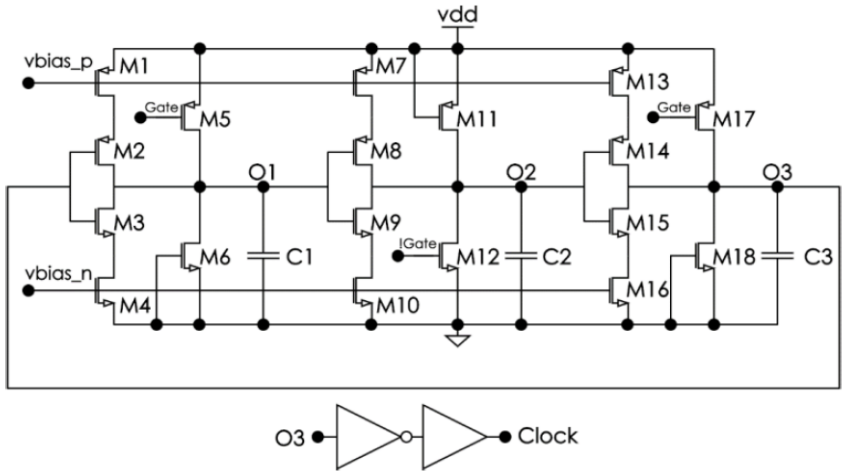

Figure 4.42: Schematic of the GO implemented in the second WUR prototype. Picture reused from [23].

of frequent transitions at the AFE output due to noise.
With reference to Figure 4.42, the nominal value for the charging and discharging currents is 2 nA to generate a free-running 1 kHz clock frequency with capacitances $C1 = C2 = C3 = 1.1\ pF$. The BC circuit guarantees a frequency error after calibration within 0.5%. At room temperature, the overall consumption of the BBL during phase 1 is 4.8 nW. Added to the 8 nW consumed by the AFE, this yields a 12.8-nW system consumption during phase 1. On the other hand, during phase 2 the consumption of the BBL is 9 nW, making the system consumption 17 nW in phase 2. Assuming the system to spend 1% of time in phase 2, its average power consumption can be estimated at 12.9 nW.

4.8 Results

Simulation results show good performances despite PVT variations of both the AFE and the BBL. As will be explained in Section 4.8.2, the matching network was not implemented and measurements were carried out using a coaxial impedance adapter. Four chips were measured. Moreover, measurements at different temperatures were not carried out.

4.8.1 Simulated results

Figure 4.43 shows the simulated AC gain of the ED proposed in the second prototype, Figure 4.30, for different temperatures, showing a similar response in the whole operating temperature range. As an example, a 1000-run Montecarlo simulation of the AC gain of the ED, shown in Figure 4.44, proves the behavior of the ED is stable with process variations. Figure 4.45 shows the simulated and measured rectification gain of the ED. Due to the absence of the cascode, the AC gain is smaller than in the

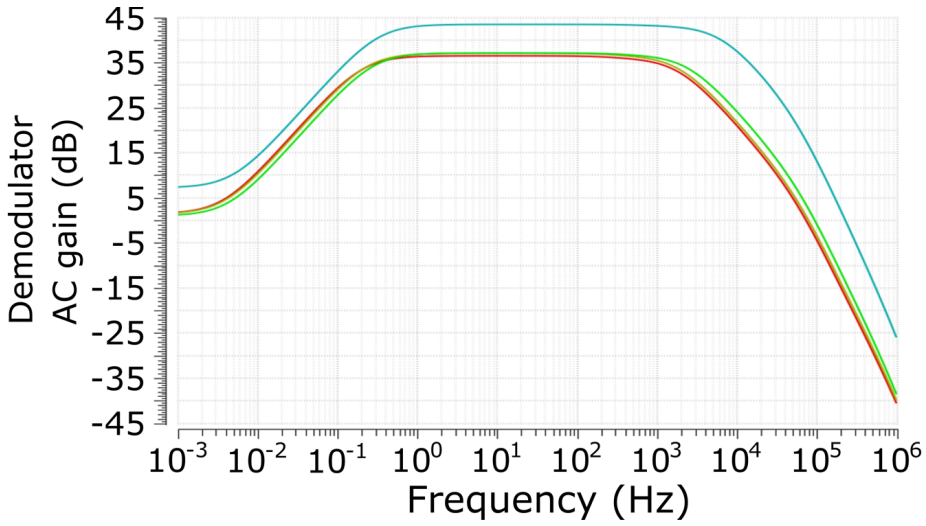

Figure 4.43: AC response of the ED proposed in the second prototype, Figure 4.30, at different temperatures: -40°C in red, -20°C in yellow, 27°C in green, 85°C in blue.

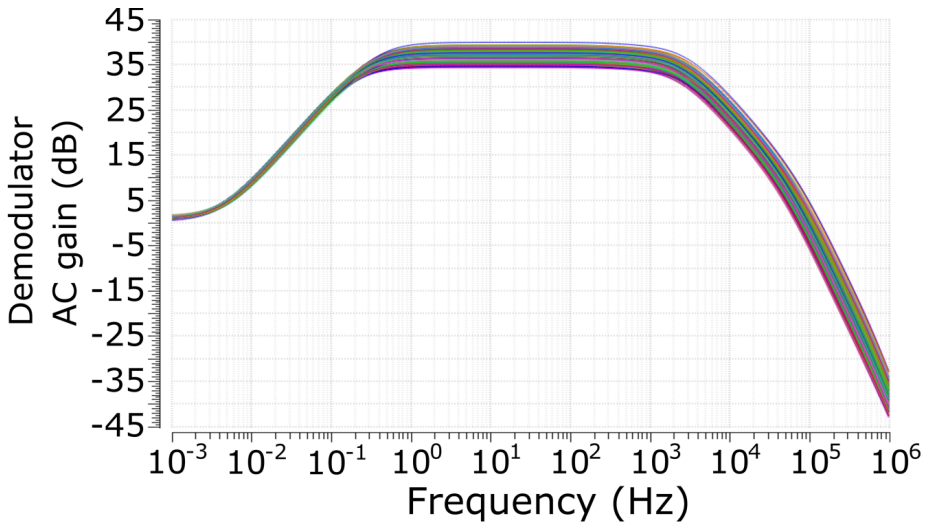

Figure 4.44: 1000-run Monte Carlo simulation of the AC gain of the ED proposed in the second prototype, Figure 4.30, at room temperature.

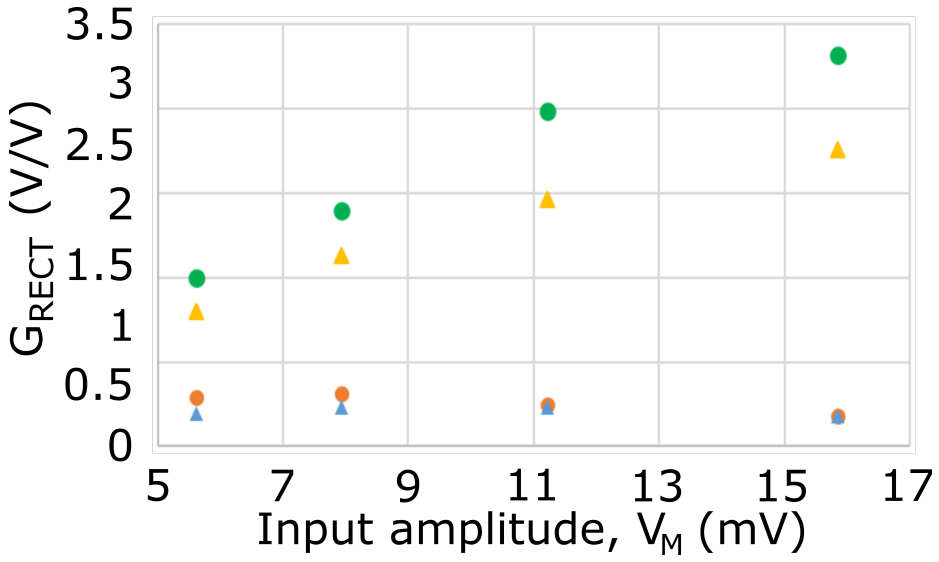

Figure 4.45: Simulated and measured (in orange and blue, respectively) ED rectification gain in a prototype without matching network. The simulated and measured ED rectification gain of the previous prototype has been reported as well, in green and yellow respectively.

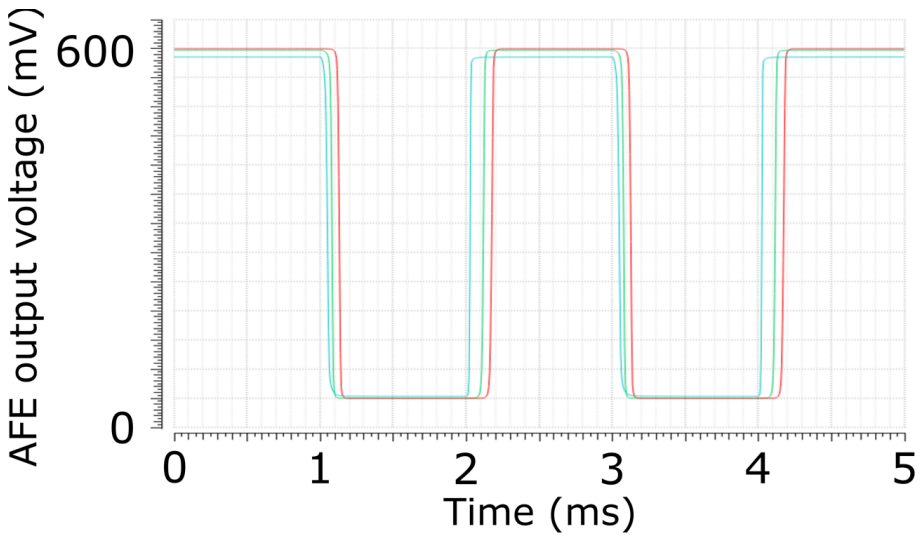

Figure 4.46: Transient response of the output of the AFE proposed in the second prototype to a bitstream with a 5-mV RF input signal carrying 0-1-0 at different temperatures: -40°C in red, 27°C in green, 85°C in blue.

Table 4.4: Simulated performances of the AFE presented in the second prototype.

Temperature [°C]	Power consumption [nW]	Estimated sensitivity [dBm]
-40	4.5	-39
27	7.5	-37.5
85	13.5	-39

previous prototype (around 35 dB instead of roughly 50). This leads the rectification gain to be smaller than in the previous prototype as well (Figure 4.22). However, gain compression appears for lower input power values due to the voltage across the 5 I_{bias} generator at the lower right in Figure 4.12 being significantly higher than the voltage across the I_{bias} generator at the lower right in Figure 4.30, leaving more room for the voltage swing occurring when a 1 is being received. As a matter of fact, $V_{DD} = 1.2\ V$ in the first prototype whereas $V_{DD} = 0.6\ V$ in the second one. Moreover, the voltage across diode D_3 in Figure 4.12 at the reception of a 1 gets smaller as compression increases, helping the circuit absorb it.
Figure 4.46 shows the transient response of the comparator, Figure 4.32, to a bitstream with a 5-mV RF input signal carrying 0-1-0 at different temperatures: simulations show correct comparator operation throughout the whole operating temperature range. As for the BBL, simulations showed the BC circuit could effectively compensate for PVT variations.
Table 4.4 summarizes the performances of the AFE in the second prototype over its simulated temperature range, from -40 to 85°C.

4.8.2 Measured results

In the following, the measurement results of the overall second WUR prototype will be presented. Both the AFE and the BBL, designed by another Ph.D. student, contribute to the functionality and performances of the prototype. As a matter of fact, unlike for the previous prototype, the MDR and FAR metrics become important in assessing WUR performances. As already mentioned, measurements at different temperatures were not carried out.
Figure 4.48 shows the employed measurement setup. It includes an RF generator for the RF input signal and its OOK modulation. An STM32 Nucleo board (Main Nucleo in Figure 4.48) was used for the generation of the bitstream, programming the SIPO register, processing the output bits generated by the WUR and managing clock frequency calibration. An additional STM32 Nucleo board was used to characterize the impact of the GO-CDR on WUR sensitivity. For the reason which will be explained in Section 4.8.2, all measurements shown hereafter have been performed with a 50-Ω resistor soldered in parallel to the input of the ED and using a commercial coaxial impedance adapter (see Figure 4.48), thus providing a unity matching

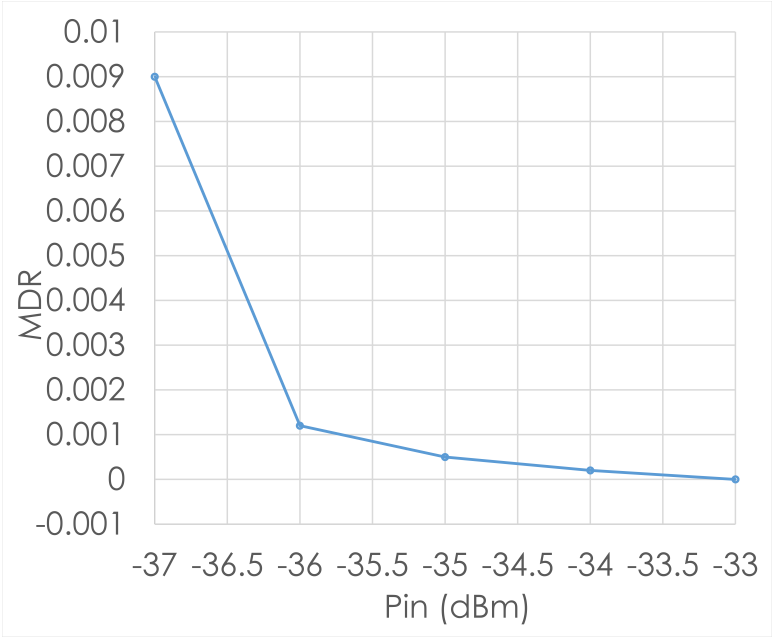

Figure 4.47: MDR measurements tolerating no errors in the received packet. Sensitivity corresponds to a 10^{-3} MDR. Picture reused from [23].

gain. Since the AFE response is broadband with respect to the RF carrier frequency, all measurements were performed using the 868 MHz European ISM band carrier frequency in order for the coaxial impedance adapter to operate in its operating bandwidth.

First, functional tests were performed to verify correct operation. Then, systematic measurements were accomplished to characterize the MDR and the FAR. Finally, the capability of the WUR to receive long sequences of data was investigated and the performance of the BC circuit was analyzed. Figure 4.49 shows the sample measured waveforms in response to a packet composed of a 3-bit preamble (100) followed by a 16-bit string matching the stored codeword (1011101101010011). This measurement was performed with a -34 dBm RF input sequence at 1 kbit/s, with a 0.5% clock frequency error measured after calibration. The ED correctly outputs the envelope of the incoming RF signal, the generated clock correctly samples DDin and the BBL correctly generates the Wake-Up pulse.

The AFE sensitivity, defined as the input power yielding a 10^{-3} BER, has been measured around -36 dBm. MDR measurements have been carried out to find the overall system sensitivity. Moreover, in order to investigate the impact of the GO-CDR on the sensitivity of the WUR, an additional Nucleo (see Figure 4.48), synchronized and running in parallel with the main one, was employed to decode the AFE output with an external precisely timed clock and to compare the received stream with the one transmitted by the main Nucleo. MDR measurements with different codewords featuring different N_m carried out with the external clock found the sensitivity to be -36.25 dBm both tolerating no errors and tolerating up to two in each received packet.

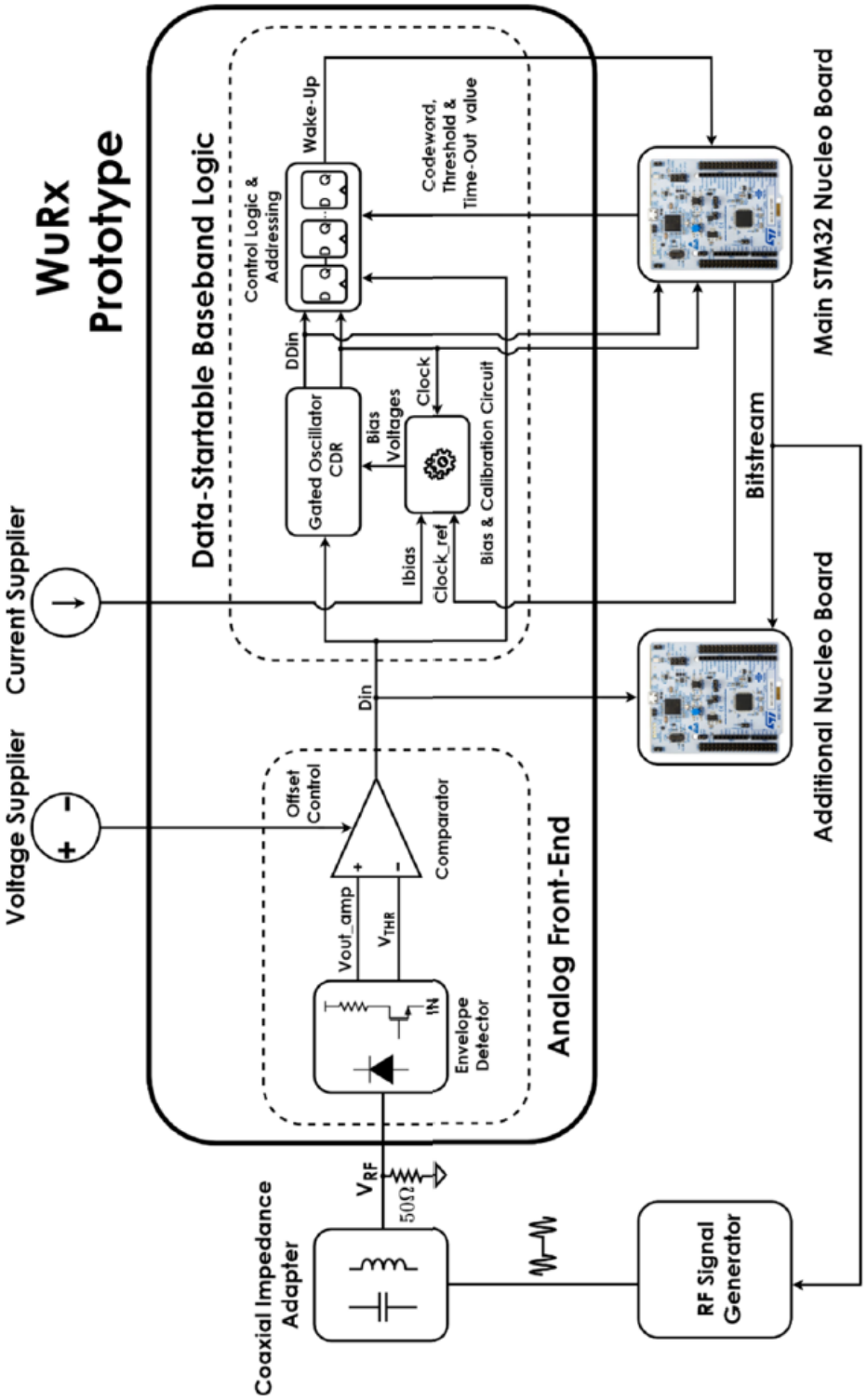

Figure 4.48: Measurement setup. Picture reused from [23].

Figure 4.49: Measured sample waveforms. With reference to Figure 4.39, from top to bottom: ED output v_{out_amp}, DDin, Clock and wake-up. Picture reused from [23].

When using the internal GO-CDR clock, the overall sensitivity has been measured to -35.75 dBm tolerating no errors in received packets, as shown in Figure 4.47, and -36 dBm tolerating up to two errors. Therefore, MDR measurements show the degradation of sensitivity due to the GO-CDR is minimal.
FAR measurements were also carried out throughout 24 hours with the input of the coaxial adapter closed on a 50-Ω resistance and resulted in zero overall false wake-ups. To evaluate the WUR capability to receive long sequences of data, additional MDR measurements were performed sending 63-bit packets with consecutive "1"s. Sensitivity was found at -35 dBm tolerating no errors and -35.5 tolerating 5 errors, again implying the degradation of sensitivity due to the GO-CDR is minimal.

The measured performances of this prototype are summarized in Tables 4.5 and 4.6. This results in a 116.2-dB FoM at 433 MHz, implying a roughly 10-dB improvement with respect to the previous prototype, mainly thanks to power consumption reduction and an enhanced sensitivity.
The two measured prototypes have performances comparable to some of the prior art. However, some solutions have higher performances, most likely due to different reasons. For instance, [8], [9] and [13] use Dynamic Threshold-voltage MOS (DTMOS) transistors, which are non-standard devices where the bulk is connected to the gate, thus boosting transconductance and achieving better gain. Moreover, [8], [13], [14], [16], [17], [18] use very low bitrates, which allow a drastic reduction in the integrated noise caused by the ED, thus improving sensitivity. [8] and [16] use low carrier frequencies, which allow higher matching gain values, which directly add to sensitivity. Finally, [10], [11], [14]-[20] use a passive ED, which inherently yields a better sensitivity as well, but, as explained in Section 3.3.2 and further discussed in

Table 4.5: Performance comparison between the second proposed prototype and several prototypes in literature which had already been presented at the time of design of the second prototype, part 1.

	Proposed prototype 1	Proposed prototype 2	[9], 2012	[10], 2013	[11], 2016	[6], 2016	[12], 2017	[13], 2017
Temperature range [°C]	-20 : 60	unknown	unknown	unknown	unknown	unknown	unknown	unknown
Carrier frequency [MHz]	771	433	915	915	2400	868	2400	403
Matching gain [dB]	12	17.3*	12	5	5	unknown	12.5	18.5
Bitrate [$kbit/s$]	2	1	100	12.5	8.192	1	200	0.3
Sensitivity at 27°C [dBm]	-46[a]	-52.3[*,b]	-41[a]	-43[b]	-56.5[b]	-55[b]	-50[a]	-63.8[b]
Supply voltage [V]	1.2	0.6	1.2	1.2/0.5	1/0.5	1.2	0.8	0.4
Power at 27°C [nW]	36	12.9[**]	98	116	236	1200	4500	4.5
Correlator	no	yes	no	yes	yes	yes	no	yes
Code length	N.A.	63	N.A.	31	31	16	N.A.	32
Technological node [nm]	90	90	130	130	65 low-power	N.A.	180	180
Die area [mm^2]	0.7	0.37	0.03	0.35	2.25	N.A.	0.456	4.5 (estimated)
External components	LC	LC*	LC	LC	LC	N.A.	LC	transformer
FoM_1 [dB]	106.9	116.2	106.1	102.8	112.3	99.2	100	129.7
FoM_2 [dB]	100.1	105.6	101.4	unknown	unknown	100	102.8	115.9

* LC matching network simulated: sensitivity projected as sum of measured AFE sensitivity and simulated matching network gain
** Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time
[a] Sensitivity defined through 10^{-3} Bit Error Rate
[b] Sensitivity defined through 10^{-3} Missed Detection Rate

Table 4.6: Performance comparison between the second proposed prototype and several prototypes in literature which had already been presented at the time of design of the second prototype, part 2.

	[8], 2018	[14], 2019	[15], 2019	[16], 2019
Temperature range [$°C$]	unknown	unknown	unknown	unknown
Carrier frequency [MHz]	113.5	434	750	152
Matching gain [dB]	25	26	13	27
Bitrate [$kbit/s$]	0.3	0.1	200	0.2
Sensitivity at 27$°C$ [dBm]	-69[b]	-79.2[b]	-50[b]	-76[b]
Supply voltage [V]	0.4	0.4	0.4	1/0.6
Power at 27$°C$ [nW]	4.5	0.42	1486[**]	7.4
Correlator	yes	yes	yes	yes
Code length	16	11	40	8
Technological node [nm]	180	65 low-power	65	130
Die area [mm^2]	4.5 (estimated)	0.05 (estimated)	0.75	1.95
External components	transformer	LC	L	LC
FoM$_1$[dB]	134.9	153	104.8	138.8
FoM$_2$[dB]	121.1	134.1	103.7	unknown

[**] Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time

[a] Sensitivity defined through 10^{-3} Bit Error Rate

[b] Sensitivity defined through 10^{-3} Missed Detection Rate

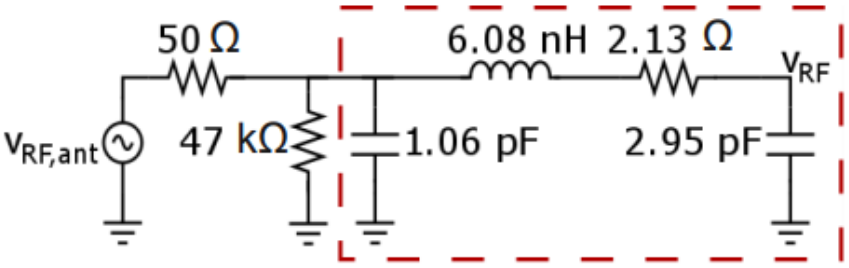

Figure 4.50: Input impedance lumped element model of the ED at the SMA connector. $v_{RF,ant}$ indicates the voltage supplied by the RF generator while v_{RF} is the input voltage of the ED. Picture reused from [23].

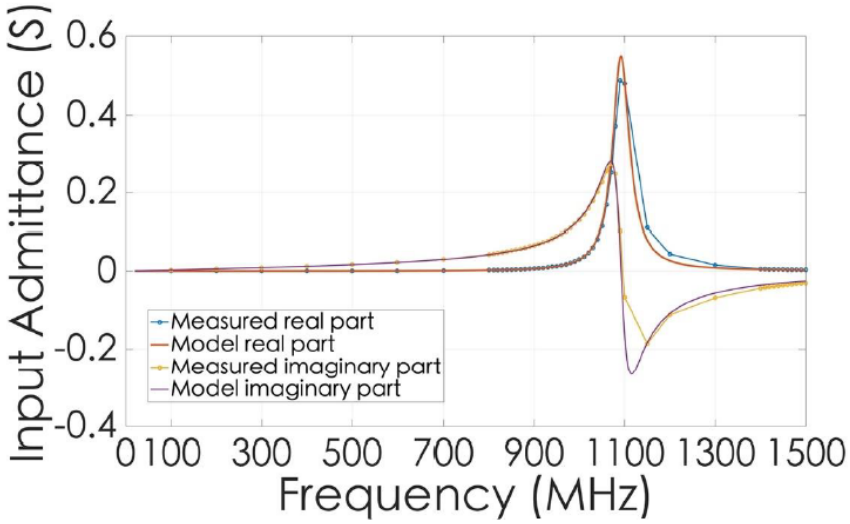

Figure 4.51: Input admittance vs. frequency. Blue: measured real part of the input admittance, Orange: simulated real part of the input admittance using the model in Figure 4.50. Picture reused from [23].

Chapter 5, they cause high propagation delays, which typically makes only bitrates lower than 1 kbit/s possible. Although a few examples of passive EDs with higher bitrates than 1 kbit/s are present in literature, passive EDs are not suitable for bitrate programmability.

Matching Network

The input impedance at the SMA connector has been characterized by means of a Vector Network Analyzer in the 10 MHz–1.5 GHz range, as shown in Figure 4.51. The resonance frequency is clearly visible around 1.1 GHz and is most likely due to the wire inductance and the input capacitance, respectively 6.08 nH and 2.95 pF in the equivalent lumped element circuit which has been fitted to the measurement results, Figure 4.50. In particular, the 2.95-pF capacitance is way higher than the one extracted from post-layout simulations, roughly 1.7 pF, and is most likely due to the parasitic capacitance of the pad, a standard analog one. In a final implementation,

such a pad would need to be replaced by a low-capacitance RF one. Due to these limitations, the matching network has not been implemented for this prototype, as its gain would be much smaller than the expected one.

However, for the sake of completeness, *LC* matching networks for different carrier frequencies have been designed using the extracted input impedance lumped element model in Figure 4.50 to estimate the obtainable matching gain. Assuming the use of inductances with $Q = 80$, the simulated matching gain values at 100 MHz, 433 MHz and 868 MHz are 24.8 dB, 17.3 dB and 14.3 dB, respectively. These gain values are to be added to the measured circuit sensitivity to get the projected overall WUR sensitivity, leading to an overall projected system sensitivity of -59.8 dBm, -52.3 dBm and -49.3 dBm at 100 MHz, 433 MHz and 868 MHz, respectively.

4.9 Conclusions on active Envelope Detectors

Chapter 4 presents two implementation of an AFE for an integrated medium range WUR employing an active ED. Active EDs perform both envelope extraction, i.e. OOK demodulation, and baseband amplification of the extracted envelope. The first implementation features an active ED with a BP response, whereas the second one an active ED with a LP response.

In the BP case, the ED output peaks at each transition in the input bit. This forces the subsequent decision circuit to be hysteretic, i.e. with two thresholds. These thresholds need to be generated by a dedicated block, increasing the complexity of the system. Moreover, the risk of multiple errors in the event of long sequences of equal consecutive bits is to be avoided making use of a coding with frequent transitions, possibly Manchester coding. In turn, this causes the halving of the bitrate. However, a system employing a BP ED requires no off-chip capacitances, as the capacitance needed in the ED can easily be integrated.

In the LP case, the ED outputs a signal with the same shape as the input envelope, leading to the use of a standard comparator instead of a hysteretic one. Only one threshold is needed and it can be generated inside the ED itself. This leads to a simpler system. Moreover, there is no risk of multiple errors in the event of long sequences of equal consecutive bits, which implies Manchester coding is not a requirement. However, there is a maximum packet length and a minimum packet spacing and an off-chip capacitance is needed.

Chapter 5

Passive Envelope Detector implementation

This chapter presents the features and advantages of passive EDs, which are employed in the third implemented prototype presented in this thesis.
The third prototype was taped out in 2021 and includes both the AFE, partly designed by another Ph.D. student, and the BBL, very similar to that of the second prototype and again designed by another Ph.D. student. The AFE is very different from those of the two previous versions, since it features a passive ED instead of an active ED. The employed ED architecture is based on the topology recently published in [14] with the original addition of a block performing temperature compensation of the ED and the matching network, in order to reduce sensitivity temperature dependence. Moreover, a differential approach has been adopted to prevent SNR degradation due to the presence of the additional block. The baseband amplification of the extracted envelope is thus performed by a separate block. The comparator is again a standard one and its threshold is generated by an RC filter. This is a very simple solution but it forces the use of coding featuring a low number of equal consecutive bits, like in [6].

5.1 ED structure

As usual, the first stage of the AFE is the ED, whose purpose is to extract the envelope of the input OOK-modulated signal. As discussed in Section 4.2, active EDs have an unfavorable current-to-sensitivity relationship, Equation 4.16. As a result, with active EDs a non-negligible improvement in sensitivity requires a much higher current consumption. On the other hand, passive EDs solve this issue.
As anticipated in Section 3.3.2, the simplest passive ED is a standard Dickson charge pump, shown in Figure 3.6 and redrawn in Figure 5.1. Actually, since MOSFETs are symmetrical, the drain and source terminals are interchangeable and terminal names were assigned in Figure 5.1 and all the subsequent figures for the sake of simplicity.

As explained in Chapter 3, both passive and active EDs leverage the second-order non-linearities generated in subthreshold currents. However, passive EDs have two major advantages over active ones:

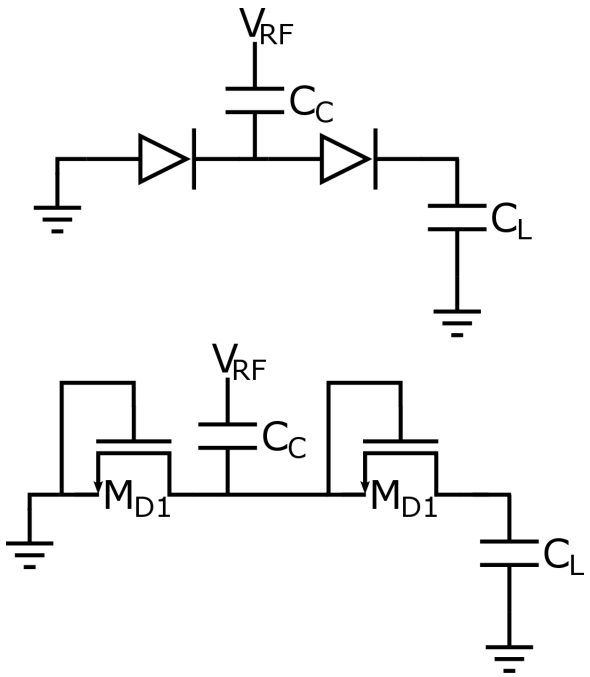

Figure 5.1: Dickson charge pump.

- zero power consumption
- no flicker noise

Passive EDs feature transistors biased at zero DC current, which results not only in a zero power consumption but also in the native absence of flicker noise. The reason lies in the fact that the flicker noise current of a MOSFET is linked to its DC current through its g_m [31]:

$$\frac{\overline{i_n^2}}{\Delta f} = \frac{K_f g_m^2}{WL\mu C_{ox} f^{\alpha}} \tag{5.1}$$

where K_f is the flicker noise coefficient, W and L are the MOSFET dimensions, μ is the carrier mobility, C_{ox} is the gate oxide thickness and $\alpha \sim 1$. The inherently lower noise shown by this type of ED directly translates into a better sensitivity than those of active EDs.

A variation of the Dickson charge pump was presented in [14]. This variation yields the same output amplitude as the standard Dickson charge pump in Figure 5.1, but allows the biasing of diode gates regardless of the input signal, which is very useful for sensitivity optimization, as will be clear in Section 5.2. This ED architecture is composed of a chain of diode-connected MOSFETs, whose basic element features two diodes in two different configurations in series, as shown in Figure 5.2. In the first configuration the input RF signal is applied only to the drain of M_{D1}, whereas in the second one the input RF signal is applied both to the source and the gate of M_{D2} through coupling capacitances C_{gate} and C_{drain}. Capacitance C_{gnd} filters the RF component at the output. The output useful signal is the drain-to-source DC voltage generated by the second-order non-linearities of the MOSFET.

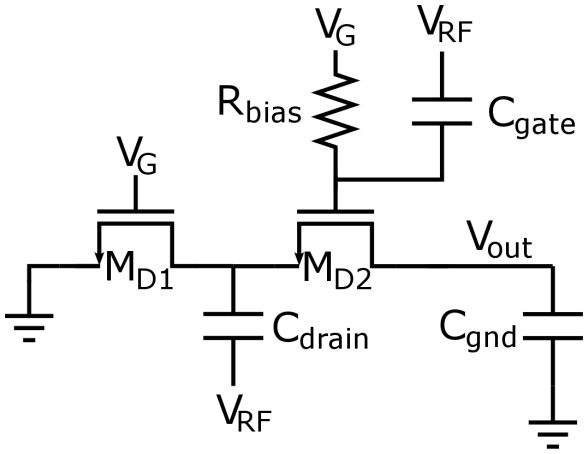

Figure 5.2: Modified Dickson charge pump as presented in [14]

All capacitance values are thus heavily dependent on the choice of carrier frequency. In both configurations, the DC voltage at the gates of M_{D1} and M_{D2} can be set. The input of the ED is again $v_{RF}(t) = v_M(t)\cos(\omega t)$, where $v_M(t)$ is the envelope, whereas the unknown output voltage is the mean drain-to-source voltage V_{DS0}. Since the MOSFETs are biased at zero current, they operate in subthreshold, resulting in a current during operation

$$I_D = I_S \frac{W}{L} e^{\frac{V_{GB}}{nV_T}} \left[e^{-\frac{V_{SB}}{V_T}} - e^{-\frac{V_{DB}}{V_T}} \right], \tag{5.2}$$

where $I_S = I_{S0}\, e^{-\frac{V_{th}}{nV_T}}$. I_{S0} is the reverse saturation current of a p-n junction, related to technology, and V_{th} is the threshold voltage of the MOSFET. Considering the first configuration, M_{D1}, and $V_{SB} = 0$, this can be rewritten as

$$I_D = I_S \frac{W}{L} e^{\frac{V_{G0}}{nV_T}} e^{-\frac{V_{S0}}{V_T}} \left[1 - e^{-\frac{V_{DS0}+v_{RF}}{V_T}} \right] = K \left[1 - e^{-\frac{V_{DS0}+v_{RF}}{V_T}} \right], \tag{5.3}$$

where K is constant with respect to the useful output signal. $\left[1 - e^{-\frac{V_{DS0}+v_{RF}}{V_T}} \right]$ may be written in Taylor series as

$$1 - \left(1 - \frac{V_{DS0} + v_{RF}}{V_T} + \frac{1}{2} \left(\frac{V_{DS0} + v_{RF}}{V_T} \right)^2 \right). \tag{5.4}$$

Truncating at the second order has proven effective to accurately model the behavior of this circuit while using simple math. Since $\cos^2 \alpha = (1 + \cos 2\alpha)/2$,

$$1 - \left(1 - \frac{V_{DS0} + v_{RF}}{V_T} + \frac{1}{2} \left(\frac{V_{DS0} + v_{RF}}{V_T} \right)^2 \right) = \frac{V_{DS0}}{V_T} + \frac{v_M \cos \omega t}{V_T} - \frac{V_{DS0}^2}{2V_T^2} - \frac{V_{DS0} v_M \cos \omega t}{2V_T^2} - \frac{v_M^2}{4V_T^2} - \frac{v_M^2 \cos 2\omega t}{4V_T^2}. \tag{5.5}$$

The time average of this quantity must be zero as the RF is supposed to get filtered out, resulting in

$$< I_D >= 0 => K\left[1 - e^{-\frac{V_{DS0}+v_{RF}}{V_T}}\right] = 0. \tag{5.6}$$

$K = 0$ is not useful as K is independent of the output signal, whereas $1 - e^{-\frac{V_{DS0}+v_{RF}}{V_T}} = 0$ leads to a solution. Assuming $V_{DS0} \ll V_T$, which is the case for most plausible input power levels and especially at the sensitivity, the result is

$$V_{DS0} \cong \frac{v_M^2}{4V_T}. \tag{5.7}$$

The same procedure can be applied to the second configuration, M_{D2}, where the gate voltage is $V_G = V_{G0} + v_{RF}$ instead. All other hypotheses hold. In this case,

$$V_{DS0} \cong \frac{(2-n)\, v_M^2}{4nV_T}. \tag{5.8}$$

Among all diode stages, half are in the first configuration and half in the second one. Therefore, the average contribution of a diode stage is the average of the contributions of the two configurations, Equations 5.7 and 5.8:

$$V_{DS0} = \frac{\frac{v_M^2}{4V_T} + \frac{(2-n)v_M^2}{4nV_T}}{2} = \frac{v_M^2}{4nV_T}. \tag{5.9}$$

The diode stages appear in parallel with respect to the RF input but in series with respect to the output baseband signal, resulting in each stage adding its contribution to that of the preceding stages. This implies the total output signal is

$$v_{ed,o} = \frac{Nv_M^2}{4nV_T}, \tag{5.10}$$

where N is the number of diode stages.
At RF, the ED can be seen as a load composed of a resistance and a capacitance in parallel, R_{in} and C_{in}, for the preceding matching network. This makes matching gain A_v heavily dependent on the values of R_{in} and C_{in}. Assuming C_{gnd} and C_{drain} are big enough with respect to the C_{gs} of the MOSFETs and they have negligible parasitics, R_{in} is the parallel of all the equivalent resistances seen at the inputs connected to C_{drain}, yielding

$$R_{in} = \frac{r_{DS}}{N}, \tag{5.11}$$

whereas C_{in} is the parallel of the C_{gs} of all the diodes and the pad capacitance C_{pad}, that is

$$C_{in} = NC_{gs} + C_{pad}, \tag{5.12}$$

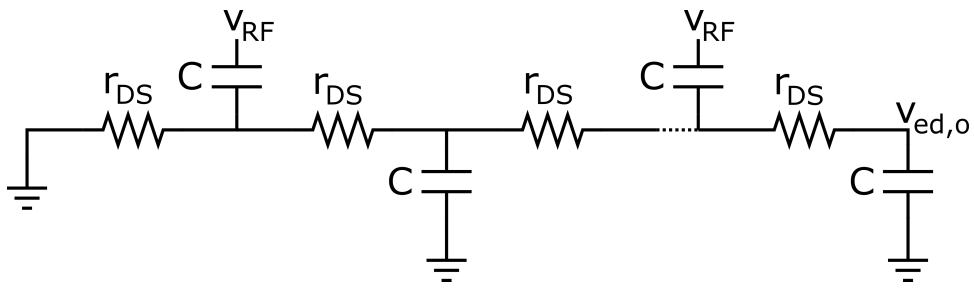

Figure 5.3: Estimation of the propagation delay generated by the ED in Figure 5.2

due to the filtering capacitances being shorts at RF. This does not hold in case C_{gate} and C_{drain} are of the same order as C_{gs} and have non-negligible parasitics, in which case a much more complex model should be used to determine the analytical expression of R_{in} and C_{in}. Indeed, it is more convenient to perform simulations to determine R_{in} and C_{in} numerically. On the other hand, the output resistance, which is defined at baseband, is the series of the channel resistances of all diodes, that is

$$R_{out} = Nr_{DS}. \tag{5.13}$$

5.1.1 Propagation delay estimation

In order to estimate the delay introduced by the ED, that is the cut-off frequency of the ED, Figure 5.3 may be considered, assuming all capacitances to have the same value.

Applying the open circuit time constant method, the delay time constant of the circuit can be estimated as

$$\tau \cong Cr_{DS}\left[N + (N+1) + ... + 2 + 1\right] \sim Cr_{DS}\frac{N\left(N+1\right)}{2} = CR_{in}\frac{N^2\left(N+1\right)}{2}. \tag{5.14}$$

Therefore, τ is almost proportional to the cube of N, which results in any increase in N having a strong impact on the delay introduced by the circuit.

5.1.2 Loss of signal due to non-idealities

Detailed study of this architecture showed there are two different sources for loss of signal within the ED with respect to the theoretical value of $v_{ed,o}$. The first one is MOSFET junction leakage whereas the second one is non-perfect coupling of the input signal v_{RF} to the internal gate node in the second configuration. This phenomenon directly translates into a reduction in the gain of the ED, which was modelled as a multiplication of the theoretical gain by a factor $k < 1$, yielding

$$v_{ed,o} = \frac{kNv_M^2}{4nV_T}. \tag{5.15}$$

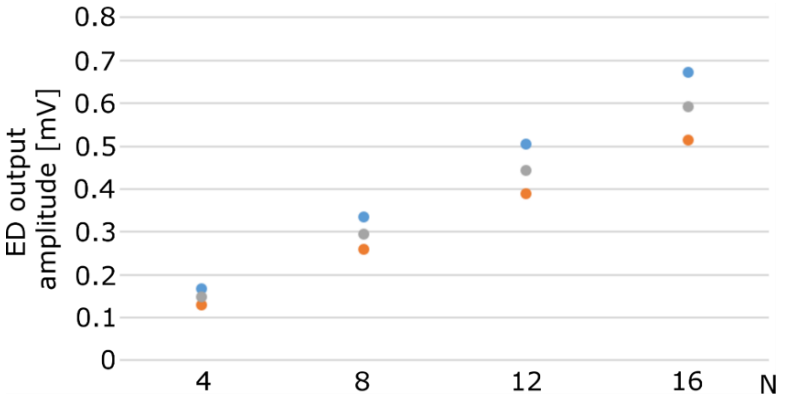

Figure 5.4: ED output amplitude in theory (blue), using ideal 1-pF capacitances (gray) and using real 36-fF capacitances (orange) as a function of the number of diode stages. This graph clearly shows the loss of signal due to imperfect signal coupling and to MOSFET leakage.

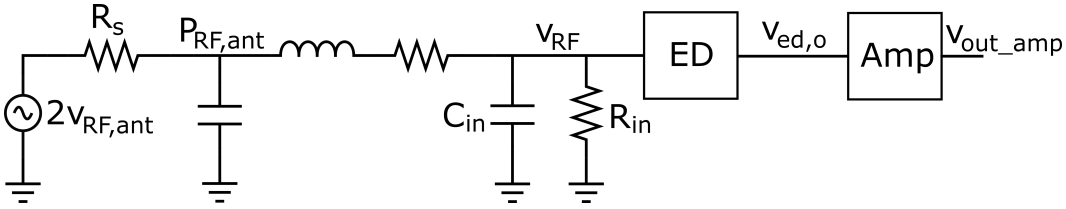

Figure 5.5: Simplified schematic to calculate sensitivity as presented in [14].

This loss appears to be independent of the number of diode stages N. This was proven by simulating the ED with different values for N first with arbitrarily big C_{gate} and C_{drain} and with the actual chosen values for the same capacitances. Figure 5.4 shows the theoretical amplitude of the ED output in blue, the amplitude with ideal 1-pF capacitances in gray and using real 36-pF capacitances in orange. The losses between the blue and the gray points are due to MOSFET junction leakage whereas those between the gray and the orange points are due to the non-perfect coupling. In this implementation roughly half of the loss seems to be due to one cause and half to the other cause. In order to estimate the ultimate value for constant k it is sufficient to multiply the reductions due to each of the causes.

5.2 Sensitivity optimization

Let's assume the input signal to be provided by an RF generator. Assuming the matching condition to be met, an input power $P_{RF,ant} = v^2_{RF,ant}/2R_S$, where $R_S = 50\ \Omega$ is the resistance of the source, as shown in Figure 5.5.

This implies the voltage amplitude virtually needed within the generator before R_S is 2 $v_{RF,ant}$. The peak-to-peak signal at the output of the matching network is then $v_{RF} = A_v\ v_{RF,ant}$, whereas the peak-to-peak signal at the output of the ED is

$$v_{ed,o} = \frac{NA_v^2 P_{RF,ant} 2R_S}{4nV_T}. \tag{5.16}$$

As mentioned earlier, the diodes are biased at zero DC current. This implies they emit no flicker noise, only thermal white noise. The power spectral density of the thermal noise emitted by diodes is related to their channel resistance, r_{DS}:

$$v_{Ned,o}^2(f) = 4k_B TNr_{DS} = 4k_B TN^2 R_{in}. \tag{5.17}$$

Therefore, the noise power at the output of the ED is $4k_B TN^2 R_{in} f_S$, where f_S is the baseband frequency. The SNR at the ED output is then

$$SNR_{ed,o} = \frac{v_{ed,o}^2}{v_{Ned,o}^2(f) f_S} = \frac{N^2 A_v^4 P_{RF,ant}^2 4R_S^2}{(4nV_T)^2 4k_B TN^2 R_{in} f_S} = \frac{A_v^4 P_{RF,ant}^2 R_S^2}{(4nV_T)^2 k_B TR_{in} f_S}. \tag{5.18}$$

Therefore, the SNR at the ED output does not depend on N, but it heavily depends on R_{in}, as A_v is also a function of R_{in}. Again with reference to Figure 5.5, let's assume the ED is followed by a baseband amplifier with Noise Factor NF and Noise Figure $10\log_{10} NF = NF(dB)$. NF of the amplifier is defined as

$$NF = \frac{SNR_{ed,o}}{SNR_o}, \tag{5.19}$$

where SNR_o is the total SNR at the output of the baseband amplifier. Therefore, it indicates the extent to which the amplifier degrades the SNR of the signal it receives as input. It can also be expressed as

$$NF = 1 + \frac{N_{ampl,o}}{\int G(f)^2 v_{Ned,o}^2(f) df}, \tag{5.20}$$

where $N_{ampl,o}$ is the integrated noise at the amplifier output due to the amplifier itself, G is the amplifier frequency response and $v_{Ned,o}^2$ is as defined above. Therefore, the denominator of Equation 5.19 corresponds to the noise due to the ED referred to the amplifier output. Integration must be carried out over a wide frequency interval, starting from way below f_S ending way above the cut off frequency of G. From Equation 5.19 it is straightforward to conclude that

$$SNR_o = \frac{SNR_{ed,o}}{NF}. \tag{5.21}$$

If the minimum SNR at the output of the baseband amplifier is known and equal to SNR_{req}, the overall sensitivity can be calculated as

$$P_{SEN} = \sqrt{\frac{SNR_{req} NF (4nV_T)^2 k_B TR_{in} f_S}{A_v^4 R_S^2}} \tag{5.22}$$

or in dBm

$$P_{SEN}\,(dBm) = 10\log_{10}\left(\frac{P_{SEN}}{10^{-3}}\right) = \frac{1}{2}\left[SNR_{req}\,(dB) + NF\,(dB)\right] - A_v\,(dB) + 10\log_{10}\left[\sqrt{\frac{(4nV_T)^2\,k_BTR_{in}f_S}{R_S^2}}/10^{-3}\right], \tag{5.23}$$

where $A_v(dB) = 20\log_{10} A_v$. Considering loss of signal due to non-idealities, the final formula is

$$P_{SEN}\,(dBm) = 10\log_{10}\left(\frac{P_{SEN}}{10^{-3}}\right) = \frac{1}{2}\left[SNR_{req}\,(dB) + NF\,(dB)\right] - A_v\,(dB) + 10\log_{10}\left[\sqrt{\frac{(4nV_T)^2\,k_BTR_{in}f_S}{k^2R_S^2}}/10^{-3}\right]. \tag{5.24}$$

In Equation 5.22, the numerator obviously increases with increasing R_{in}. The denominator is related to R_{in} through A_v, as shown in Equation [14]

$$A_v = \sqrt{\frac{R_{in}}{R_S}}/\sqrt{\left(1 + \frac{\omega R_{in}C_{in}}{Q_{ind}},\right)} \tag{5.25}$$

Where ω is the carrier RF frequency and Q_{ind} is the quality factor of the inductor in the matching network, with reference to Figure 5.5. In this equation, the numerator corresponds to the ideal relationship between $v_{RF,ant}$ and v_{RF}, whereas the denominator accounts for the losses due to the finite quality factor of the inductance. A_v then increases with increasing R_{in} but eventually saturates. Therefore, $P_{SEN}\,(R_{in})$ in Equation 5.22 has a minimum.

Figure 5.6 shows $P_{SEN}\,(R_{in})$ at room temperature. Since the total SNR does not depend on N, the sensitivity optimization process does not directly involve this parameter. Indeed, Equation 5.10 shows that $v_{ed,o} \propto N$, resulting in a big value for N to be desirable. Moreover, as shown in Equation 5.23, $NF(dB)$ adds to sensitivity, thus degrading it. Hence, for sensitivity optimization it is best to have a small value for $NF(dB)$, which can be achieved again with a big value for N. Therefore, N can be chosen arbitrarily large, namely the maximum value which allows the desired delay between the input and the output of the ED (more on this later). This leads to the choice of an optimum r_{DS}, $\overline{r_{DS}} = N\overline{R_{in}}$ for sensitivity optimization. Thanks to the possibility of setting the gate-to-source DC voltage of the diodes, that is their gate voltage since their source is at virtual ground, $\overline{r_{DS}}$ can be defined according to the following formula:

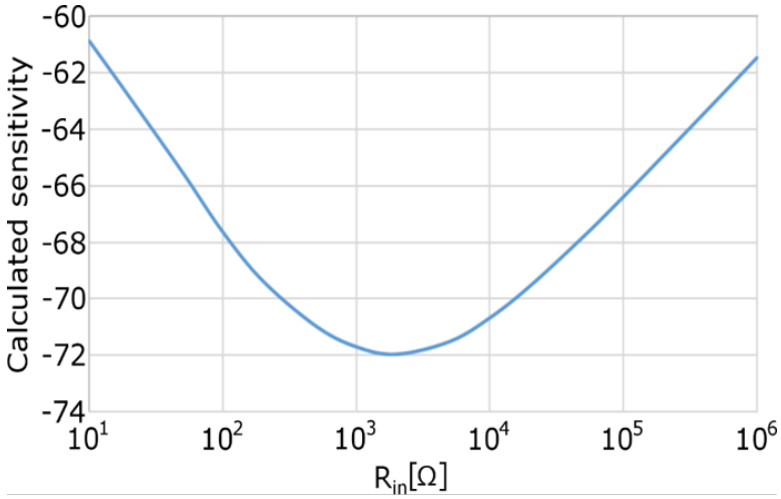

Figure 5.6: P_{SEN} as a function of R_{in} presented in [14].

$$\overline{r_{DS}} = \frac{1}{\overline{g_{DS}}} = \left[\frac{\partial i_{DS}}{\partial v_{DS}}\right]^{-1} = \left[\frac{\partial \left(I_S \frac{W}{L} e^{V_{GS}/nV_T}\left[1 - e^{-V_{DS}/V_T}\right]\right)}{\partial v_{DS}}\right]^{-1} \simeq \frac{V_T}{I_S\,(W/L)} e^{-\frac{\overline{V_{GS}}}{nV_T}}, \tag{5.26}$$

assuming $V_{DS} \geq 4V_T$. Ultimately, by setting the appropriate gate-to-source voltage of the diodes it is possible to set the correct R_{in} for sensitivity optimization.

5.3 Sensitivity temperature compensation

Up to this point, the discussion presented in [14] has been reported. The proposed prototype also implements temperature compensation of the ED and the matching network, in order to reduce sensitivity temperature dependence with an additional block.

As mentioned earlier, sensitivity depends on R_{in}, which is in turn related to r_{DS} of the diodes. As shown in Equation 5.22, $P_{SEN} \propto \sqrt{T^3 R_{in}}$. Therefore, in order to reduce the temperature dependence of P_{SEN}, the removal of the temperature dependence of r_{DS}, thus R_{in}, was implemented. Since r_{DS} is mainly determined by the gate-to-source voltage of the diodes, a dedicated block, shown in Figure 5.7 and again similar to a Brokaw cell [30], was designed to provide the diodes with such a voltage in an appropriate way. All transistors within this block are biased in subthreshold and thus show an exponential IV characteristic. Thanks to the fact that the diodes have the same V_{GS} as M_4, they have roughly the same V_{th} and I_{S0}, resulting in

$$g_{DS} = \frac{I_S\,(W/L)}{V_T} e^{\frac{V_{GS}}{nV_T}} \simeq \frac{I_S\,(W/L)}{V_T} e^{\frac{V_{GS4}}{nV_T}} \simeq \frac{(W/L)}{(W/L)_4} \frac{I_{DC4}}{V_T} \tag{5.27}$$

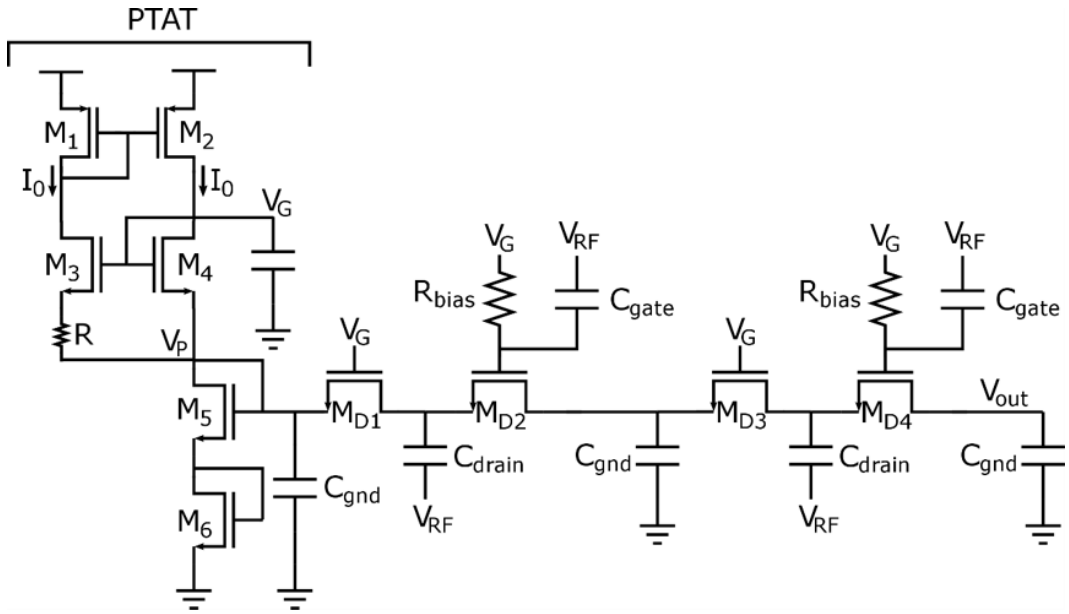

Figure 5.7: ED equipped with a dedicated block for the temperature compensation of sensitivity.

Since the block dedicated to the generation of the V_{GS} of the diodes is very similar to a Brokaw cell and M_3 and M_4 have an exponential characteristic, the current through M_3 and M_4 is PTAT and in particular

$$I_{DC4} = \left(\frac{W}{L}\right)_4 \frac{nV_T \ln \frac{\left(\frac{W}{L}\right)_3}{\left(\frac{W}{L}\right)_4}}{R}, \tag{5.28}$$

yielding

$$g_{DS} \simeq \frac{(W/L)}{(W/L)_4} \frac{n \ln \frac{\left(\frac{W}{L}\right)_3}{\left(\frac{W}{L}\right)_4}}{R} \tag{5.29}$$

Equation 5.29 proves the conductance of the diode stages is independent from temperature. Moreover, since the source terminal of M_{D1}, which constitutes the beginning of the diode chain, is connected to node V_P and the chain is biased at zero DC current, $v_{ed,o}$ corresponds to V_P when $v_M(t) = 0$. This implies the input DC voltage of the amplifier is V_P itself. To avoid biasing the amplifier at a low gain operating point due to a low V_P, two stacked transistors connected in the diode configuration, M_5 and M_6, were employed. Finally, this dedicated PTAT block can act as a current generator for the rest of the AFE.

5.3.1 PTAT start-up

Notably, all circuits which have a stable or metastable DC operating point corresponding to zero current require a start-up circuit. Its aim is to push the main circuit

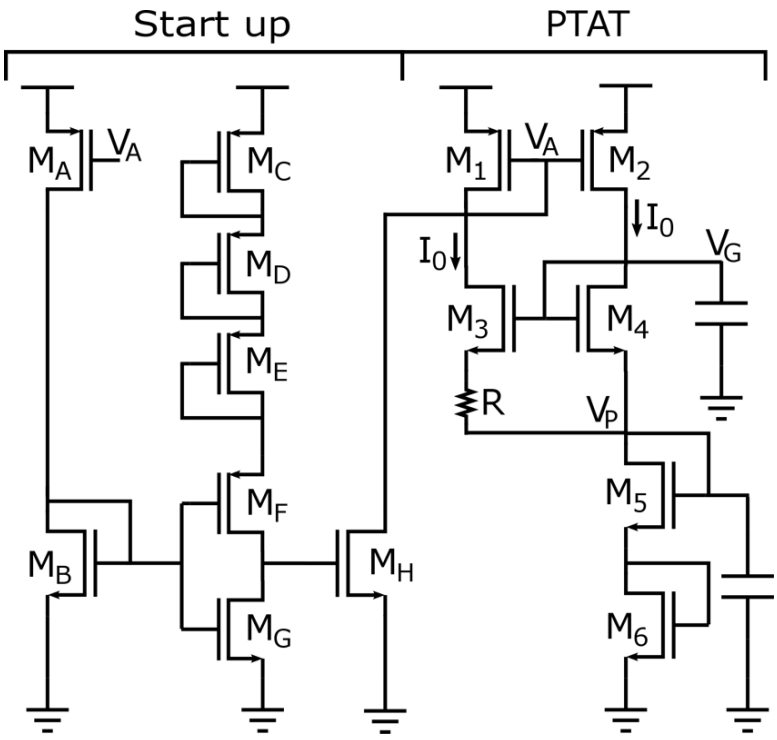

Figure 5.8: PTAT block equipped with a start up circuit.

away from the zero-current DC operating point and push it to the actual DC operating point it was designed for. The start-up circuit implemented is shown in Figure 5.8. Its operation is based on the fact that M_A shares both the gate and the source with M_1. Therefore, in the event of the PTAT block being stuck at the zero-current operating point, neither M_1 nor M_A conduct any current, resulting in zero voltage across M_B. M_C, M_D and M_E act as a shifter for the supply voltage, making M_F and M_G an effective inverter for the voltage at the gate and drain of M_B. When there is no voltage across M_B, M_F turns on and pulls its drain to the same voltage as its source, which in turn causes M_H to turn on as well. The small current sinking into M_H forces the conduction of M_1. Current I_0 then starts flowing and voltage V_A decreases. This brings the voltage across M_B up and turns on M_G to force the gate of M_H to ground and turn M_H off. At the end of this process, the PTAT is in the correct operating point.

5.3.2 ED differential approach

The addition of the PTAT block causes its own noise, which includes both thermal and flicker, unlike that of the ED, to propagate through the signal chain. Most of the noise due to the PTAT block is actually flicker and is due to M_3 and M_4. In order to prevent a heavy degradation in the achievable sensitivity, a differential approach was chosen: the implemented ED features two chains of diodes connected to the PTAT cell instead of one, as shown in Figure 5.9.

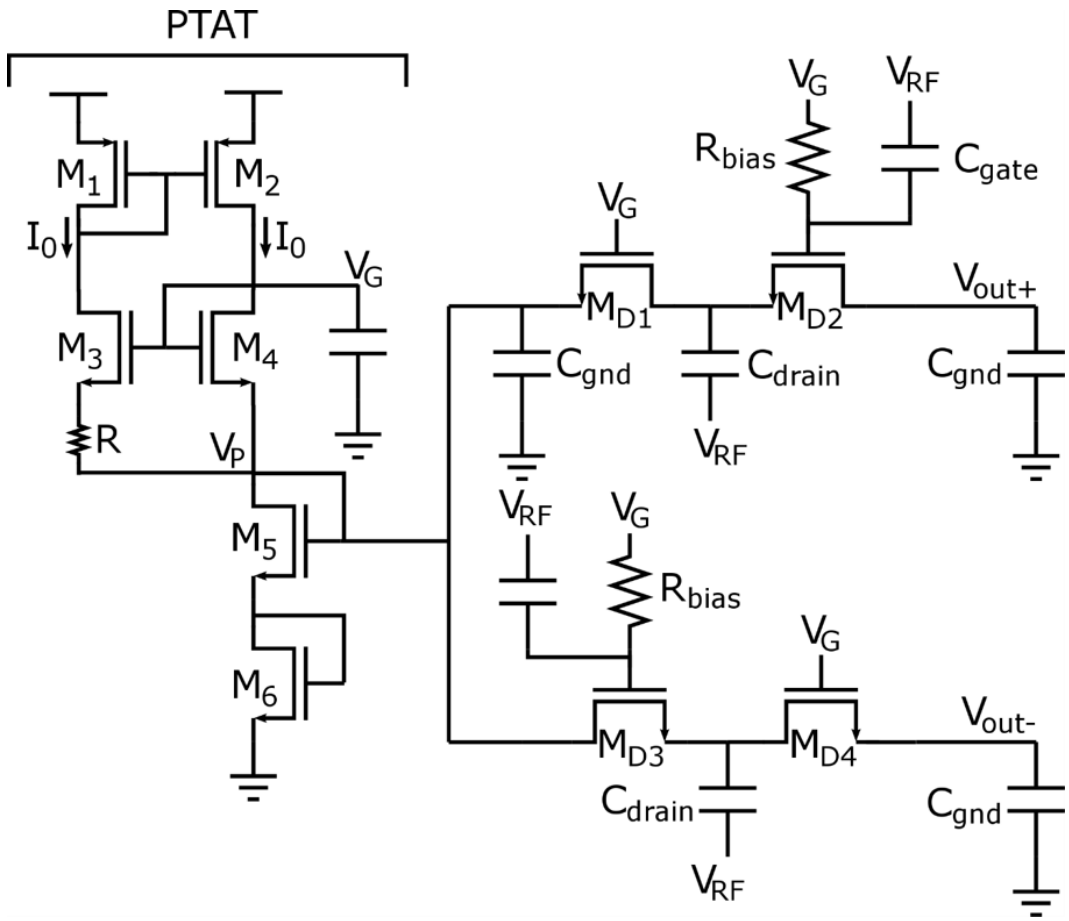

Figure 5.9: Differential ED equipped with a PTAT block.

The noise due to the PTAT cell is then seen as common-mode at the output of the ED and thus gets cancelled out. The new circuit has a differential output signal of

$$v_{ed,o} = \frac{2kNv_M^2}{4nV_T}, \tag{5.30}$$

an input resistance $R_{in} = r_{DS}/2N$ and an input capacitance $2NC_{gs} + C_{pad}$, which yields the same SNR at the ED output as in the single-ended case. The latter is proved in the following:

- Let's consider the *single-ended ED* first. As thoroughly described in Section 5.2, the input power is $P_{RF,ant} = v_{RF,ant}^2/2R_S$ and the signal amplitude after the matching network is $v_{RF} = A_v\, v_{RF,ant}$, whenever the matching condition is met. The input resistance is $R_{in} = r_{DS}/N$. The output amplitude is $v_{ed,o} = Nv_M^2/4nV_T$, whereas the ED output noise power spectral density is $v_{Ned,o}^2(f) = 4k_BTN^2R_{in}$. The resulting SNR of a single-ended ED is

$$SNR_{ed,o} = \frac{v_{ed,o}^2}{v_{Ned,o}^2(f)\, f_S} = \frac{A_v^4 P_{RF,ant}^2 R_S^2}{(4nV_T)^2\, k_BTR_{in} f_S}, \tag{5.31}$$

  which corresponds to Equation 5.18.

- In case of a *differential ED*, the input power and the signal amplitude after the matching network are the same as in the single-ended case. On the other hand, the input resistance is $R_{in} = r_{DS}/2N$, the ED output amplitude is $v_{ed,o} = 2Nv_M^2/4nV_T$ and the ED output noise power spectral density is $v_{Ned,o}^2(f) = 2 \times 4k_BTN^2 2R_{in} = 4 \times 4k_BTN^2R_{in}$. This yields an SNR of

$$SNR_{ed,o} = \frac{v_{ed,o}^2}{v_{Ned,o}^2(f)\, f_S} = \frac{16N^2A_V^4P_{RF,ant}^2R_S^2}{(4nV_T)^2\, 16k_BTN^2R_{in}f_S} = \frac{A_V^4P_{RF,ant}^2R_S^2}{(4nV_T)^2\, k_BTR_{in}f_S}, \tag{5.32}$$

  which is the same as for the single-ended ED case.

Moreover, switching to a differential approach causes no increases in power consumption due to the zero-current biasing of the diodes. Since the two chains work independently, the system does not get slowed down either. Finally, let's assume the ED is followed by a standard amplifier, for instance a common-source one, with no residual flicker noise and transconductance g_m. It shall be proven that the noise factor is the same when adopting a differential approach as it is in the single-ended case, with reference to Figures 5.10 and 5.11.

If the ED and the amplifier were single-ended, Equation 5.20 would result in

$$NF_{SE} = 1 + \frac{2 \times 4k_BTg_mR_{out}^2f_S}{g_mR_{out} \times 4k_BTN^2R_S} = 1 + \frac{2}{g_mN^2R_S} \tag{5.33}$$

On the other hand, in the differential case

$$NF_{diff} = 1 + \frac{4 \times \frac{4k_BT}{\frac{g_m}{2}} \left(\frac{g_mR_{out}}{2}\right)^2 f_S}{\frac{g_mR_{out}}{2} \times 4k_BT\,(2N)^2\, R_S} = 1 + \frac{2}{g_mN^2R_S}. \tag{5.34}$$

This shows that no degradation occurs in the overall SNR at the output of the amplifier by adopting the differential approach.

5.4 ED design phase

5.4.1 Choice of component types

The ED, shown in Figure 5.9, includes three types of components: MOSFETs, resistors and capacitors. The choice of MOSFET type is based mainly on the need for ED gain maximization. Since ED gain is proven to be inversely proportional to the non-ideality factor n of the MOSFET performing demodulation by Equation 5.10 ($v_{ed,o} = \frac{Nv_M^2}{4nV_T}$), it is best to choose the MOSFET type featuring the lower n available within the chosen technology. As for the MOSFETs constituting the PTAT block, it may be necessary to choose the MOSFET type with the lower threshold depending on the supply voltage.

The choice of resistor type is based on two main reasons. First, as the current levels

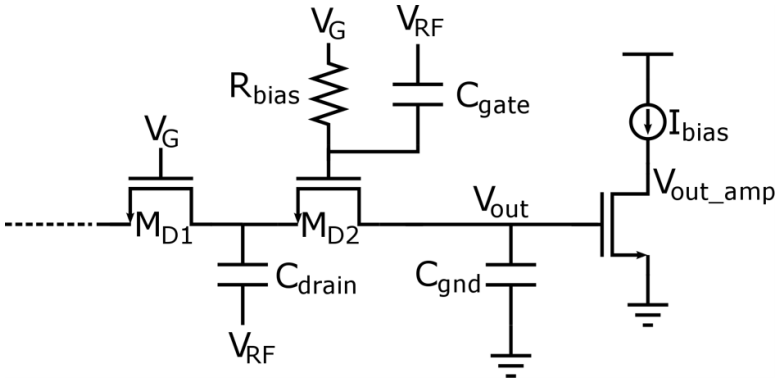

Figure 5.10: Single-ended ED followed by a conventional common-source amplifier.

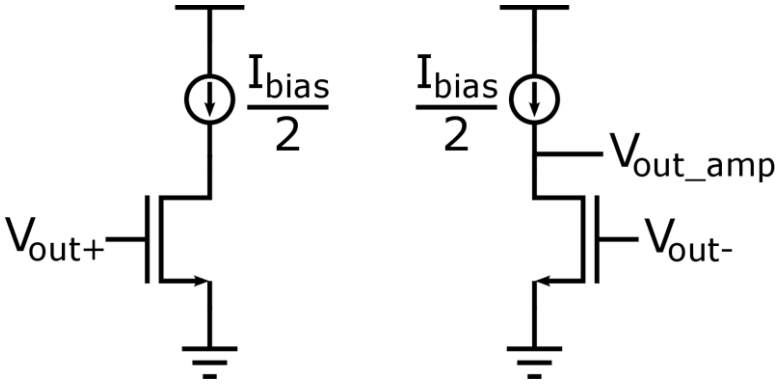

Figure 5.11: Differential ED followed by a differential common-source amplifier.

considered in this implementation are in the order of nanoAmpères, it is necessary to use MΩ resistors in order to generate non negligible voltages across them. Therefore, the higher resistivity available is desirable. Second, the operation of the PTAT block is based on the fact that its current, in Equation 5.28 ($I_{DC4} \propto V_T$), is actually proportional to temperature. Any additional temperature dependence due to the resistor would be unwanted, which leads to choosing the resistor type with the lower temperature coefficient available.

Finally, the choice of capacitor type has important implications mainly due to coupling capacitors C_{gate} and C_{drain}. Of course, losses may also impact the filtering efficiency of C_{gnd}, but simulation models for capacitors in the employed technology do not include losses at all. Instead, any possible parasitic capacitance associated with C_{gate} and C_{drain} may cause capacitive dividers resulting in loss of signal at the MOSFET terminals. Therefore, it is necessary to carefully assess the impact of the type and polarity of the chosen capacitors in terms of loss of output signal amplitude and increase in the overall input capacitance of the ED C_{in}. This entails sensitivity estimation by the use of Equation 5.24 using different types of capacitor with different polarity, on equal terms, to determine which case causes the lower degradation in

the sensitivity itself.

5.4.2 Preliminary choice of component sizes

As mentioned in Section 5.1, it is desirable to have big N for high output amplitude and low amplifier NF, possibly in the order of several tens. Then, it is possible to make an estimation of the optimum ED input resistance, $\overline{R_{in}}$, assuming R_{in} and C_{in} to be equal to their ideal expression (Equations 5.11, $R_{in} = r_{DS}/N$, and 5.12, $C_{in} = NC_{gs} + C_{pad}$) and a plausible NF, such as 2-3 dBs. Once the optimum has been chosen, V_G can be set accordingly. Moreover, circuit operation was improved using the following rules, with reference to Figure 5.9:

- High L in M_1 and M_2 for good mirroring
- High W in M_1 and M_2 for low V_{GS} to make room for M_3 and M_4
- $\frac{\left(\frac{W}{L}\right)_3}{\left(\frac{W}{L}\right)_4} = 2$ to minimize the static current of the PTAT block
- $\frac{W}{L}$ of M_3 and M_4 for correct V_{GS} in order to set the target g_{DS} of the diodes
- Non minimal $W \times L$ for M_3 and M_4 to reduce mismatch
- Either low $\frac{W}{L}$ for M_5 and M_6 or more stacked MOSFETs, since their V_{GS} determines the input DC voltage of the subsequent stage
- Non minimal $W \times L$ for M_5 and M_6 to reduce mismatch
- $(W/L) = (W/L)_4$ for all diodes M_{D1}-M_{D4} to reduce mismatch and for simpler calculations
- Highest possible R_{bias} allowed by the technology and considering area occupation; the resistor may also be implemented as a MOSFET
- Lowest possible C_{gnd} for effectively cutting the carrier RF frequency
- Careful assessment of the values of C_{drain} and C_{gate} due to the following trade-off: higher capacitances offer better coupling but cause a higher ED C_{in}, which results in a lower matching gain, as shown in Equation 5.25 [14]; the optimum corresponds to the higher overall SNR, using Equation 5.18

It is important to consider the actual values for R_{in} and C_{in}, which can be found by performing an AC simulation, instead of the ideal values $R_{in} = r_{DS}/N$ and $C_{in} = NC_{GS} + C_{pad}$ when implementing the above rules. As a matter of fact, the real and the ideal values may even differ substantially.

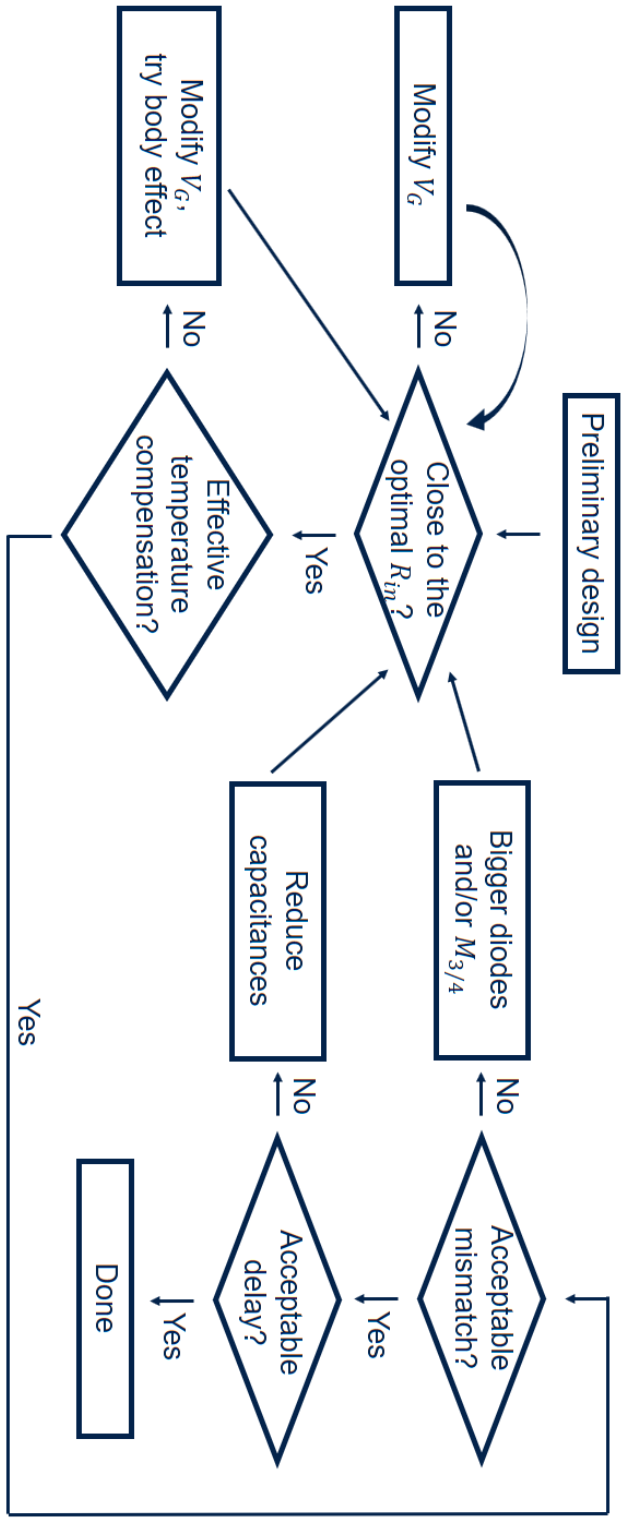

Figure 5.12: Parameter optimization process of the ED implemented in the third prototype, along with the PTAT block.

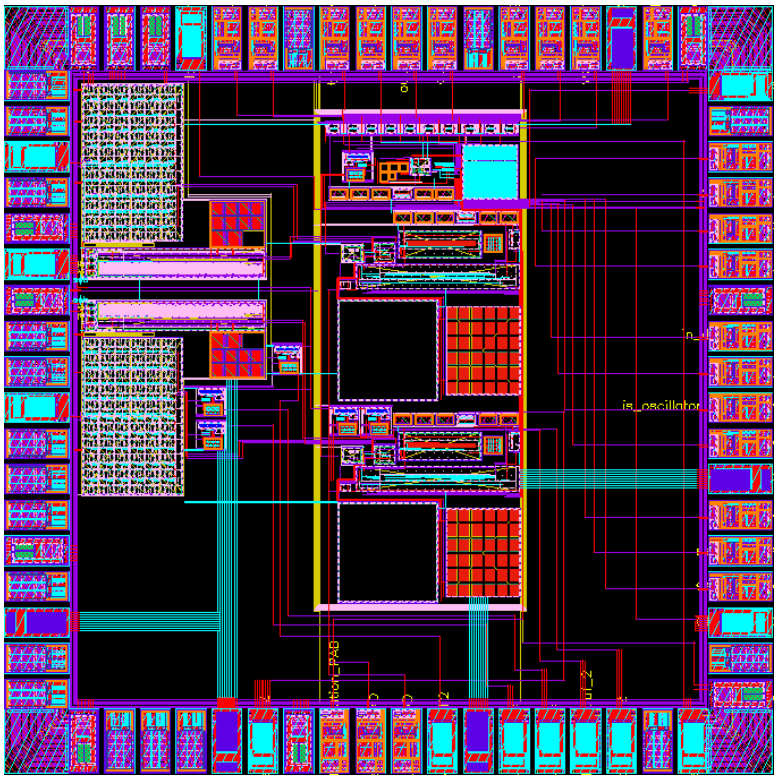

Figure 5.13: Layout of the third WUR prototype presented in this thesis, taped out in 2021.

5.4.3 Parameter optimization

Once a preliminary design is ready, it is possible to perform an optimization of all parameters involved. The aim of this process, shown in Figure 5.12, is to obtain a design with the best possible compromise between optimum sensitivity, speed and temperature compensation. Adding body effect to either M_{D1}-M_{D4} or M_3-M_4 (or both groups of MOSFETs) in Figure 5.9 may be empirically tried to improve sensitivity optimization and temperature compensation.

5.5 ED implementation

The implemented circuit is shown in Figure 5.9, composed of a differential passive ED and a PTAT block for temperature compensation. Simulations were carried out in the -40 : 120$^{\circ}C$ temperature range. Each chain of diode-connected MOSFETs is composed of $N = 60$ stages. The values for the capacitors are $C_{gate} = C_{drain} = C_{gnd} = 36\ fF$. The dimensions of the diodes are not the minimal ones for the technology in order to avoid excessive mismatch, $W = 450\ nm \times 2$ and $L = 450\ nm$. The PTAT block, start-up included, consumes 6.2 nW at 27$^{\circ}C$, 1.4 nW at -40$^{\circ}C$ and 12.2 nW at 120$^{\circ}C$. The overall area of the ED-PTAT section is roughly 0.3 mm^2, 60% of which is due to the capacitors. The chip layout is shown in Figure 5.13. The overall circuit area is roughly 1 mm^2.

The matching network, which will be implemented with a lumped off-chip capacitor and inductor as shown in Figure 5.5, is yet to be designed. As a matter of fact, only indicative values for the capacitance and the inductance can be estimated on the basis of the simulated input impedance of the chip, especially the capacitive part as explained in Section 4.8.2. The chosen components for the matching network with tentative values are a 24-nH Coilcraft inductance and a 30-pF KEMET capacitance. Simulations with the models for these components yield a 14.7-dB simulated matching gain at 433 MHz. The actual values will be chosen according to the measured input impedance of the ED.

5.6 Simulated results

Figure 5.14 shows the transient response of the ED with an input power corresponding to the estimated sensitivity. This estimation has been performed as follows. In [14], it is stated that a 12.3 dB SNR is needed at the input of the comparator in order to have a 10^{-3} MDR, which translates into a signal amplitude that is 4.1 times bigger than the RMS integrated noise at the same point, i.e. $SNR_{req} = 4.1$. Performing an AC simulation allows to calculate the RMS noise voltage from the integrated noise and the estimated sensitivity can be found as the input RF voltage yielding an output amplifier voltage corresponding to 4.1 times the calculated RMS noise voltage. Considering a simulated 14.7-dB matching gain, the estimated sensitivity is -63.1 dBm.

A maximum input power also exists: for powers higher than this, the ED output differential voltage becomes distorted, which also causes a distortion in the amplifier and comparator outputs. Figure 5.15 shows the transient response of the ED with an input power above the maximum permitted power. Its value has been estimated by noticing the fact that v_{out-} cannot go below a minimum value, beyond which the diodes cannot work anymore. According to simulations, this value is around 25 mV at room temperature. This means the maximum excursion v_{out-} can perform is V_P-25 mV, where V_P is defined in Figure 5.9. Therefore, the maximum ED differential output amplitude is $2 \times (V_P - 25\ mV) = 350\ mV$. The input RF power yielding a 350-mV ED output voltage, thus the maximum power which causes no distortion, is -38 dBm, also considering the simulated matching gain. If the input power is higher than this, it may become hard for the BBL to properly recognize the incoming signal.

Figure 5.16 shows the simulated values for the ED input resistance, R_{in}, as a function of temperature. Without the PTAT block, R_{in} ranges from 150 kΩ to 3 kΩ (yellow trace), whereas the use of the PTAT block allows to drastically reduce this variation (red trace). Actually, in this case R_{in} ranges from 35 kΩ to 10 kΩ . This has a positive impact on the stability of sensitivity over temperature, especially at high temperatures, as reported in Figure 5.17. This Figure plots the simulated values for $2\ v_{RF,ant}$ which correspond to sensitivity, i.e. for which the SNR at the output of the amplifier, v_{out_amp}, is again equal to $SNR_{req} = 4.1$, when the PTAT block is used (in

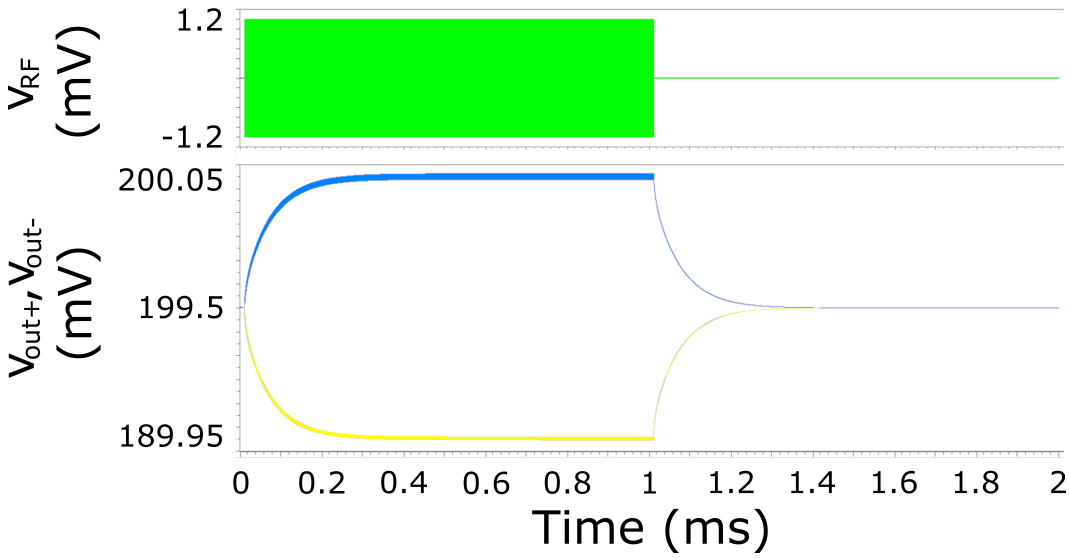

Figure 5.14: Transient simulation showing a 1-0 RF input voltage at the estimated sensitivity (top, green trace) and the voltages at nodes v_{out+} and v_{out-} (bottom, blue and yellow traces), as defined in Figure 5.9. The difference between the latter two constitutes the differential output voltage of the ED.

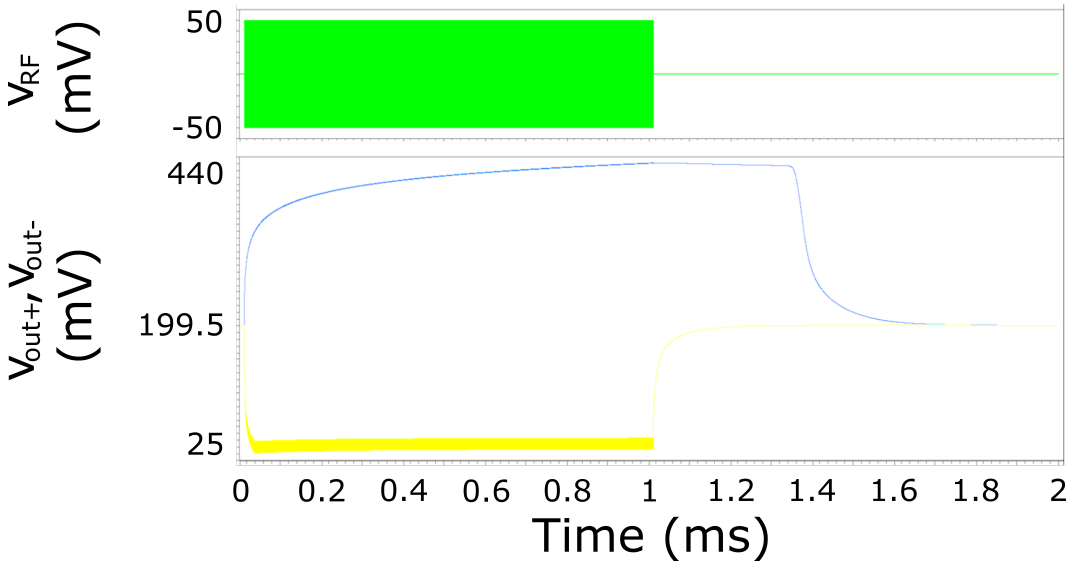

Figure 5.15: Transient simulation showing a 1-0 RF input voltage at a power higher than the maximum input power (top, green trace) and the voltages at nodes v_{out+} and v_{out-} (bottom, blue and yellow traces), as defined in Figure 5.9. A clear distortion occurs.

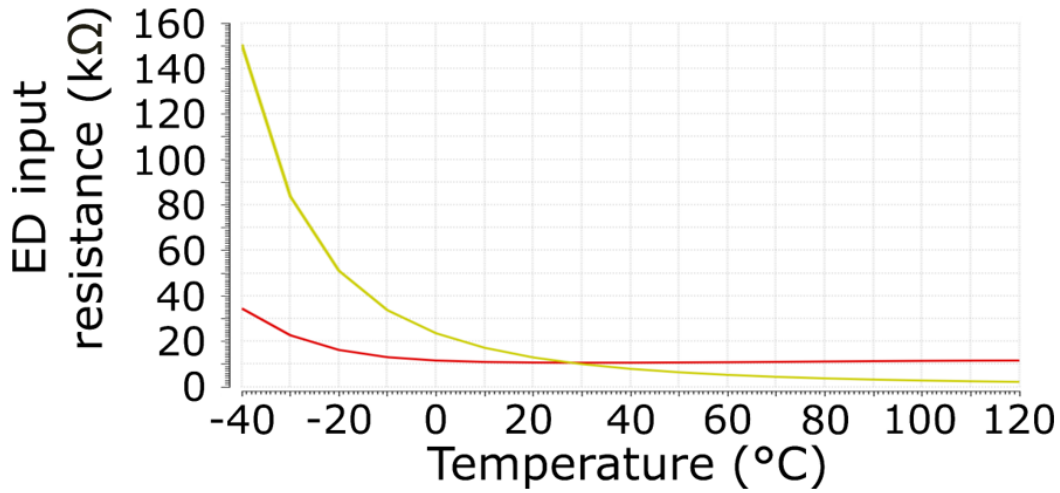

Figure 5.16: Simulated values for the input resistance of the ED, R_{in}, as a function of temperature when the PTAT block is used (in red) and not used (in yellow).

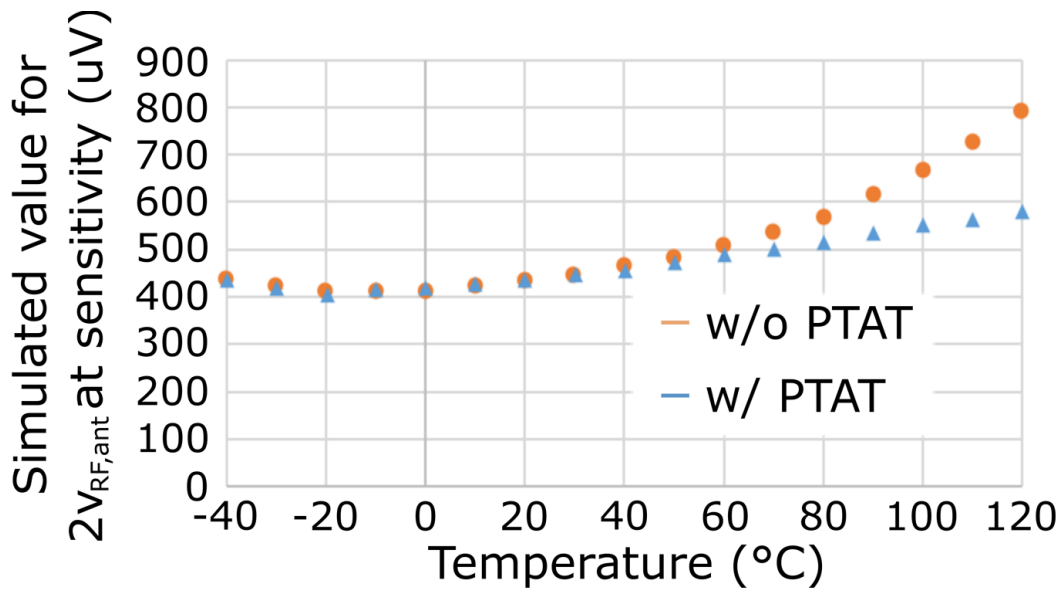

Figure 5.17: Simulated values for 2 $v_{RF,ant}$ corresponding to sensitivity when the PTAT block is used (in blue) and not used (in orange).

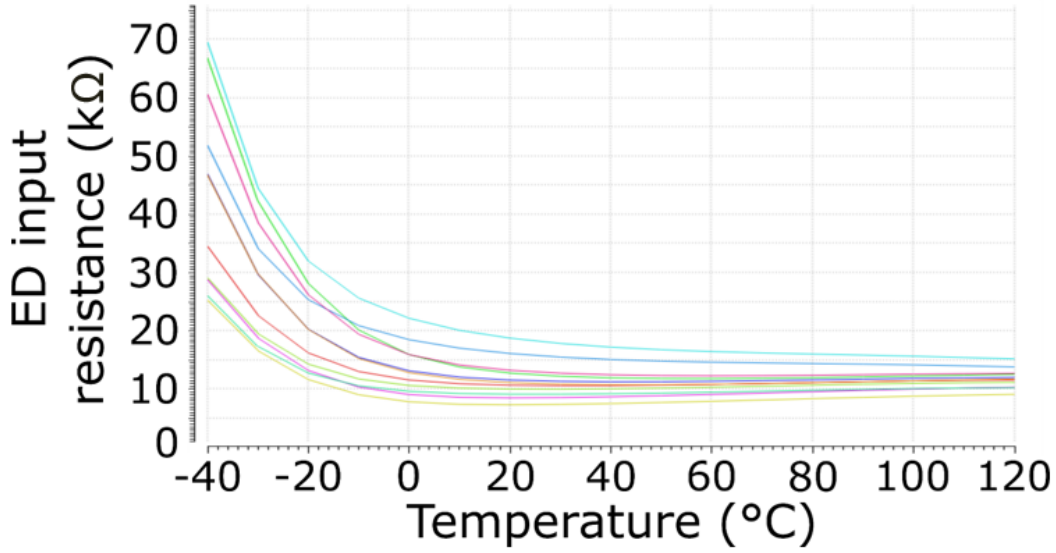

Figure 5.18: 10-run Montecarlo simulation showing the input resistance R_{in} of the ED in Figure 5.9.

blue) and when the PTAT block is not used (in orange). It is necessary to consider the voltage before R_S to account for the loss of matching, occurring especially when the PTAT block is not used. In both cases, the amplifier and the comparator are fed with noiseless currents for a fair comparison. As mentioned before, the input matching network is simulated using models for a commercial inductor and capacitor having the tentative values of $L = 24\ nH$ and $C = 30\ pF$. When the PTAT block is not used, V_G and V_P are set to the values corresponding to those set by the PTAT block at room temperature, in order for the subsequent amplifier to operate correctly. Moreover, the PTAT block performs the biasing of V_G and V_P, which would have to be set otherwise in its absence. However, the impact of its use is not as strong as expected probably due to a relatively small change in the overall resistance seen at the input of the matching network even without the use of the PTAT block: simulations show that with the PTAT block this resistance ranges from 55 to 59 Ω , whereas without the PTAT block it ranges from 43 to 61 Ω . Finally, Figure 5.18 shows a 10-run Montecarlo simulation of the input resistance R_{in} of the ED. It is apparent that this aspect needs to be carefully considered when designing for high yield.

5.7 Remaining AFE blocks and BBL

5.7.1 AFE: amplifier, comparator and offset cancellation

The block diagram of the prototype is shown in Figure 5.19. Actually, the rest of the AFE was designed by another Ph.D. student. Once the input OOK-modulated signal has been rectified by the ED, it is fed to a separated baseband amplifier. Since the output of the ED is differential for the reasons explained in Section 5.3.2, the input of the amplifier is differential as well. Instead, the output of the amplifier is

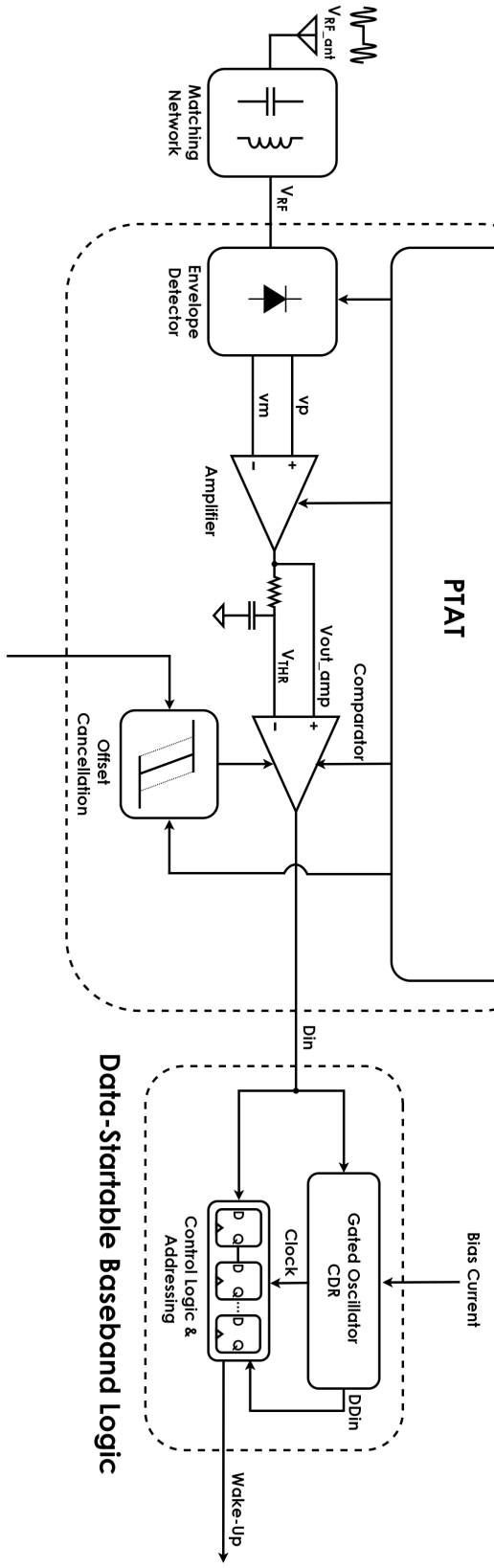

Figure 5.19: Block diagram of the third WUR prototype presented in this thesis, taped out in 2021.

Table 5.1: Performance comparison between the proposed prototypes and several prototypes in literature, part 1.

	Proposed prototype 1	Proposed prototype 2	Proposed prototype 3 (simulated)	[9], 2012	[10], 2013	[11], 2016	[6], 2016	[12], 2017
Temperature range [°C]	-20 : 60	unknown	-40 : 120	unknown	unknown	unknown	unknown	unknown
Carrier frequency [MHz]	771	433	433	915	915	2400	868	2400
Matching gain [dB]	12	17.3*	14.7	12	5	5	unknown	12.5
Bitrate [$kbit/s$]	2	1	0.5	100	12.5	8.192	1	200
Sensitivity at 27°C [dBm]	-46[a]	-52.3*,[b]	-63.1[b]	-41[a]	-43[b]	-56.5[b]	-55[b]	-50[a]
Supply voltage [V]	1.2	0.6	0.6	1.2	1.2/0.5	1/0.5	1.2	0.8
Power at 27°C [nW]	36	12.9**	16.8	98	116	236	1200	4500
Correlator	no	yes	yes	no	yes	yes	yes	no
Code length	N.A.	63	256	N.A.	31	31	16	N.A.
Technological node [nm]	90	90	90	130	130	65 low-power	N.A.	180
Die area [mm^2]	0.7	0.37	1	0.03	0.35	2.25	N.A.	0.456
External components	LC	LC*	LC	LC	LC	LC	N.A.	LC
FoM_1 [dB]	106.9	116.2	124.3	106.1	102.8	112.3	99.2	100
FoM_2 [dB]	100.1	105.6	114.3	101.4	unknown	unknown	100	102.8

* LC matching network simulated: sensitivity projected as sum of measured AFE sensitivity and simulated matching network gain
** Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time
[a] Sensitivity defined through 10^{-3} Bit Error Rate
[b] Sensitivity defined through 10^{-3} Missed Detection Rate

Table 5.2: Performance comparison between the proposed prototypes and several prototypes in literature, part 2.

	[13], 2017	[8], 2018	[14], 2019	[15], 2019	[16], 2019	[17], 2020	[18], 2020	[19], 2020	[20], 2021
Temperature range [$°C$]	unknown	unknown	unknown	unknown	unknown	unknown	-10 : 40	unknown	-40 : 85
Carrier frequency [MHz]	403	113.5	434	750	152	450.8	9000	430	915
Matching gain [dB]	18.5	25	26	13	27	26	13.5	unknown	unknown
Bitrate [$kbit/s$]	0.3	0.3	0.1	200	0.2	0.1	0.033	42	16.4
Sensitivity at 27$°C$ [dBm]	-63.8[b]	-69[b]	-79.2[b]	-50[b]	-76[b]	-80.9[b]	-69.5[b]	-108[b]	-70.2[c]
Supply voltage [V]	0.4	0.4	0.4	0.4	1/0.6	0.54	0.4	unknown	unknown
Power at 27$°C$ [nW]	4.5	4.5	0.42	1486[**]	7.4	40	7.3	41000	2700
Correlator	yes	yes	yes	yes	yes	yes, analog	yes	yes	yes
Code length	32	16	11	40	8	11	18	31	unknown
Technological node [nm]	180	180	65 low-power	65	130	65 low-power	65/180[***]	65 low-power	65
Die area [mm^2]	4.5 (estimated)	4.5 (estimated)	0.05 (estimated)	0.75	1.95	0.15	0.14/ 4.5 cm^2[***]	3	0.33
External components	transformer	transformer	LC	L	LC	LC	transformer	LC	LC
FoM$_1$ [dB]	129.7	134.9	153	104.8	138.8	134.9	128.5	145	117
FoM$_2$ [dB]	115.9	121.1	134.1	103.7	unknown	126.6	115.8	unknown	unknown

[**] Clockless solution: average power computed assuming the system in Phase 2 for 1% of the time
[***] Antenna included
[b] Sensitivity defined through 10^{-3} Missed Detection Rate
[c] Sensitivity defined through 10^{-3} Packet Error Rate

single-ended. The baseband amplifier is critical in terms of noise, as it could significantly degrade sensitivity, as shown in Equation 5.24, by adding non negligible noise to a very small signal. In order to prevent this, its biasing current must be carefully chosen so as to keep NF at acceptable levels. For instance, a biasing current around 10 nA is enough to yield an $NF\ (dB)$ around 3 dB. The output of the baseband amplifier is fed to the comparator along with its LP RC filtered version, which acts as a threshold. The main reason for adopting this approach for threshold generation is its simplicity but, unfortunately, it requires a non-negligible amount of area. This is even more so since with such low currents the output impedance of the amplifier is very high, resulting in an even higher R value to prevent R itself from excessively loading the amplifier. Moreover, the number of equal consecutive bits is limited to prevent the threshold from getting too close to the two voltages at the output of the amplifier corresponding to "0" and "1". Actually, the choice of a Manchester coding is preferred. Since the AFE is clockless, latched comparators cannot be implemented, along with the offset compensation techniques they make possible. Therefore, comparator offset cancellation is carried out by an additional block. This allows to indirectly compensate for the offset of the whole chain, in particular that of the amplifier, which may significantly degrade sensitivity. The PTAT block provides bias currents for the amplifier, the comparator and the offset cancellation block.

5.7.2 BBL: GO-CDR, Control Logic and Addressing Unit

The BBL of this prototype is very similar to that of the second prototype, described in Section 4.7.2. The only differences are the absence of a Bias and Calibration Circuit for the GO-CDR and a longer correlator length. In the previous prototype, the BC circuit was present to set the correct bias current for the GO to oscillate at 1 kHz with a 0.5% tolerance. Since this prototype is supposed to operate with the Manchester coding or a similar coding, phase errors due to subsequent bits cannot accumulate. Therefore, there is no need for a clock frequency calibration. On the other hand, this time the goal of receiving long sequences of bits, which may be either data sequences or encrypted addresses, resulted in the implementation of a much longer correlator, namely 256 bit-long.

5.8 AFE simulated results

The simulated results of the whole AFE in this prototype are summarized in Tables 5.1 and 5.2. Again a roughly 10-dB improvement with respect to the previous prototype has been achieved, mainly thanks to an expected 10-dB sensitivity enhancement in the face of a slight power consumption increase. Moreover, AFE operation over a significant temperature range is expected to be demonstrated, which has rarely happened in literature.

5.9 Conclusions on passive Envelope Detectors

Chapter 5 presents an implementation of an AFE for an integrated medium range WUR employing a passive ED. Passive EDs only perform envelope extraction, i.e. OOK demodulation, while baseband amplification of the extracted envelope has to be performed by another block.
Passive EDs are composed of a chain of diode-connected MOSFETs biased at zero static current. The ED implemented in the third prototype presented in this thesis differs from the standard Dickson charge pump by the possibility of biasing diode gates regardless of the input signal [14]. The voltage at diode gates determines their channel resistance, thus the overall input resistance of the ED, and can be chosen so as to optimize sensitivity. In the proposed prototype, a mechanism for sensitivity stabilization with temperature is added through an additional PTAT block. To avoid degrading ED SNR due to the addition of the PTAT block, a differential approach for the ED has been chosen.
Passive EDs have a major advantage: thanks to zero-static current biasing, not only do they have zero power consumption but they also inherently have a better sensitivity. However, they also have a few drawbacks. For instance, EDs with bitrates above 1 kbit/s can hardly be implemented since the number of diode stages N needs to be big and the ED propagation delay is related to the cube of N. Moreover, they have low flexibility, as their bitrate and carrier RF frequency cannot be changed once they have been designed. They also occupy significant silicon area due to the presence of the coupling and filtering capacitances. Finally, they have a high input capacitance, which makes the design of a matching network with a significant gain hard.

Chapter 6

Conclusions

This Ph.D. research activity aims at designing the Analog Front-End (AFE) of a Wake-Up Radio (WUR) to be integrated within a sensor or an actuator node in a Wireless Sensor and Actuator Network (WSAN). A WUR is a minimal receiver which is constantly on and scanning the RF channel in the place of the main radio. Its aim is to reduce the power consumption of sensor and actuator nodes while enabling asynchronous communication, thus reducing latency as well. A WUR is composed of two subsystems: the AFE, whose task is to turn the input OOK-modulated signal into a stream of bits, and the Baseband Logic (BBL), which compares the received bitstream with the address of the specific node and, if the two match, issues a Wake-Up interrupt.
The goal of this Ph.D. research activity is the design of an AFE which receives OOK-modulated input signals with RF carrier frequencies in the Sub-GHz ISM band and has a power consumption in the order of nanoWatts, thus targeting medium range IoT applications, i.e. with a wake-up distance within 100 meters.
Over the course of the Ph.D. activity, three prototypes implementing both active and passive EDs have been designed using a 90-nm STMicroelectronics technology, as shown in Figure 6.1. The main challenge this topic poses is the sensitivity-power trade-off, that is a longer wake-up distance requires a higher power consumption. All proposed implementations are clockless and leverage the second-order nonlinearities of a MOSFET in subthreshold for envelope extraction.

- *The first prototype* is supplied with 1.2 V and has no addressing capabilities as the BBL is not present. It features an active band-pass (BP) ED performing both envelope extraction and amplification of the extracted envelope. The BP response results in the ED output signal peaking at each transition in the input bit, which in turn forces the use of a hysteretic decision circuit. In order to avoid multiple errors in case of long sequences of equal bits, the use of a type of coding with frequent transitions, such as Manchester coding, is required. The ED is then followed by a Schmitt trigger. Since the information lies in the ED output peaks, the trigger is AC-coupled to the ED and its input DC voltage and two threshold voltages are generated in a temperature-compensated manner by a dedicated block. The biasing block is quite complex and consumes a current which is exponential with temperature.

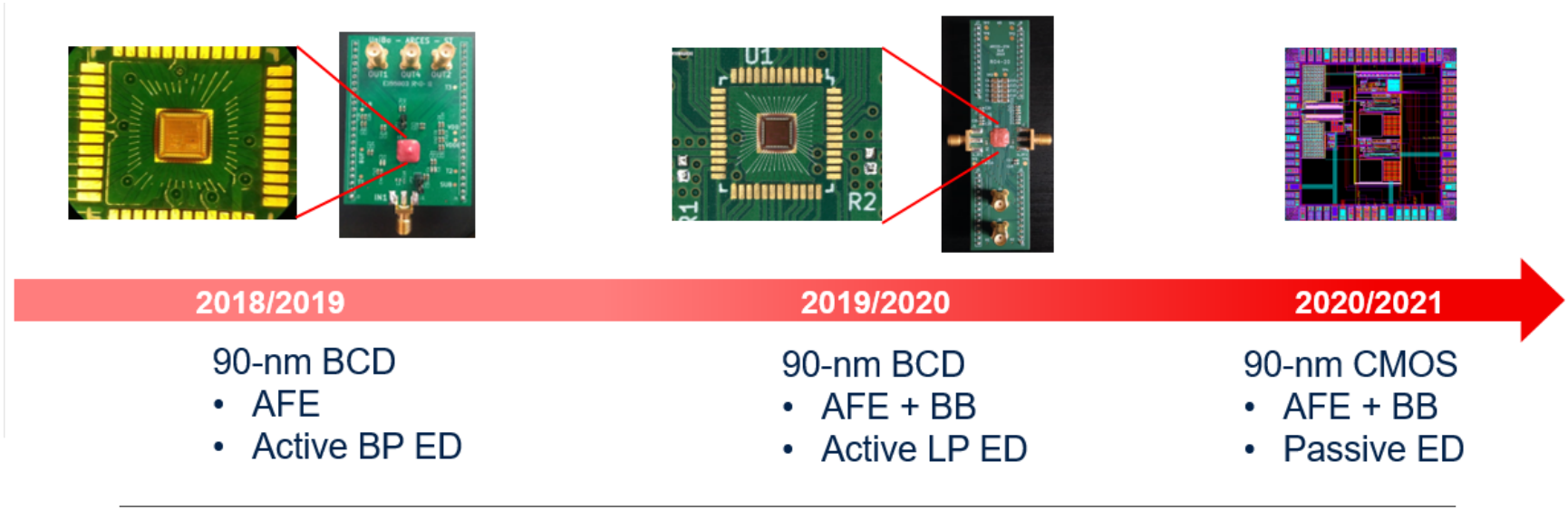

Figure 6.1: The three prototypes designed and implemented over the course of this Ph.D. activity.

Measurements have been carried out on this prototype, yielding a -46-dBm sensitivity at 771 MHz with a 36-nW power consumption over a 1.2-V supply and a 2-kbit/s bitrate. This results in 106.9-dB Figure of Merit (FoM).
This prototype has been presented in [2], [1] and its ED topology has been filed as a patent [22]. [2] also includes a detailed analysis on the active ED topology, which can be implemented as BP or LP.

- *The second prototype* is supplied with 0.6 V and also includes a BBL designed by another Ph.D. student. It solves the issue posed by the need for Manchester coding by using an active low-pass (LP) ED instead of the BP one. This implies the ED outputs a signal with the same shape as the input envelope, resulting in the use of a standard comparator instead of one with a hysteresis. Moreover, again thanks to the LP response of the ED, the reference voltage for the comparator is generated within the ED itself and the comparator is DC-coupled to the ED. As a drawback, the LP response forces the need for a capacitance that is too big to be integrated, i.e. an external capacitance is needed. The bulk voltages of the comparator input MOSFETs are available on the outside to set its equivalent threshold. Finally, biasing currents for the ED and the comparator are provided by a dedicated reference current generator, which is much simpler and consumes lower current than the biasing block in the previous version. Therefore, the use of a LP response instead of a BP allows system simplification at the expense of the presence of an external capacitance to implement the LP response itself. Moreover, there is no need for Manchester coding, which causes bitrate halving. However, all active EDs have an unfavorable current-to-sensitivity relationship. This leads to the use of passive EDs, which inherently have a better sensitivity since they have no flicker noise.
  Measurements have been carried out on this prototype, yielding a projected -52.3-dBm sensitivity at 433 MHz with a 12.9-nW power consumption over a 0.6-V supply and a 1-kbit/s bitrate. This results in 116.2-dB FoM.
  This prototype has been presented in [23] and [24].

- *The third prototype* is supplied with 0.6 V and includes the same BBL as the second one. A passive ED composed of a chain of diodes biased at zero current and based on a recent publication is employed to enhance sensitivity without having to drastically increase ED current consumption. Therefore, the rectification and amplification functions have been separated. Temperature compensation of the ED and the external matching network is performed by an original block similar to a Brokaw cell with the aim of stabilizing the ED input resistance, thus sensitivity, with temperature. The block performing the amplification of the extracted envelope has been designed by another Ph.D. student, along with the subsequent comparator, to which an offset calibration circuit has been added.

Table 6.1: Important parameters to make a comparison between active and passive EDs.

	Active BP	Active LP	Passive
Sensitivity			
Power consumption			
Subsequent stage	Trigger (Manchester needed)	Standard comparator	Amplifier + Comparator
Matching gain (input capacitance)			
Maximum input power			
Die area			
External components			
Flexibility			
Design resources			

This prototype is currently under fabrication. Simulated results yield a -63.1-dBm sensitivity at 433 MHz with a 16.8-nW power consumption over a 0.6-V supply and a 0.5-kbit/s bitrate. This results in 124.3-dB FoM.

In conclusion, a tentative comparison between the active and the passive ED topologies can be performed, as shown in Table 6.1. Both solutions have both advantages and drawbacks and the choice of one over the other mainly depends on the most important constraints in the specific application.
As far as sensitivity is concerned, a passive solution is undoubtedly more suitable than an active one. As a matter of fact, a passive ED has no flicker noise, whereas an active one requires a drastic increase in current consumption for a significant sensitivity improvement. Moreover, if the active solution is LP its main noise contributing MOSFETs may require their area to be enlarged for a lower flicker noise, posing a noise-speed trade-off.
The same can be stated about power consumption, as a passive ED is biased with zero static current, unlike an active one. However, this advantage may not be so significant due to the fact that a separated amplifier has to be added when using a passive ED.
As for system complexity, the use of an active BP ED certainly leads to a more complex system, as a trigger with two thresholds is needed. Moreover, Manchester coding is required. On the other hand, employing an active LP ED allows the use of a standard comparator and no constraints are put on coding. As for passive EDs, the choice of the comparator topology is related more to the amplifier implementation than to the ED.
It is drastically different when considering the realizable matching gain, as a passive ED typically has a much higher input capacitance than an active one. Depending on pad capacitance, the ED input capacitance may become the dominant component in

the capacitance which the matching network sees as a load. If such a capacitance becomes too high, a reduction in the carrier RF frequency may be necessary to realize a significant matching gain.
The use of active EDs is also beneficial if the specific application allows nodes to receive high input power levels. As a matter of a fact, gain compression makes it possible for active EDs to correctly receive input OOK modulated signals with high power levels, whereas with passive EDs the output signal gets distorted if the input power is too high. This may make correctly processing the signal hard or even impossible for the subsequent stages. This problem can only be alleviated by increasing the supply voltage and the number of diode-connected MOSFETs in the lower part of the PTAT block.
Area occupation, both silicon area and area on the board, is another important parameter when choosing what ED architecture to use. As for area occupation on the board, the only solution which imposes the use of external components, namely an external capacitance, is the active LP one. On the other hand, both passive EDs and active EDs occupy a significant silicon area. In both cases, passives are responsible for most of the silicon area occupation, namely coupling and filtering capacitances in the case of the passive ED and the RC filter in the case of the active ED. In both cases, additional area is also due to the resistances in the biasing circuit and in the PTAT block, which need to be in the order of several MΩ for nA currents. Among active EDs, LP ones occupy a smaller silicon area because their capacitance is too big to be integrated and is therefore external. In conclusion, it is complicated to make a comparison between active and passive EDs on such terms and it is probably possible to determine which case is worse only for the specific application.
Flexibility is also an important parameter. When using active EDs it is possible to increase their bitrate by increasing their biasing current, whereas with passive EDs the maximum bitrate cannot be easily changed once the design has been carried out. Actually, a bitrate-sensitivity trade-off is set during the design phase through the number of diode stages. Moreover, the presence of coupling and filtering capacitances in passive EDs makes the design and operation of passive EDs strictly related to the chosen RF frequency, whereas in active EDs the RF frequency is filtered out immediately by the ED, thus making the matching network the only frequency dependent section. As a result, an active ED can be used for different carrier RF frequencies if the matching network is changed accordingly.
The situation is similar for design resources, especially simulation time. Since the design and operation of passive EDs is strictly related to the chosen RF frequency, simulations must be carried out using the chosen RF frequency when designing a passive ED, leading to long and resource consuming simulations. On the other hand, when designing active EDs simulations can be carried out using much lower RF frequencies, as long as non-linearities are solicited.
Ultimately, passive EDs are probably a better choice when high sensitivity is the most important constraint, whereas less stringent constraints are placed on bitrate,

flexilibity, design resources and maximum input power. In all other cases, an active ED is probably to be preferred.

Appendix A

Definition of a new Figure of Merit for the Analog Front End of a medium range WUR

Most of the material reported in this chapter is reused from [2] (©2020 IEEE), in agreement with IEEE copyright policy on theses and dissertations.

As mentioned in Section 3.2, the most widely used FoM to assess the performances of the AFE of a medium range WUR is

$$FoM_1(dB) = -P_{SEN}(dBm) + 5\log BW - 10\log\frac{P_{dc}}{1mW} \tag{A.1}$$

This FoM does not allow entirely fair comparisons between different architectures because it is not an invariant for a given architecture. This can be proven by considering the example in Section 4.2.

In order to make fairer comparisons, a new FoM was defined based on some assumptions:

- Rectification leverages an exponential characteristic, in particular second-order non-linearities, as shown in Equation 4.11.
- $g_m \propto I_{bias}$ and $R_0 \propto 1/I_{bias}$, therefore $g_m R_0 \cong const$ with respect to I_{bias}.
- In-band noise is mostly thermal, i.e. $i_n^2(f) \propto g_m$, therefore the RMS noise voltage at the ED output

$$v_{RMS,out} = \sqrt{g_m {R_0}^2 BW} = \alpha\sqrt{BW/I_{bias}}, \tag{A.2}$$

  where α is constant.

- The required signal-to-noise ratio SNR_{SEN} at the ED output determines sensitivity

$$V_{OUTM\_SEN} = SNR_{SEN}\, v_{RMS,out}. \tag{A.3}$$

Therefore, if v_{M_SEN} is the input RF amplitude which yields V_{OUTM_SEN} at the ED output, we obtain using Equations 4.11 and A.3

$$v_{M\_SEN}^2 \, k \, \frac{g_m R_0}{4nV_T} = \alpha \; SNR_{SEN} \sqrt{\frac{BW}{I_{bias}}} \tag{A.4}$$

where k is a proportionality constant. Equation A.4 shows that

$$v_{M\_SEN}^2 \sqrt{\frac{I_{bias}}{BW}} = \alpha \; SNR_{SEN} \; \frac{1}{k} \; \frac{4nV_T}{g_m R_0}, \tag{A.5}$$

where the right-hand side is constant. Therefore, the left-hand side is an invariant for any ED architecture sharing the above assumptions, making it a suitable figure to compare the performances of different architectures. Obviously, it is desirable for this quantity to be as small as possible, so the FoM can be defined as its inverse expressed in dBs, $10 \log \left[v_{M\_SEN}^2 \sqrt{I_{bias}/BW} \right]$, yielding

$$FoM_2(dB) = -P_{SEN}(dBm) + 5 \log BW - 5 \log I_{bias} \tag{A.6}$$

Bibliography

[1] A. Elgani, M. Magno, F. Renzini, *et al.*, "Nanowatt wake-up radios: Discrete-components and integrated architectures," in *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, IEEE, 2018, pp. 793–796.

[2] A. M. Elgani, F. Renzini, L. Perilli, *et al.*, "A clockless temperature-compensated nanowatt analog front-end for wake-up radios based on a band-pass envelope detector," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 8, pp. 2612–2624, 2020. DOI: 10.1109/TCSI.2020.2987850.

[3] D. Spenza, M. Magno, S. Basagni, L. Benini, M. Paoli, and C. Petrioli, "Beyond duty cycling: Wake-up radio with selective awakenings for long-lived wireless sensing systems," in *2015 IEEE International Conference on Computer Communications (INFOCOM)*, 2015, pp. 522–530.

[4] F. A. Aoudia, M. Magno, M. Gautier, O. Berder, and L. Benini, "A low latency and energy efficient communication architecture for heterogeneous long-short range communication," in *2016 Euromicro Conference on Digital System Design (DSD)*, 2016, pp. 200–206. DOI: 10.1109/DSD.2016.97.

[5] T. Le, M. Magno, A. Pegatoquet, O. Berder, O. Sentieys, and E. Popovici, "Ultra low power asynchronous mac protocol using wake-up radio for energy neutral wireless sensor networks," *ENSSys 2013 - Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems*, Nov. 2013. DOI: 10.1145/2534208.2534221.

[6] M. Magno, V. Jelicic, B. Srbinovski, V. Bilas, E. Popovici, and L. Benini, "Design, implementation, and performance evaluation of a flexible low-latency nanowatt wake-up radio receiver," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 633–644, 2016.

[7] P. Mercier, "Design of low-power wake-up radios for iot applications," in *2018 IEEE Custom Integrated Circuits Conference (CICC)*, 2018, pp. 1–71. DOI: 10.1109/CICC.2018.8357117.

[8] P.-H. P. Wang, H. Jiang, L. Gao, *et al.*, "A near-zero-power wake-up receiver achieving 69-dbm sensitivity," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 6, pp. 1640–1652, 2018. DOI: 10.1109/JSSC.2018.2815658.

[9] N. E. Roberts and D. D. Wentzloff, "A 98nw wake-up radio for wireless body area networks," in *2012 IEEE Radio Frequency Integrated Circuits Symposium*, 2012, pp. 373–376. DOI: 10.1109/RFIC.2012.6242302.

[10] S. Oh, N. E. Roberts, and D. D. Wentzloff, "A 116nw multi-band wake-up receiver with 31-bit correlator and interference rejection," in *Proceedings of the IEEE 2013 Custom Integrated Circuits Conference*, 2013, pp. 1–4. DOI: 10.1109/CICC.2013.6658500.

[11] N. E. Roberts, K. Craig, A. Shrivastava, *et al.*, "26.8 a 236nw 56.5dbm-sensitivity bluetooth low-energy wakeup receiver with energy harvesting in 65nm cmos," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 450–451. DOI: 10.1109/ISSCC.2016.7418101.

[12] K.-W. Cheng and S.-E. Chen, "An ultralow-power wake-up receiver based on direct active rf detection," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 7, pp. 1661–1672, 2017. DOI: 10.1109/TCSI.2017.2664919.

[13] P.-H. P. Wang, H. Jiang, L. Gao, *et al.*, "A 400 mhz 4.5 nw 63.8 dbm sensitivity wake-up receiver employing an active pseudo-balun envelope detector," in *ESSCIRC 2017 - 43rd IEEE European Solid State Circuits Conference*, 2017, pp. 35–38. DOI: 10.1109/ESSCIRC.2017.8094519.

[14] V. Mangal and P. R. Kinget, "Sub-nw wake-up receivers with gate-biased self-mixers and time-encoded signal processing," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 12, pp. 3513–3524, 2019. DOI: 10.1109/JSSC.2019.2941010.

[15] M. Elhebeary, L.-Y. Chen, S. Pamarti, and C.-K. KenYang, "An 8.5pj/bit ultra-low power wake-up receiver using schottky diodes for iot applications," in *ESSCIRC 2019 - IEEE 45th European Solid State Circuits Conference (ESSCIRC)*, 2019, pp. 205–208. DOI: 10.1109/ESSCIRC.2019.8902825.

[16] J. Moody, P. Bassirian, A. Roy, *et al.*, "Interference robust detector-first near-zero power wake-up receiver," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 8, pp. 2149–2162, 2019. DOI: 10.1109/JSSC.2019.2912710.

[17] V. Mangal and P. R. Kinget, "Clockless, continuous-time analog correlator using time-encoded signal processing demonstrating asynchronous cdma for wake-up receivers," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 8, pp. 2069–2081, 2020. DOI: 10.1109/JSSC.2020.2980526.

[18] H. Jiang, P.-H. P. Wang, L. Gao, *et al.*, "A 22.3-nw, 4.55 cm2 temperature-robust wake-up receiver achieving a sensitivity of 69.5 dbm at 9 ghz," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 6, pp. 1530–1541, 2020. DOI: 10.1109/JSSC.2019.2948812.

[19] A. Dissanayake, J. Moody, H. L. Bishop, *et al.*, "A -108dbm sensitivity, -28db sir, 130nw to 41µw, digitally reconfigurable bit-level duty-cycled wakeup and data receiver," in *2020 IEEE Custom Integrated Circuits Conference (CICC)*, 2020, pp. 1–4. DOI: 10.1109/CICC48029.2020.9075907.

[20] K.-K. Huang, J. K. Brown, N. Collins, *et al.*, "21.3 a fully integrated 2.7µw -70.2dbm-sensitivity wake-up receiver with charge-domain analog front-end, -16.5db-sir, fec and cryptographic checksum," in *2021 IEEE International Solid-State Circuits Conference (ISSCC)*, vol. 64, 2021, pp. 306–308. DOI: 10.1109/ISSCC42613.2021.9365806.

[21] B. Razavi and B. Wooley, "Design techniques for high-speed, high-resolution comparators," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 12, pp. 1916–1926, 1992. DOI: 10.1109/4.173122.

[22] A. M. Elgani, F. Renzini, L. Perilli, *et al.*, "Envelope edge detector in weak inversion," 2018.

[23] M. D'Addato, A. M. Elgani, L. Perilli, *et al.*, "A gated oscillator clock and data recovery circuit for nanowatt wake-up and data receivers," *Electronics*, vol. 10, no. 7, 2021. DOI: 10.3390/electronics10070780.

[24] M. D'Addato, A. Antolini, F. Renzini, *et al.*, "Nanowatt clock and data recovery for ultra-low power wake-up based receivers," in *Proceedings of the 2020 International Conference on Embedded Wireless Systems and Networks*, ser. EWSN '20, Lyon, France: Junction Publishing, 2020, 224–229.

[25] STMicroelectronics, *BCD technology*. [Online]. Available: https://www.st.com/content/st_com/en/about/innovation---technology/BCD.html.

[26] M. Rose and H. J. Bergveld, "Integration trends in monolithic power ics: Application and technology challenges," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 9, pp. 1965–1974, 2016.

[27] A. Andreini, C. Contiero, and P. Galbiati, "A new integrated silicon gate technology combining bipolar linear, cmos logic, and dmos power parts," *IEEE Transactions on electron devices*, vol. 33, no. 12, pp. 2025–2030, 1986.

[28] F. Renzini, "Design techniques to enhance low-power wireless communication soc with reconfigurability and wake up radio," *Doctoral Thesis*, 2019.

[29] L. W. Couch, *Fondamenti di Telecomunicazioni: Sistemi di Comunicazione Analogici e Digitali [Digital and Analog Communication Systems]*. Pearson, 2008.

[30] A. Brokaw, "A simple three-terminal ic bandgap reference," in *1974 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, vol. XVII, 1974, pp. 188–189. DOI: 10.1109/ISSCC.1974.1155346.

[31] F. Maloberti, *Analog Design for CMOS VLSI Systems*. Kluwer Academic Publishers, 2001.