Wireless wire - ultra-low-power and high-data-rate wireless communication systems

Wireless Wire — Ultra-Low-Power and High-Data-Rate Wireless Communication Systems

Xia Li

Front and back covers:

“The persistence of injection locking”, oil painting by Xia Li

The invitation bookmark:

“Burn the witch”, oil painting by Xia Li

Wireless Wire — Ultra-Low-Power and High-Data-Rate Wireless Communication Systems

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op donderdag 29 september 2011 om 14.00 uur

door

Xia Li

geboren te Xi'an, China

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. P.G.M. Baltus
en
prof.dr.ir. A.H.M. van Roermund

Copromotor:
dr. D. Milosevic

A catalogue record is available from the Eindhoven University of Technology Library.

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Xia Li

Wireless Wire — Ultra-Low-Power and High-Data-Rate Wireless Communication Systems / by Xia Li. – Eindhoven : Technische Universiteit Eindhoven, 2011.
Proefschrift. – ISBN: 978-90-386-2548-5
NUR 959
Key words: Low power / high data rate / CMOS integrated circuit design / millimeter wave / injection locking / wake-up receiver / wireless phased array receiver frontends

Lisa · Maverick

Everybody has a witch in one's soul
She entices you to be a maverick
and keeps your secret wounds being healed
You love yourself as she has told
You become beloved by the rest once she would go

I used to love my witch
But now I have no choice but to burn her into ashes

Life is a live show
No rehearsal and the audience is low
Staying with her, I am a weirdo but full with hope
Without her, I become somebody but smile like a joke
My dream is sealed
My witch is killed
I want to grow
I want to glow

-By Xia Li

Dedicated to my beloved ones

Samenstelling Promotiecommissie:

prof.dr.ir. P. van den Bosch	TU Eindhoven (voorzitter)
prof.dr.ir. P.G.M. Baltus	TU Eindhoven (promotor)
prof.dr.ir. A.H.M. van Roermund	TU Eindhoven (2nd promotor)
dr.ir. D. Milosevic	TU Eindhoven (co-promotor)
prof.dr.ir. J.R. Long	TU Delft
prof.dr.ir. B. Nauta	TU Twente
prof.dr.ir. J.P. Linnartz	TU Eindhoven / Philips Research
dr.ir. G. Dolmans	Holst Center /IMEC-NL

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACK	Acknowledgement
ADC	Analog-to-digital convertors
AF	Array factor
AoA	Angle-of-arrival
BAN	Body-area network
BER	Bit-error-rate
BPF	Band-pass filter
BW	Bandwidth
CA	Collision avoidance
CIR	Carrier-to-interference ratio
CSMA/CA	Carrier sense multiple access with collision avoidance
DAC	Digital-to-analog convertor
DCF	Duty cycle factor
DO	Directional to omni-directional
DRL	Dual resonated load
EIRP	Equivalent isotropic radiation power
EVM	Error vector magnitude
FCC	Federal Communications Commission
FOM	Figure-of-merit
GaAs	Gallium Arsenide
IC	Integrated circuit
IIP3	Input-referred third-order intercept point
IJLO	Injection-locked oscillator
IMEC	Inter-Universitair Micro-Electronica Centrum
InP	Indium Phospide
IR-UWB	Impulse-radio ultra-wideband
ISI	Inter-symbol interference
ISM	Industrial, scientific and medical
LNA	Low-noise amplifier
LO	Local oscillator
LoS	Line-of-sight
LR-WPAN	Low-rate wireless personal-area network
MAC	Media access control
MB-OFDM	Multiband orthogonal frequency-division multiplexing
mmW	Millimeter-wave
NF	Noise figure

OOK	On-off-keying
P2P	Point-to-point
PA	Power amplifier
PAE	Power-added efficiency
PEIF	Power efficiency improvement factor
PL	Pathloss
PLL	Phase-locked-loop
PSK	Phase-shift-keying
QoE	Quality-of-experience
QoS	Quality of service
RFID	Radio-frequency identification
RTPS	Reflection type phase shifter
Rx	Receiver
SiGe	Silicon Germanium
SNR	Signal-to-noise ratio
SR	Short-range
SW-CPW	Slow-wave coplanar waveguide
TL	Transmission lines
Tx	Transmitter
VCO	Voltage-controlled oscillator
VGA	Viable-gain amplifier
Wi-Fi	Wireless fidelity
WiGig	Wireless gigabit
WLAN	Wireless local-area network
WuRx	Wake-up receiver

LIST OF SYMBOLS

B	Bandwidth
B_{inst}	Instantaneous bandwidth
B_{eff}	Effective bandwidth
C	Channel capacity
D_{ant}	Antenna directivity
$E_{ave,corr}$	Average energy per bit of correctly received bits
E_b	Energy transmission per bit
E_{bit}	Energy consumption per bit
F	Noise factor
f	Frequency
G_{ant}	Antenna gain
g_m	Transconductance
G_{Tx}	Antenna gain of the Tx
G_{Rx}	Antenna gain of the Rx
G_{TRx}	Total antenna gain
I_{inj}	Injection current
I_{osc}	Oscillation current
L_{pkt}	Packet length
k_e	Electronics factor
k_{Tech}	Technology factor
n_{event}	Event frequency
$P_{b,Tx,DC}$	Power consumption of the wake-up beacon generation
$P_{DC,tot}$	Total power consumption of the entire front end
PL	Path loss
$P_{LO,DC}$	Power consumption of an LO
P_{out}	Output power of a Tx
$P_{pr,DC}$	Pre-receiving power
$P_{Rx,DC}$	Power consumption of a Rx
$P_{Rx,DC,n\text{-}path}$	Power consumption of an *n*-element phased array Rx front end
$P_{Tx,DC}$	Power consumption of a Tx
$P_{Tx,DC,n\text{-}path}$	Power consumption of an *n*-element phased array Tx front end
$P_{TRx,DC,n\text{-}path}$	Total Power consumption of an *n*-element phased array front end
$P_{WuRx,DC}$	Power consumption of a wake-up receiver
Q	Quality factor
R	Data rate
R_{effect}	Effective data rate

R_{raw}	Raw data rate
S	Receiver sensitivity
S_{WuRx}	Wake-up receiver sensitivity
SNR_{in}	Input signal to noise ratio
SNR_{out}	Output signal to noise ratio
T_{DC}	Active time of one cycle of a wake-up receiver
$t_{Rx,overhead}$	Packet overhead
$t_{settling}$	Total settling time
T_{tot}	Total observe time
T_{wu}	Active time of one cycle of a duty-cycled wake-up receiver
α	Conduction angle of a PA
θ_m	Beam-pointing angle
θ_{ss}	Steady-state phase difference
κ	Power-linearity factor
λ	Wave length
η_{PA}	Efficiency of a PA
η_{Tx}	Efficiency of a transmitter
ω_{inj}	Injection signal frequency
ω_0	Free oscillation frequency
ω_L	Locking range of an injection locked oscillator
ω_{res}	Frequency sweep resolution (sweep step)

CHAPTER 1 INTRODUCTION

1.1 Background

Thanks to Moore's law, capabilities of integrated circuits (IC) have been improving exponentially in the past decades. The size of transistors has consistently shrunk, and the speed, which is normally described by the cut-off frequency f_T, is doubled about every two years [1]. Along with technology progress and rapid development of wireless technologies, the wireless communication market has been growing explosively and various consumer electronics products have been developed accordingly[1], such as Bluetooth earphone and wireless gaming consoles. In today's market of interest, two types of prevalent applications can be distinguished: low-power and high-speed wireless communications. In the first category, most of the systems are relatively low frequency, low rate and low cost. For example, as regulated in the IEEE 802.15.4 low-rate wireless personal-area network (LR-WPAN) standard, Zigbee systems can provide low-cost and low-power wireless connections for equipment that requires battery life as long as several months to several years. Its maximum raw data rate is only 250 kbps at 2.4 GHz [2]. Quite on the contrary, the second category targets at high-speed communications for massive data transferring. This type of communication systems tends to offer wireless connections for short-range (SR) personal data transfer, which is also known as the last-meter problem. The IEEE 802.15.3c task group has worked on a 60-GHz WPAN standard that is able to provide a data rate as high as at least 1 Gbps for short-range video streaming and high-speed downloading [3]. Unsurprisingly, the power dissipation of such systems would easily reach a level of tens to hundreds of milliwatts [4].

There are also some wireless systems that are in between of those two categories, e.g. Bluetooth systems (IEEE 802.15.1), and wireless local-area networks (WLAN, IEEE 802.11 standards family)[2]. However, a trend of data rate increase has been observed in these standards. For instance, the maximum data rate is extended from 11 Mbps of IEEE 802.11.b to 150 Mbps of IEEE 802.11.n [5]. In 2010, wireless gigabit (WiGig) alliance has published a tri-band (2.4 GHz, 5 GHz and 60 GHz) system that commercializes the

[1] Although the performance improvement of wireless radios lags behind the digital basebands that are predicted by Moore's law, rapid progress can be observed in improvement of the entire system.

[2] In fact, it is not very reasonable to compare WPAN and WLAN systems due to their different applications and requirements like range and mobility. However, there is a very interesting phenomenon that along with increasing of data rate and particularly emerging of WiGig, their definitions seem to be blurred and thus can be compared to some extent.

60 GHz band for indoor personal communications and wireless fidelity (Wi-Fi) applications. It is able to provide up to 7 Gbps data rate and is backward compatible to 802.11 standards. Its applications include wireless synchronization, wireless display, cordless computing and internet access. New scheduled access mode to reduce power consumption is also included in WiGig systems [6].

Some IC designers believe that for given applications and physical constraints (e.g. certain communication quality of service (QoS), range, and IC technology level, etc.), high data rate always results in high power dissipation. However, the application-driven market apparently does not agree with that. There is an increasing demand for a solution to achieve high speed and low power of communications simultaneously, so that people can transmit large files with very low power dissipation[3] and in a short time (e.g. several seconds). One of the most important applications is point-to-point (P2P) data links, as shown in Fig. 1.1.

Fig. 1.1 Wireless P2P commutation.

According to Shannon's theory, some efforts have been made by investigating broadband solutions, such as the impulse-radio ultra-wideband (IR-UWB) system. In principle, if the bandwidth is sufficiently wide, it can support high data rate communication with very low RF power density. This field has gained momentum since Federal Communications Commission (FCC) regulations allowing unlicensed communication using the 3.1 to 10.6 GHz frequency band in 2002 [7]. However, due to the inherent low immunity to noise and interference, difficulties to capture multi-path energy, high power dissipation of high-speed analog-to-digital converters (ADC), coexistence problems, and difficulties in

[3] It usually refers to the average power dissipation level which determines battery life or suitability of energy scavenging technologies.

receiver design (particularly, synchronization issues), this type of systems still cannot fulfill the job as a low power and high speed solution at this moment. For instance, multiband orthogonal frequency-division multiplexing (MB-OFDM) UWB systems are announced by WiMedia Alliance, which are able to provide 480 Mbps maximum data rate [8]. However, it still has co-existence problems with wireless systems below 10 GHz. Besides that, the transmit power density is stringently limited below -41.3 dBm/Hz, which limits the signal-to-noise ratio (SNR) level and therefore degrades the design flexibility of such systems.

1.2 Research objectives

The primary goals of this work are to investigate and propose a new type of wireless system solution which is able to support high data rate communication with low power consumption (as illustrated in Fig. 1.2), and to provide circuit-level designs and implementations to verify the effectiveness of the system.

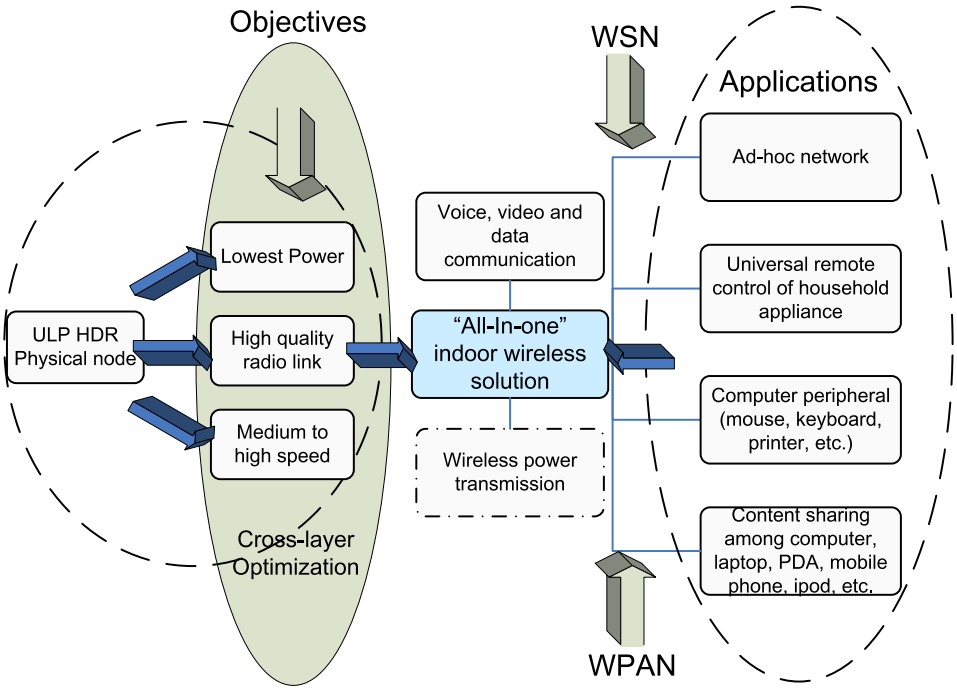


Fig. 1.2 Research objectives and potential applications.

Efforts will be made but not be limited to:

- Investigating trade-offs among PHY-layer parameters and identifying optimum frequency bands and key techniques that fundamentally enable high energy efficiency communication.
- Identifying the suitability of high data rate communication in a rendezvous system

and finding out the optimum data rate level with respect to different application requirements.
- Cross-layer analyses and optimizations towards low average power consumption. Some potential medium access control (MAC) and network methods are discussed in order to achieve overall power optimization.
- Proposing a suitable power management scheme that is compatible with high-data-rate systems based on above results. The overall goal is to minimize average power consumption on system level.
- Design and optimization of system architectures based on link budget model, application requirements as well as the system-level optimization results.
- Circuit designs of key blocks of the proposed system.
- Implementation and characterization of key blocks to verify the new concepts and system design.

1.3 Challenges and research method

This thesis aims at providing generic system solutions for ultra-low power and high-speed personal communications. Besides, key blocks will be designed, implemented and characterized in order to verify new concepts, system architectures and circuit theories. Compared with those design-oriented projects that mainly focus on limited aspects or sub-blocks of a large system, this work is arranged in a top-down manner: i.e. starting from very broad high-level practical issues, going down to realization of a feasible physical system, and then down to the detailed designs and implementations. Innovations are made on aspects where problems are encountered and identified. Results will be cross-checked iteratively in order to maintain the overall targets of design and optimization. Potential research challenges can be classified as flows:

A. Challenges during cross-layer optimization

According to the wooden barrel theory[4], the overall performance of a complicated system is usually determined or limited by its weakest chain. As a result, a cross-layer research scope becomes important. The main necessities include:

- Design of a communication system is actually a practice-oriented work. Before solving a specific physical problem, one should make it clear why this problem must be solved and to what extent it should or can be improved. That kind of information

[4] A barrel composed of many pieces of wood. If the composition of these pieces vary in length, the maximum capacity of the wooden barrel does not depend on the longest piece of wood, but on the shortest one [8].

is normally given by higher levels of a communication system instead of merely PHY-layer parameters.

- Some detailed problems like power consumption reduction or interference alleviation of a communication system can only be optimally solved through cross-layer cooperation. It does not mean that one cannot "solve" a problem by only looking into one specific layer of the communication model, but if doing so, a global optimization will not be guaranteed. Cross-layer trade-offs have to be concerned during the process of solving a practical problem.
- In principle, by knowing relevant figure-of-merits (FOMs) and state-of-the-art benchmarks of certain IC technologies, the boundary conditions (or some values close to these) of a circuit block can be obtained. Trade-offs and transforms can then be made accordingly in order to optimize this block with regard to certain aspects such as power consumption, linearity and noise. However, considering the entire communication system, FOMs or proper benchmarks are hard to obtain due to disparate applications. Approaching physical boundary conditions (e.g. minimum noise figure or power) of each block does not necessarily lead to the overall optimum of the entire system. As a result, a cross-layer view is also required to guide or check the physical systems and circuits design.

When realizing a cross-layer optimization, however, it turns out to be quite a challenging work, because:

- The degree of optimization is determined not only by the complexity of the system, or the degree of difficulty of research questions, but it is limited by the knowledge of designers as well, because a designer is normally specialized in limited fields. As a result, in practice, the cross-layer optimization process has different weights for each layer depending on the designer's own preference. In this thesis, the PHY-layer systems and circuit realization and optimization are our focused aspects. Some higher-layer parameters are not involved in the optimization process but just assumed within a reasonable range.
- It can be seen that there are various trade-off parameters among and across each layer. It is usually difficult and nontrivial to achieve an ultimate optimization. In order to reduce complexity while keeping in-depth analyses over certain essential aspects, the scope of the research work should be limited to a reasonable and controllable range. Therefore, only the most direct cross-layer parameters will be considered during the optimization flow and others will be assigned as assumptions or prerequisites. This activity will reduce the design complexity and number of iterations significantly at the expense of optimization degree. For instance, in our 60-GHz phased array system, line-of-sight (LoS) communication is assumed at the first place, because: (i) in the network layer, a short-range (less than 1-meter relay distance) ad hoc topology and a multi-hop routing method are assumed; (ii) in the chosen MAC power management

scheme, the radio operates in a high-speed burst-mode rendezvous manner, and it will be shut off immediately when communication is done; (iii) with a narrow beamwidth of the array and a directional communication method, diffraction, refraction and reflection phenomena are less serious. As a result, in the link budget calculation, the path loss index is taken as 2, i.e. as the free-space case.

B. Challenges due to lack of methodology

Design and optimization within one specific layer of the communication model is complicated and iterative, too. For instance, as shown in Fig. 1.3, in order to minimize the overall average power consumption of a system, different aspects have to be considered, and innovations may be required in each aspect (details will be specified later in following chapters), but sometimes they may conflict with each other. For example, compared to digital binary amplitude modulation schemes like on-off-keying (OOK), phase-shift-keying (PSK) scheme requires a lower signal-to-noise ratio per bit (E_b/N_0) for a given bit-error-rate (BER). Consequently, with the same data rate and receiver (Rx) noise figure (*NF*), the PSK system requires less output power of the transmitter (Tx) in order to maintain the same communication range. Therefore, the Tx power consumption can be reduced. However, the Tx for a PSK system requires a higher level of linearity. As a result, a higher-efficiency non-linear power amplifier (PA) is not allowed to be used in this case, which may cause Tx power consumption increase in the end.

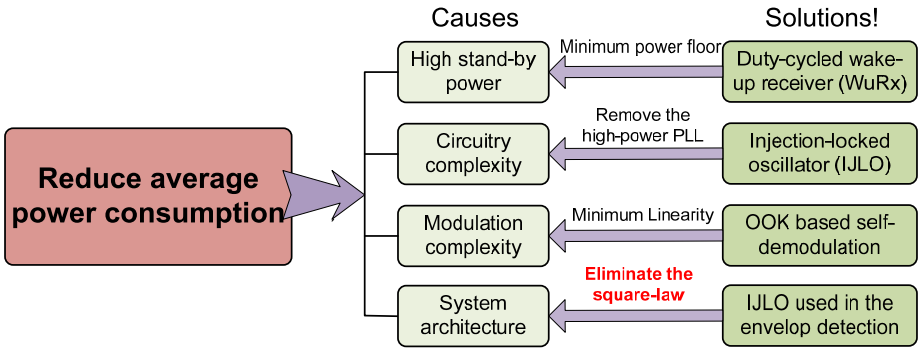


Fig. 1.3 Aspects related to average power consumption reduction of a system.

Situations would get worse when the systems are working at very high frequency, with very high speed or when they must achieve ultra-low power consumption, because in such cases, limits of technology such as the limited cut-off frequency (f_T), high noise and high parasitics levels, will be critical and even decisive. A structuralized design and optimization methodology would therefore be very useful to solve this kind of problem. In [9], a minimum-power design methodology is proposed that provides solutions to optimize gain, linearity and noise of a system, and also offers a detailed front-end design procedure. This method will be used during the system architecture design process in this thesis. More details and relevant parameters will be added with regards to real design

problems. Considering system architecture innovations, a new design method (not that generic as [9], but more specific for our designed architecture) will be proposed as well.

C. Challenges related to some practical issues

Besides fundamental limitation factors like propagation pathloss at a certain frequency, some practical limitations also hamper performance optimization of a communication system. These practical issues include:

- Implementation technology. Firstly, it is known that technologies that use III-V compounds, such as Gallium Arsenide (GaAs) and Indium Phospide (InP), normally offer high f_T, low noise and high speed compared to Silicon based technologies like CMOS. However, due to its high integration level and low cost (per unit), CMOS technology would normally be preferred for mass production and therefore becomes an interesting potential technology for research work as well. Consequently, it is inevitable that the circuits will suffer from poor noise, worse linearity and lower gain (at the same bias level) compared to these III-V technologies [10]. Secondly, trade-offs have to be made when choosing a certain version among different silicon based technologies. For instance, Silicon Germanium (SiGe) technologies would be relatively cheaper than the most advanced CMOS technologies and it is able to offer even better performance. However, it is reported in [11] that the technology development of BiCMOS technology lags about three to four generations behind compared to CMOS technology because of the lack of product drivers in the market. Consequently, it is worth to keep on with the most leading technology, i.e. CMOS for our research at the expense of some performance degradation. Thirdly, in order to achieve high data rate communication, systems normally choose to work at higher frequency bands like the millimeter-wave (mmW) bands due to their wider bandwidth. This would cause potential problem in CMOS technology due to their limited f_T and maximum available gain of transistors. Along with the downscaling process of CMOS technology, the cut-off frequency continues to increase. For example, the maximum oscillation frequency increased from 200 GHz in 65-nm to 250 GHz in 45-nm CMOS technologies [12]. However, the parasitics level and leakage problems, on the other hand, become more severe. All these imperfections of the implementation technology will degrade the performance of a communication system, especially in high-frequency systems.
- Regulation and standardization
  The realization and optimization of a communication system should abide by specific regulations (e.g. regulations of FCC) and be compatible with certain communication protocols as well. For instance, the most achievable system bandwidth and maximum allowed output power of the 60 GHz band is limited as shown in Table 1.1.

Table 1.1 Regional spectrum allocations and EIRP constraints of 60 GHz bands [11]

Regions	Frequency bands (GHz)	Maximum EIRP[5] (dBm)
US, Canada and Korea	57-64	43
Europe and Japan	59-66	57
Australia	59.4-62.9	51.8

Moreover, diverse communication protocols exist. If a certain frequency band is chosen, normally the related standard should be stringently adhered. Otherwise there would be co-existence problems, causing communication failures. For instance, the power spectral density of UWB systems is stringently limited to -41.3 dBm/Hz, and there is no design freedom with this parameter during the system optimization process. For IR-UWB systems, the PHY-layer parameters are well described in IEEE 802.15.4a standard [7]. As a result, the design freedom is relatively reduced.

1.4 Scope of the thesis

As discussed in Section 1.3 A, the degree of system optimization will be determined by the author's scope and knowledge, or deliberatively limited for complexity reduction. The main limitations on the scope of this thesis are:

- Cross-layer optimization with emphasis on the PHY-layer. The possibilities and potential solutions of higher layers will be discussed in order to show the feasibility of the proposed physical system. However, the discussions are not in-depth for all aspects, except the closely related ones. A new power management method is provided that suits the most to our physical system. However, the detailed algorithm is not given.
- Generic solution. As a first attempt, this thesis focuses mainly on proving new communication methods and system architectures and focuses on the design of key blocks which are compatible with the proposed communication method and system. As a result, the proposed system is not made to adhere to any communication standard. However, as a short-range, burst-mode, high-speed and directional communication system, it will cause less problems compared to those long-range, always-on, low-rate and omni-directional communication systems in the PHY-layer as long as it does not conflict with existing protocols. On the other hand, for a suitable MAC protocol, more design work will be needed in a next step.
- Choice of some PHY parameters. In Chapter 3, it is shown that the energy efficiency can be optimized at higher frequency bands with high data rate and directional communication methods. The available candidates include: 60 GHz band, 71 to 76

[5] It refers to equivalent isotropic radiated power, i.e. the output power plus antenna gain.

GHz bands, 81 to 86 GHz, 92 to 95 GHz and even higher industrial, scientific and medical (ISM) bands. However, due to the limitations of the chosen technology, only one or two bands can be chosen for the design of the system and the implementation in circuit level.

- Design methodology. A new method is proposed in this thesis to design a communication system that is able to achieve high speed and low power communication simultaneously. In Chapter 5, a design example is given and the design methodology is summarized. However, due to the complexity of the system, there are too many inputs and optimization factors among all communication layers. As a result, some parameters are fixed as assumptions in our proposed design methodology, which would make it less generic but more specific, e.g. towards particular modulation schemes or specific circuitry. Nevertheless, it can still offer an overall solution with regards to our research question.
- Limited design cycle. It may also happen that the specific circuits are not fully optimized, because the initial target of designing these circuits like phase shifters and low-noise amplifiers (LNA) is to verify the system design and innovation, but not to push circuits themselves to the level with the most aggressive FOMs. In fact, however, the optimum gain, NF and linearity distribution of each block in a system should follow a way which leads to overall optimization (for example, following the model described in [9]) instead of separate optimization.
- Chosen technology. In this work, the TSMC 65-nm CMOS technology is chosen to implement our proposed system and blocks due to its high integration level, low cost (per unit) and acceptable leakage, f_{max} and noise level at our system frequencies. However, the systems and designs are not constrained by any particular technology. When shifting to other technologies like III-V ones, no fundamental difficulties or limitations are expected.

1.5 Original contributions

The main contributions of the work presented in this thesis include:

- A cross-layer power optimization. A new FOM is derived to evaluate the power and communication efficiency of a wireless system. In the PHY-layer, optimum frequency bands for high efficiency communication are identified by analyzing the fundamental relationships among noise, pathloss, antenna gain and frequency. It is shown that by adopting directional antennas, the frequency-induced loss (e.g. pathloss, higher noise, etc.) can be compensated by the frequency induced gain (i.e. higher antenna gain due to better directivity at higher frequency). The optimum peak data rate is identified for the rendezvous and directional communication system. A new MAC power management, i.e. asynchronous duty-cycled wake up scheme is

proposed especially for our high-speed system in order to minimize its average power consumption. Some network layer issues are discussed as well.

- The complete "wireless wire" system (as discussed in Chapter 4). It is designed and optimized in a hierarchical manner. Key blocks (e.g. the injection-locking oscillator (IJLO)) are identified and discussed by using mathematical models. System level trade-offs are discussed and optimized. The critical trade-off between sensitivity and locking range of an IJLO is solved by using our proposed frequency-sweeping locking method. Its effectiveness is discussed as well.
- Complete 70-GHz wake-up receiver (WuRx). It is designed and implemented in TSMC 65-nm CMOS technology. Measurement results are provided as well. By adopting our proposed self-mixing demodulation structure, typical high-power blocks like a phase-locked-loop (PLL) and local oscillator (LO) generation module are avoided while the frequency accuracy is not impaired. In such a WuRx, key blocks as IJLO, passive mixer and output buffer amplifier are designed, implemented and tested. The digital control blocks are simulated by using TSMC 65-nm CMOS models.
- Design and simulation of 60-GHz 4-element phased array Rx. A design methodology for a high-frequency power-constrained LNA is provided followed by several design examples of 60-GHz low-power LNAs. Two types of 60-GHz RF phase shifters are designed and simulated, which are able to reach the same performance level (e.g. power consumption, insertion gain and NF) compared to state-of-the-art results in similar technologies. Simulations are done on two versions of the complete Rx chain by adding the self-mixing demodulation module to the phased array Rx.

1.6 Outline of thesis

The state-of-the-art results about low-power Rx architectures, low-power WuRx architectures, and mmW phased array techniques are stated in Chapter 2.

In Chapter 3, the cross-layer power optimization is carried out. Some key PHY parameters are identified towards the maximum power efficiency and minimum average power consumption. The asynchronous duty-cycled wake-up scheme MAC power management method is proposed and compared with other widely used power-efficiency MAC methods. Possible network configurations, routing methods and interference issues are reviewed and discussed.

The complete system architecture is given in Chapter 4. Afterwards, several energy scavenging technologies are discussed and compared. The power optimization models of a multi-element directional communication system are provided. Several system level trade-offs and system mathematical models are analyzed and optimized.

A 70-GHz IJLO-based self-mixing WuRx is shown in Chapter 5. By using the frequency-sweeping locking method, the trade-off between sensitivity and locking range (bandwidth) is converted into the trade-off between sensitivity and circuit sweeping time. Besides, the WuRx sensitivity is also in trade-off with the WuRx detection time. As a result, the WuRx is designed with wideband and narrowband modes. The former is used in situations that require fast response of the WuRx, and the latter is used when the sensitivity is the ultimate important factor, e.g. when the communication range is large.

Two versions of 60-GHz phased array Rxs are designed and simulated in Chapter 6. The main work includes: design and simulations of the 60-GHz LNAs, the 60-GHz low-loss reflection-type RF phase shifter, the 60-GHz 4-bit vector-sum type phase shifter and the complete Rx chains.

Last but not least, conclusions and recommendations are given in Chapter 7.

References

[1] Moore's Law, www.wikipedia.org, 2011.

[2] IEEE 802.15 WPAN Task Group 4 (TG4), *Low Rate Alternative PHY*, 2011.

[3] IEEE 802.15 WPAN Task Group 3c (TG3c), *Millimeter Wave Alternative PHY,* 2009.

[4] Yikun Yu, Design Methods for 60GHz Beamformers in CMOS, PhD dissertations, Eindhoven University of technology, 2010.

[5] IEEE 802.11, www.wikipedia.com.

[6] Wireless Gigabit Alliance, *WiGig White Paper — defining the future of Multi-gigabit Wireless Communications*, www.wigig.org, 2010.

[7] IEEE 802.15.4a WPAN Task Group 4a (TG4a), *Low Rate Alternative PHY for WPANs*, 2011.

[8] Liu Huabo et al., "The Barrel Theory Based Decision-making Algorithm and Its Application," in *2009 International Conference on Computational Intelligence and Natural Computing*, China, 2009.

[9] Peter Baltus, *Minimum Power Design of RF Front Ends*, PhD dissertations, Eindhoven University of technology, 2004.

[10] Hammad Mehmood Cheema, *Flexible Phase-Locked Loops and Millimeter Wave PLL Components for 60-GHz Wireless Networks in CMOS*, PhD dissertations, Eindhoven University of technology, 2010.

[11] Radio Frequency and Analog/Mixed-Signal Technologies for Wireless Communications, *International Technology Roadmap for Semiconductors*, 2009 edition.

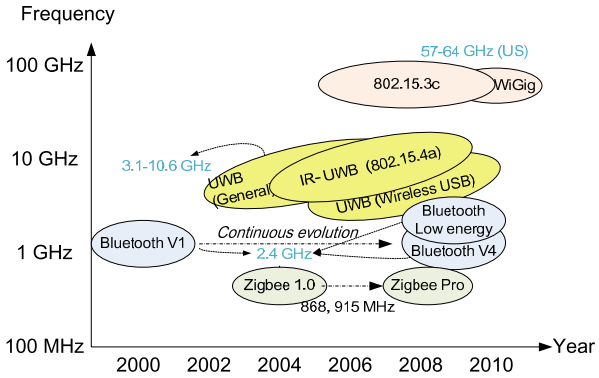
[12] Usha Gogineni *et al*., "Power Potential of 45 nm CMOS Technology," in *Proceedings of IEEE 10th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF2010)*, US, 2010.

CHAPTER 2 STATE OF THE ART

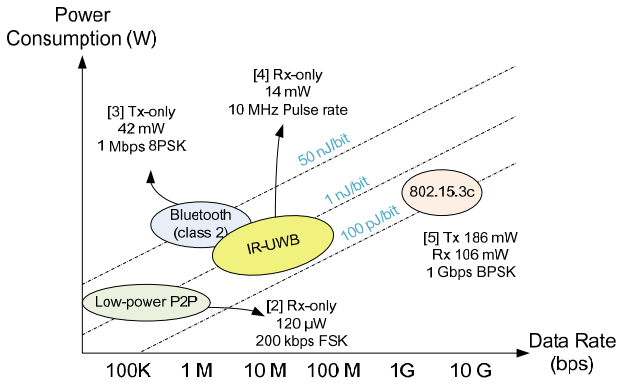
In this chapter, a short summary of the current state and the trends of wireless systems (particularly about personal communications) is given. State-of-the-art design examples and development trends of low power, mmW and WuRx systems are discussed as well.

2.1 Status and trends of wireless systems

Communication data rates have been increasing during the last decade, as can be observed in Fig. 2.1 (a). Although some (relatively) low frequency systems like Bluetooth also grow horizontally, emerging high frequency technologies like the mmW communications (57-64 GHz, 71-76 GHz and 81-86 GHz, etc.) become increasingly important. One of the reasons is shown in Fig. 2.1 (b): with higher data rates, it is possible for these mmW systems to achieve better energy efficiency.



(a)



(b)

Fig. 2.1 Status and trends of wireless systems on personal communications.

Besides the high speed and the possibility to achieve low energy per bit (E_{bit})[6], high frequency (e.g. mmW) communications also provide advantages like smaller electronics feature sizes, smaller directional antennas, better security and higher frequency reuse factor [6]. In addition, compared to the crowded spectrum at lower frequency bands like 2.4 GHz, mmW communication spectrum is relatively empty. As a result, more versatile applications and design freedoms become feasible. For instance, an 87 GHz transceiver front-end is reported in [7] that utilizes unconventional baseband-less carrier and data recovery circuitry and it is able to achieve up to 3.5 Gbps data rate with less than 10^{-11} BER. In 2010, 60 GHz indoor personal communications was standardized by WiGig as part of a tri-band communication system. By adopting beamforming techniques, the 60 GHz system is able to provide up to 7 Gbps data rates that can be used for wireless P2P, wireless display, cordless computing and also Wi-Fi over 60 GHz [8].

2.2 Low power techniques

In this section, some widely used low-power architectures and techniques are discussed followed with state-of-the-art design approaches and examples.

2.2.1 Impulse radio UWB

IR-UWB is a well-known low power and low cost wireless technology for short range personal communication [9]. According to the Shannon theorem, for a constant channel capacity, the required SNR at the demodulator of an IR-UWB system can be very small due to its wide bandwidth. As a result, a low power Tx can be realized. Early in 2005 to 2007, Inter-Universitair Micro-Electronica Centrum (IMEC) designed the first fully digital IR-UWB Tx that is compatible to the IEEE 802.15.4a standard. Its peak power consumption is 0.65 mW with 1 Mbps data rate and E_{bit} is 0.65 nJ/bit [10]. However, some drawbacks make the development of IR-UWB system slacken afterwards. Firstly, conventional IR-UWB systems have low immunity to interference. Also, it is difficult for them to capture multi-path energy. Secondly, the requirement of high speed (e.g. Gsamples per second) and wide dynamic range ADC increases Rx power consumption significantly [11]. Thirdly, when adopting coherent detection methods in the Rx front-end, large amounts of power would be consumed in the synchronization process, which results in E_{bit} value of the Rx ten times larger than that of the Tx [9].

In order to reduce the Rx power consumption, several methods are adopted in IR-UWB systems such as simple modulation schemes (e.g. OOK), duty cycling and non-coherent

[6] It should be emphasized that in this thesis, E_{bit} refers to energy consumption per bit while E_b (normally used in E_b/N_0, i.e. the SNR per bit) refers to the energy transmitted per bit.

energy detection approaches. The basic principle of energy-detection non-coherent IR-UWB Receivers is illustrated in Fig. 2.2 [12], while various design details and techniques can be found in the literature. In [12], the proposed energy detector (i.e. the combination of squarer and integrator in Fig. 2.2) is able to achieve a sensitivity of -89 dBm and an E_{bit} of 0.2 nJ/bit with 100 kbps data rate. With a similar non-coherent detection approach, the OOK IR-UWB Rx front end and baseband in [9] consume 1.64 mW to 2.18 mW (with 40% duty cycle) with 1 Mbps data rate, which corresponds to E_{bit} values from 1.64 nJ/bit to 2.18 nJ/bit.

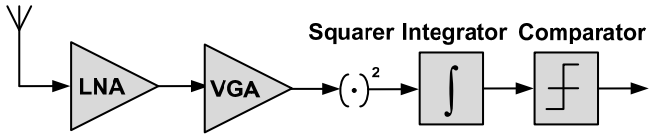


Fig. 2.2 Block diagram of a non-coherent IR-UWB Rx with energy detection approach.

2.2.2 Super-regenerative receiver

The conventional super-regenerative receiver uses a second low frequency oscillator[7] to periodically quench the main RF oscillation in order to provide high but non-linear gain. The basic topology is shown in Fig. 2.3. Although it is disputed due to drawbacks like poor selectivity, susceptibility to front end overload, lack of stability and restriction to certain types of modulation schemes like OOK modulation, it is, in principle, able to achieve very high RF gain with very low power consumption [11]. For that reason, it is included in this overview of low-power techniques.

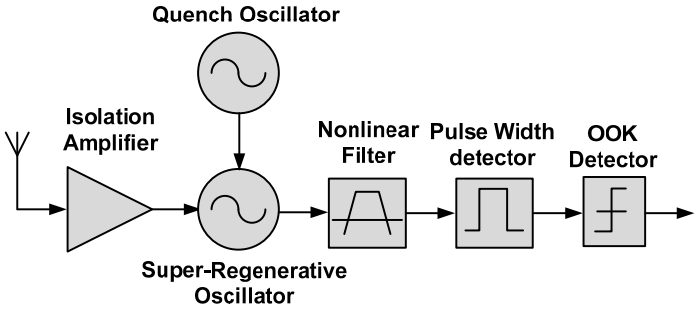


Fig. 2.3 Block diagram of a conventional super-regenerative Rx.

---

[7] In fact, the super-regenerative Rx can operate in linear or logarithmic modes. In the linear mode, the amplitude of the core oscillator is measured before it goes into saturation and then the core oscillator is disabled immediately. In the logarithmic mode, however, two different implementation approaches can be used: full quenching by the second low frequency oscillator or self-sequencing. In the sequencing super-regenerative Rx, no quenching oscillator is required [13].

A 402-to-405MHz super-regenerative receiver is presented in [14]. It is implemented in 0.18-μm CMOS technology. It consumes 0.5mW power at a data rate of 120 kbps and achieves a sensitivity of -95dBm, which corresponds to 4.17 nJ/bit.

In [15], an ultra-low power super-regenerative RF front end is reported for medical body-area network (BAN) applications. It is implemented in 90-nm CMOS technology. The proposed Rx operates at the 2.36-2.4 GHz and 2.4-2.485 GHz ISM bands, and it consumes 500 μW power with 1 to 5 Mbps data rate (i.e. 0.5 to 0.1 nJ/bit). In addition, it is able to achieve a sensitivity of -67 dBm at a BER of 10^{-3}.

Although these state-of-the-art super-regenerative Rxs are able to achieve sub-mW power consumption, one significant problem is that they have speed limitation besides other common disadvantages listed at the beginning of this section.

2.2.3 Back-scattering architecture for asymmetrical links

The back scattering technique is widely used in asymmetric wireless applications like radio-frequency identification (RFID) that uses radio waves to exchange data between a reader and a transponder or a tag [16]. Its basic architecture is illustrated in Fig. 2.4 (a).

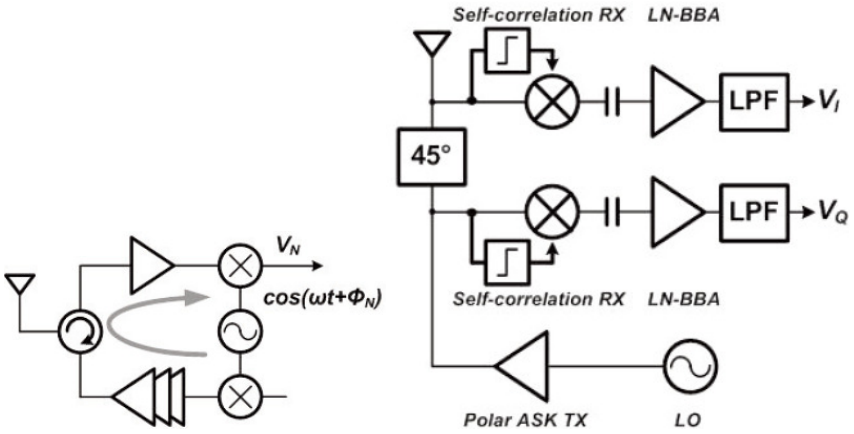


Fig. 2.4 (a) Block diagram of a conventional back-scattering transceiver and (b) a transceiver consisting of a polar Tx and a self-correlation Rx for RFID application [17].

By adopting the self-correlation structure in the Rx, as shown in Fig. 2.4 (b) [17], the requirement of a Tx-Rx isolator is eliminated, and the proposed Rx is able to achieve -75 dBm sensitivity with 10 mA current consumption from a 3.3-V supply. An obvious disadvantage of backscattering technique is that this approach is only suitable for asymmetric wireless systems instead of a generic solution for low power radios.

2.3 Millimetre-wave communications

Three major topics are observed for mmW wireless communications:

- Low power and low cost
  Compared to lower frequency communication systems, low power and low cost mmW transceivers tend to adopt simpler transceiver architectures, such as direct conversion, sliding/zero IF and non-coherent envelope detection (for Rx) [18]. In [18], a 60-GHz direct conversion Rx is realized in 65-nm CMOS technology, which is able to achieve 30-dB front-end gain, 9.2 dB NF and -36 dBm input-referred third-order intercept point (IIP3) with 43 mW power consumption. A digitally controlled 60-GHz Rx is reported by IMEC in 2009. Its noise figure is about 6 dB and power consumption is only 21 mW[8]. Furthermore, fully digital control makes it can be readily integrated in a multi-element phased array [19].
- High performance
  In order to achieve high performance like high data rate and low BER, usually much more complicated transceiver architectures and/or modulation schemes are required. Consequently, it is not surprising that system power consumption of such high performance systems is also higher. In 2011, a fully integrated sliding IF 60GHz transceiver in a 65nm CMOS technology for wireless high-definition video streaming has been presented in [20]. The CMOS chip is compatible with the WirelessHD^{TM} standard and is able to cover four channels. The 16-QAM OFDM modulation scheme is adopted. The proposed transceiver front end consumes 454mW in receiver mode and 1.357W in the transmit mode (357mW for the transmitter and 1W for the PA). The Rx is able to achieve 3.8 Gbps data rate with 11% error vector magnitude (EVM) over 1 meter communication distance. In [21], a 60-GHz direct conversion transceiver in 65 nm CMOS technology is reported, which is compatible with the IEEE 802.15.3c standard. The proposed transceiver front end consumes 186 mW, 106 mW and 66 mW for Tx, Rx and PLL, respectively. The maximum data rates with an antenna built-in the package are 8Gbps in QPSK mode and 11Gbps in 16QAM mode within a BER less than 10^{-3}.
- Multi-element phased array front end
  Besides advantages like spatial filtering, high security and additional antenna gain, the directional communication system can be optimized to achieve either low power or high performance. Among several directional communication methods, the multi-element phased array beamforming technique provides benefits like high integration level, potential for low cost and low power, and suitability for digital

[8] Additional power consumed in the phase-locked-loop (PLL) is 78 mA from a 1.1 V supply.

control[9] and decent beam pointing resolution. As a result, it is widely used. In [22], a low power 60GHz 4-element 65nm CMOS phased array transceiver is presented. The total measured power consumption is 137mW in both Rx and Tx mode, including 29mW for the synthesizer. With 3-bit digitally controlled phase shifters (plus an extra polarity bit for controlling the sign of the phase), the Rx is able to achieve a worst-case phase resolution of 11°. In the high performance direction, a 60-GHz multi-Gbps phased array transceiver is shown in [23]. The proposed Rx is able to support 4-Gbps data rate with a 10 meters communication distance and a BER less than 10^{-11}. Its total power consumption is up to 2 W. As a result, the potential applications as well as device mobility are limited.

2.4 Wake-up receiver

Power management is one of the most important topics in wireless communication, especially for low power applications. Among all the different methods, wake-up receivers (WuRx) form a very interesting category, because, in principle, it can reduce the standby power of the main radio very efficiently. If the power consumption of the WuRx itself can be minimized, the mobility of the entire communication system can be increased significantly. The state-of-the-art low-power WuRxs are listed in Table 2.1.

Table 2.1 Review of state-of-the-art WuRxs

Types	**Performance**				
	Frequency	**Sensitivity (dBm)**	**Power consumption**	**Data Rate**	**E_{bit} (nJ/bit)**
Diode [24]	1.9 GHz	-48	64 μW	100 kbps	0.64
Diode with simple Rx [25]	915 MHz	-80	51 μW	10 kbps	5.1
	2.4 GHz	-69	51 μW	10 kbps	5.1
Uncertain-IF [26]	1.9 GHz	-72	52 μW	100 kbps	0.52
Super-regenerative [15]	2.4 GHz	-67	500 μW	1-5Mkbps	0.1 to 0.5
Duty-cycling [27]	2.4 GHz	-82	415 μW	500 kbps	0.83
Multi-stage charge pump [28][10]	Generic	NA	0 (passive)	0	-

Though very different approaches and methodologies are used, several common features

[9] It depends on detailed circuit designs and baseband algorithms.

[10] This is a concept without circuit design and implementation. Sensitivity of the proposed structure in [27] is traded off with circuit latency. Due to the zero-bias feature, this radio-triggered WuRx can be used as an RF power harvester as well.

can be observed in the WuRxs above:

- Relatively low signal frequencies, narrow bandwidths and low data rates.
- Sub-mW power levels.
- Trade-offs between power consumption, data rate, and sensitivity and carrier frequency at the same BER level are observed in [24] to [27]. When increasing the carrier frequency and data rate, a WuRx will suffer from poor sensitivity under the same signal and power conditions. However, this trade-off can be avoided by using a duty-cycled scheme with lower duty-cycle factor (DCF) as suggested in [27]. Furthermore, the minimum DCF is limited by the detection time. In other words, if a fast-response[11] detector can be designed at high frequencies like mmW, it is feasible for the WuRx to operate with an ultra-low duty cycle. Consequently, a high frequency, high data rate, high sensitivity, low power and low latency WuRx could be obtained.

References

[1] Comparison of Standards, www. wikipedia.org, 2011.

[2] Jagdish Pandey, Jianlei Shi, Brian Otis, "A 120μW MICS/ISM-Band FSK Receiver with a 44μW Low-Power Mode Based on Injection-Locking and 9xFrequency Multiplication," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[3] Hiroyuki Kobayashi et al., "An All-Digital 8-DPSK Polar Transmitter with Second-Order Approximation Scheme and Phase Rotation-Constant Digital PA for Bluetooth EDR in 65nm CMOS," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[4] Ta-Shun Chu, "A Short-Range UWB Impulse-Radio CMOS Sensor for Human Feature Detection," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[5] Kenichi Okada et al., "A 60GHz 16QAM/8PSK/QPSK/BPSK Direct-Conversion Transceiver for IEEE 802.15.3c," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[6] Cheolhee Park *et al*., "Short-Range Wireless Communications for Next-Generation Networks: UWB, 60 GHz Millimeter-Wave WPAN, and ZigBee," in *IEEE Wireless Communications,* pp. 70-78, 2007.

[7] Shih-Jou Huang *et al.*, "An 87GHz QPSK Transceiver with Costas-Loop Carrier Recovery in 65nm CMOS," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[8] Wireless Gigabit Alliance, *WiGig White Paper — defining the future of Multi-gigabit Wireless*

[11] The speed of the WuRx is determined by application requirements.

Communications, www.wigig.org, 2010.
[9] Marco Crepaldi *et al*., "An Ultra-Low-Power Interference-Robust IR-UWB Transceiver Chipset Using Self-Synchronizing OOK Modulation," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2010.
[10] Ryckaert, J. *et al*., "A 0.65-to-1.4 nJ/Burst 3-to-10 GHz UWB All-Digital TX in 90 nm CMOS for IEEE 802.15.4a," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2860-2869, 2007.
[11] Emanuele Lopelli, Transceiver Architectures and sub-mW fast Frequency-Hopping Synthesizers for Ultra-Low Power WSNs, PhD Dissertation, Eindhoven University of Technology, 2010.
[12] Andrea Gerosa et al., "An Energy-Detector for Noncoherent Impulse-Radio UWB Receivers," in *IEEE Transactions on Circuits and Systems—I: Regular Papers*, vol. 56, no. 5, pp.1030-1040, 2009.
[13] Brian P. Otis, *Ultra-Low Power Wireless Technologies for Sensor Networks*, PhD Dissertations, University of California, Berkeley, 2005.
[14] Anis, M. *et al*., "Fully integrated UWB impulse transmitter and 402-to-405MHz super-regenerative receiver for medical implant devices," in *IEEE International Systems and Circuits Conference (ISCAS)*, 2010.
[15] Vidojkovic, M. et al., "A 500mW 5Mbps ULP super-regenerative RF front-end," in *IEEE European Solid-State Circuits Conference (ESSCIRC)*, 2010.
[16] Radio Frequency Identification, www.wikipedia.org, 2011.
[17] Eun-Hee Kim et al., "An Isolator-less CMOS RF Front-End for UHF Mobile RFID Reader," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.
[18] Michael M. Kraemer, *Design of a low-power 60 GHz transceiver front-end and behavioral modeling and implementation of its key building blocks in 65 nm CMOS*, PhD Dissertation, L'Institut National des Sciences Appliquées de Toulouse, 2010.
[19] Low-Cost Low-Power 60GHz Solutions In Digital 45nm CMOS, www.sciencedaily.com, 2009.
[20] Alexandre Siligaris et al., "A 65nm CMOS Fully Integrated Transceiver Module for 60GHz Wireless HD Applications," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.
[21] Kenichi Okada et al., "A 60GHz 16QAM/8PSK/QPSK/BPSK Direct-Conversion Transceiver for IEEE 802.15.3c," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.
[22] Maryam Tabesh et al., "A 65nm CMOS 4-Element Sub-34mW/Element 60GHz Phased-Array Transceiver," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.
[23] Sohrab Emami *et al*., "A 60GHz CMOS Phased-Array Transceiver Pair for Multi-Gb/s Wireless Communications," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.
[24] Nathan M. Pletcher *et al*., "A 65μW, 1.9 GHz RF to Digital Baseband Wakeup Receiver for Wireless Sensor Nodes," in *IEEE Custom Integrated Circuits Conference (CICC)*, 2007.
[25] Xiongchuan Huang *et al*., "A 1.4GHz/925MHz 51μW Wake-Up Receiver with Offset and Noise Suppression," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2010.
[26] Nathan M. Pletcher *et al*., "A 52 μW Wake-Up Receiver With -72 dBm Sensitivity Using an Uncertain-IF Architecture," in IEEE Journal of Solid-State Circuits, vol. 44, no.1, January 2009.
[27] Salvatore Drago *et al*., "A 2.4GHz 830pJ/bit Duty-Cycled Wake-Up Receiver with -82dBm Sensitivity for Crystal-Less Wireless Sensor Node," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2010.
[28] Lin Gu *et al*., "Radio-Triggered Wake-Up for Wireless Sensor Networks," in *Real-Time Systems*, vol. 29, no. 2-3, pp. 157-182, 2005.

CHAPTER 3 CROSS-LAYER POWER OPTIMIZATION

The optimization process of a wireless system towards the low-power performance is not straightforward but more likely involves the coordination and trade-offs among different communication layers. For example, a low-power performance in the PHY layer would probably mean that the total power consumption or the design difficulty is not actually reduced but is shifted to the MAC or other layers. In other words, one may need an ultra-complicated and power-inefficient protocol to control the communication with using an ultra-simple and "non-smart" physical node. The overall power consumption is not necessarily optimized when different perspectives are considered or optimized orthogonally, which is, unfortunately widely seen in most of the design approaches. Therefore, a cross-layer optimization method becomes important and useful. However, there are too many variable parameters in each layer and some of them are correlated with each other, which makes the optimization unobvious and indirect, as shown in Fig. 3.1. As a result, a cross-layer figure-of-merit (FOM) is crucial to indicate the direction for the optimization as well as to evaluate its effectiveness [1].

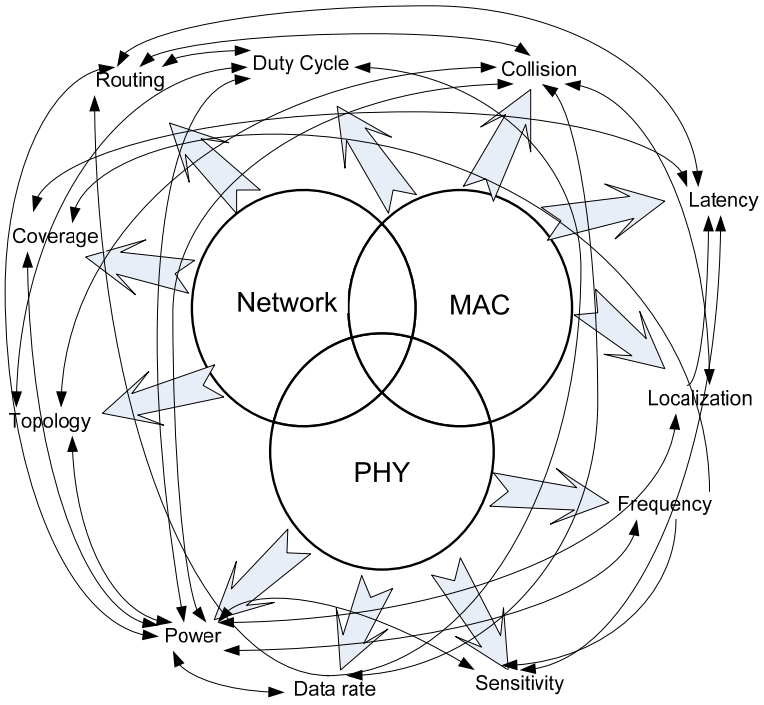


Fig. 3.1 Parameters in different layers of a wireless system.

From Fig. 3.1, it can be seen that, for example, the duty-cycle factor is correlated to the power consumption, network topology, routing method and communication data rate.

Intuitively speaking, when decreasing the duty-cycle factor, the system power consumption will be decreased because of the shorter active time of the system. However, the Rx has to spend more time to listen to the channel in order not to miss the transmission, which on the other hand will increase the overall power consumption of the radio system. Besides, the data rate of the system may need to be increased in order to finish transmission of a completed packet within one cycle, which again increases the power consumption of the physical circuits.

3.1 Figure-of-merit

The total delay of a single node in a wireless network can be expressed as

$$\tau_{tot} = \tau_{Propagation} + \tau_{Serialization} + \tau_{Queuing} + \tau_{Processing} \tag{3.1}$$

where $\tau_{Propagation}$ is the propagation delay defined by distance over wave propagation speed, $\tau_{Serialization}$ is the serialization delay, which is defined as the time of starting to ending of a reception, $\tau_{Queuing}$ is the time consumed by waiting for forwarding a packet, and $\tau_{Processing}$ is the time for the relay nodes to process a packet.

As a design assumption, the communication frequency, communication distance, network topology and packet forwarding mechanism are pre-determined in the system-level optimization. The total delay τ_{tot} can then be simplified as

$$\tau_{tot} = \tau_C + \tau_{Serialization} \tag{3.2}$$

where τ_C is the minimum latency or the summation of $\tau_{Propagation}$, $\tau_{Queuing}$ and $\tau_{Processing}$, which can be taken as a design constant for a certain type of networking configuration, and $\tau_{Serialization}$ can be calculated by the Rx overhead $t_{Rx,overhead}$ plus the packet length L_{pkt} over the peak raw data rate R_{raw}, i.e.

$$\tau_{tot} = \tau_C + t_{Rx,overhead} + L_{pkt} / R_{raw} \tag{3.3}$$

It is known that the data rate is historically sub-linear to the power consumption of front end circuits [2]. Therefore, the total power consumption of a node can be modeled as

$$P_{DC,tot} = P_{pr,DC} + P_{Rx,DC} + k_e \cdot R_{raw} \tag{3.4}$$

where $P_{pr,DC}$ is the power consumption of a Rx for pre-receiving activities (e.g. listening, synchronization and settling, etc.), $P_{Rx,DC}$ is the Rx power consumption and k_e is the electronics factor, which is determined by the PHY parameters like modulation scheme, bit-error-rate (BER) level, technology level, communication range, communication frequency, and circuits and antenna designs (The detailed model of k_e can be interpreted from (3.19)). Solving for R_{raw} from (3.3), it can be obtained that

$$R_{raw} = \frac{L_{pkt}}{\tau_{tot} - \tau_C - t_{Rx,overhead}} \tag{3.5}$$

Substituting (3.5) into (3.4), $P_{DC,tot}$ can be expressed as

$$P_{DC,tot} = \frac{L_{pkt} \cdot k_e}{\tau_{tot} - \tau_C - t_{Rx,overhead}} + P_{pr,DC} + P_{Rx,DC} \tag{3.6}$$

According to the definition of E_{bit}, the system energy efficiency can be evaluated by the power consumption over data rate. It is a fair FOM to evaluate the energy efficiency of an always-on communication system. However, some important information is missing in this FOM such as the type of operations (duty-cycled or always-on) and the quality of the communication (the BER). Consequently, it does not directly indicate the level of the system average power consumption, which is the parameter determining the battery life as well as the mobility of a node. Therefore, replacing the peak power by the average power, the average energy per bit of correctly received bits ($E_{ave,corr}$) is calculated as

$$E_{ave,corr} = \frac{P_{DC,tot}}{R_{raw,corr}} \cdot \frac{t_{on}}{t_{tot}} \tag{3.7}$$

where $R_{raw,corr}$ is the correctly-received raw data rate, which is approximately calculated by the peak raw data rate multiplied by (1-BER), t_{on} is the active time of a radio and the t_{tot} is the total observing time. Substituting (3.6) into (3.7), the FOM can be expressed as

$$FOM = \frac{\frac{L_{pkt} \cdot k_e}{\tau_{tot} - \tau_C - t_{Rx,overhead}} + P_{pr,DC} + P_{Rx,DC}}{R_{raw,corr}} \cdot \frac{t_{on}}{t_{tot}} \tag{3.8}$$

where τ_{tot} should be smaller than an application determined factor τ_{max}, i.e. the maximum user-happy delay, which can be obtained from the quality-of-experience (QoE) experiments. Equation (3.8) can be further rewritten as

$$FOM = \left(\frac{L_{pkt} \cdot k_e}{\tau_{tot} \cdot R_{effect}} + \frac{P_{pr,DC} + P_{Rx,DC}}{R_{raw,corr}}\right) \cdot \frac{t_{on}}{t_{tot}} \tag{3.9}$$

where R_{effect} indicates the effective payload transmission rate over the time period τ_{tot} and R_{effect} can be calculated by

$$R_{effect} = \left[1 - \left(\frac{\tau_C + t_{Rx,overhead}}{\tau_{tot}}\right)\right] \cdot R_{raw,corr} \tag{3.10}$$

From (3.9), it can be seen that a low FOM level may indicate an efficient communication (with a fixed amount of data length L_{pkt}), i.e. either a low-latency Rx system, or a low BER level and low average power dissipation. Some trade-offs can be observed according to (3.9) and (3.10) as:

- Reducing the duty-cycle factor t_{on}/t_{tot}, the FOM will be scaled down at the first hand. However, $P_{Pr,DC}$ will increase and R_{effect} will decrease (due to the increase of $t_{Rx,overhead}$), which will slow down the decrease trend and finally lead to an increasing FOM. Overall speaking, an optimum should exist in between.
- Increasing the raw data rate R_{raw} leads to a decrease of the system BER, which in turn decreases the $R_{raw,corr}$ and R_{effect}. Consequently, the FOM is not scaled or optimized linearly and an optimum state should exist.
- Minimizing $t_{Rx,overhead}$ seems to be a straightforward method to improve the energy efficiency. However, the degree of minimization is limited by the detailed circuit and system designs. Besides, it generally leads to an increase of the $P_{pr,DC}$ and $P_{Rx,DC}$, because a shorter overhead is normally less informative and thus the Rx may need more power to compute and interpret it in order to launch the correct receiving process. A good example is the asynchronous system. Without the time-consuming synchronization process, the receiver has to put more power on the non-coherent demodulation. As a result, the overall FOM is not directly optimized.

In conclusion, the FOM indicates the power-correlated parameters across different layers. Several transforms, trade-offs and optimizations can be made by using this model. Vice versa, when a system is designed, the effectiveness of its low-power performance can also be evaluated by using the FOM. The detailed implementations are described in the following sections.

3.2 PHY-layer power optimization

Two FOMs are widely used to optimize the power performance of a wireless system, i.e. the energy per bit E_{bit} and the average power consumption P_{ave}. In order to achieve a truly low-power and high-efficiency communication system, both should be optimized.

3.2.1 Optimum frequency for minimum E_{bit}

The system noise factor (F) can be modeled as a sub-linear function of the operation frequency (f) as [3]

$$F = 1 + \frac{f}{k_{Tech}} \tag{3.11}$$

where k_{Tech} is a constant value for a given IC technology. According to benchmarks presented in [4-10], k_{Tech} can be taken as $6 \cdot 10^9$ Hz for the 65-nm CMOS technology. System noise figures (i.e. $10 \cdot \log(F)$) can thus be modeled as a function of operation frequency, as shown in Fig. 3.2.

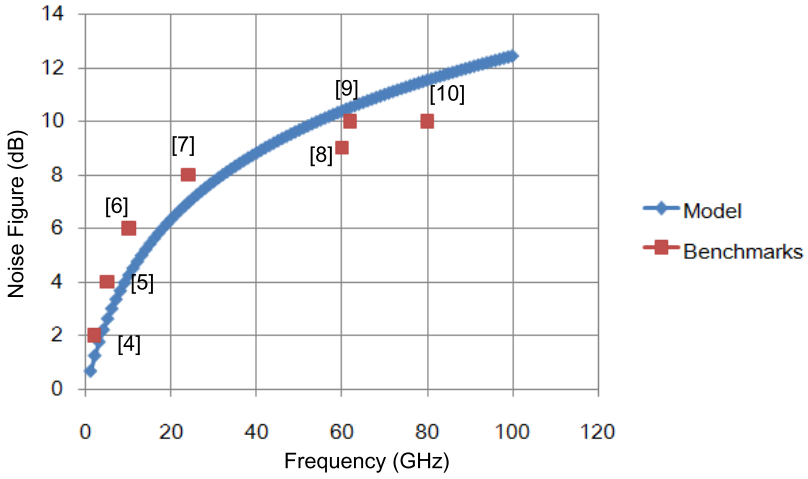


Fig. 3.2 Noise factor as a function of frequency.

According to the Frii's equation, under the line-of-sight (LoS) assumption, the free-space pathloss (PL) of the EM wave is calculated as

$$PL = (\frac{4\pi d}{\lambda})^2 = (\frac{4\pi d \cdot f}{c})^2 \propto f^2 \tag{3.12}$$

where d is the distance, λ is the wave length and c is the speed of light. At the output of a front end, the required demodulation signal-to-noise ratio (SNR) can be calculated as

$$SNR_{out} = \frac{E_b}{N_0} \cdot \frac{R}{B} \tag{3.13}$$

where E_b/N_0 is the SNR per bit, R is the data rate and B is the signal bandwidth. Given a certain BER requirement, E_b/N_0 is a function of the demodulation scheme, as shown in Fig. 3.3. In the low-power WPAN applications, BER is normally taken as 10^{-3}.

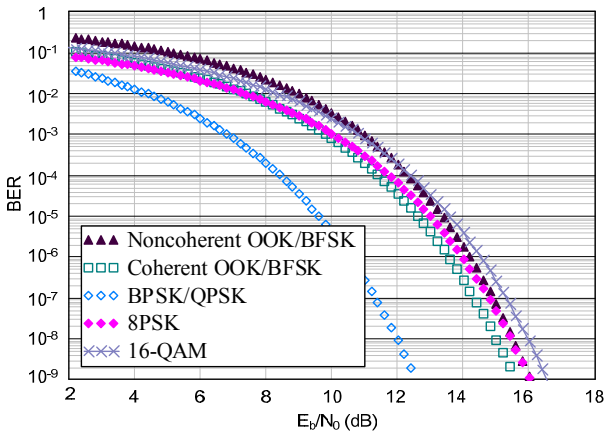


Fig. 3.3 BER as a function of E_b/N_0 in the AWGN channel.

The Rx sensitivity is calculated as a product of the system noise floor (*KTB*), noise factor and the demodulation SNR as

$$S = KTB \cdot F \cdot SNR_{out} \tag{3.14}$$

where K is the Boltzmann constant, which equals $1.38 \cdot 10^{-23}$ J/K and T is the temperature in Kelvin, which is normally taken as 300 K (the room temperature) in indoor application scenarios.

Substituting (3.11) and (3.13) into (3.14), it can be obtained that

$$\begin{aligned} S &= KTB \cdot F \cdot \frac{E_b}{N_0} \cdot \frac{R}{B} = KT \cdot \frac{E_b}{N_0} \cdot (1 + \frac{f}{k_{Tech}}) \cdot R \\ &= (KT \cdot \frac{E_b}{N_0} + KT \cdot \frac{E_b}{N_0} \cdot \frac{1}{k_{Tech}} \cdot f) \cdot R = (A_S f + B_S) \cdot R \end{aligned} \tag{3.15}$$

where A_S and B_S are constant coefficients for a given technology, modulation scheme, temperature and BER level. In other words, with fixing the technology and application related factors, the Rx sensitivity becomes sub-linearly proportional to the RF frequency and linear to the communication data rate.

The minimum output power ($P_{out,min}$) of the Tx then can be estimated as

$$P_{out,min} = \frac{S \cdot PL}{G_{Tx} \cdot G_{Rx}} \tag{3.16}$$

where G_{Tx} and G_{Rx} are the Tx and Rx antenna gain respectively. Substituting (3.12) and (3.15) into (3.16), the minimum output power of the Tx can be written as

$$P_{out,min} = \frac{(A_S f + B_S) \cdot R \cdot (\frac{4\pi df}{c})^2}{G_{Tx} \cdot G_{Rx}} \tag{3.17}$$

or

$$P_{out,min} = \frac{[KT \cdot \frac{E_b}{N_0} \cdot \frac{1}{k_{Tech}} \cdot (\frac{4\pi df}{c})^2 \cdot f + KT \cdot \frac{E_b}{N_0} \cdot (\frac{4\pi df}{c})^2] \cdot R}{G_{Tx} \cdot G_{Rx}} = \frac{(A_{out} f^3 + B_{out} f^2) \cdot R}{G_{Tx} \cdot G_{Rx}} \tag{3.18}$$

where A_{out} and B_{out} are the constant coefficients. It can be seen that the minimum output power of the Tx is related to the cube of the RF frequency (due to the "frequency-induced loss", including the pathloss and the extra noise), and it is proportional to the data rate.

Assuming the transmitter efficiency (η_{Tx}) is a design constant at a certain technology, the minimum DC power consumption of the Tx follows the same pattern as

$$P_{Tx,DC,min} = \frac{P_{out,min}}{\eta_{Tx}} = \frac{(A_{out} f^3 + B_{out} f^2) \cdot R}{\eta_{Tx} \cdot G_{Tx} \cdot G_{Rx}} \tag{3.19}$$

The DC power consumption follows the same pattern. It can be seen that communications at high frequency bands are inherently "lossy" due to the high pathloss and high noise, and they require more output power from the Tx in order to maintain the same link budget (communication range). Therefore, the power consumption of the Tx will increase significantly. The situation would get worse with using higher data rate. However, this extra power requirement can be compensated by extra antenna gain.

The generic antenna gain (G_{ant}) is calculated as the product of the radiation efficiency (η_{rad}) and its directivity (D_{ant}) as

$$G_{ant} = \eta_{rad} D_{ant} \tag{3.20}$$

The antenna directivity of a normal aperture antenna can be calculated by

$$D_{ant} = \eta_{ap} \frac{4\pi A}{\lambda^2} \tag{3.21}$$

where η_{ap} is the aperture efficiency, A is the antenna area and λ is the RF wave length. Substituting (3.21) into (3.20), it can be seen that if keeping the antenna area A constant for all frequencies and neglecting the slight frequency dependency of η_{ap}, the antenna gain becomes inversely proportional to λ^2, i.e. linearly proportional to the f^2 as

$$G_{ant} = \eta_{rad} \cdot \eta_{ap} \frac{4\pi A}{\lambda^2} = \frac{\eta_{rad} \cdot \eta_{ap} \cdot 4\pi A}{c^2} \cdot f^2 \tag{3.22}$$

Substituting (3.22) into (3.19), the minimal DC power of the Tx can be calculated as[12]

$$P_{Tx,DC,min} = \frac{P_{out,min}}{\eta_{Tx}} = \frac{(A_{out} f^3 + B_{out} f^2) \cdot R}{\eta_{Tx} \cdot (\frac{\eta_{rad} \cdot \eta_{ap} \cdot 4\pi A}{c^2} \cdot f^2)^2} = (A_{DC} \cdot f^{-1} + B_{DC} \cdot f^{-2}) \cdot R \tag{3.23}$$

where A_{DC} and B_{DC} are the symbolic constant coefficients, which are again determined only by the technology and application related factors. Consequently, the minimum E_{bit} of the Tx becomes

$$E_{bit,Tx,min} = \frac{P_{Tx,DC,min}}{R} = A_{DC} \cdot f^{-1} + B_{DC} \cdot f^{-2} \tag{3.24}$$

which is inversely linear to the frequency (f^{-2} is a negligible small value at very high frequencies) and independent with the data rate, as shown in Fig. 3.4. In this way, the

[12] The antenna areas of Tx and Rx are assumed identical for simplicity while in fact it is not compulsory.

"frequency-caused losses" can be compensated by the "frequency-induced gain" by using high-gain and high-directivity antennas.

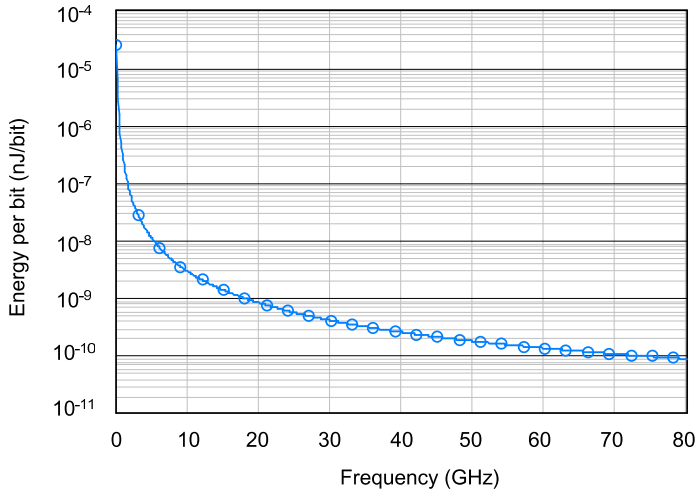


Fig. 3.4 The minimum E_{bit} of the Tx as a function of frequency.

It can be seen that the E_{bit} of the Tx keeps on scaling down when increasing the RF frequency and it is independent with the data rate. Besides, with high data rate, it needs more power for communication but the communication will also be faster, so eventually the overall power efficiency is not affected. In such a directional burst-mode communication system, higher E_{bit} can be achieved at higher frequency bands.

On the other hand, the E_{bit} of the Rx cannot be simply modeled with these link budget parameters. By integrating the benchmarks of the state-of-the-art 65 nm CMOS Rx designs presented in [6-11] into the model, the overall E_{bit} of a transceiver can be illustrated in Fig. 3.5.

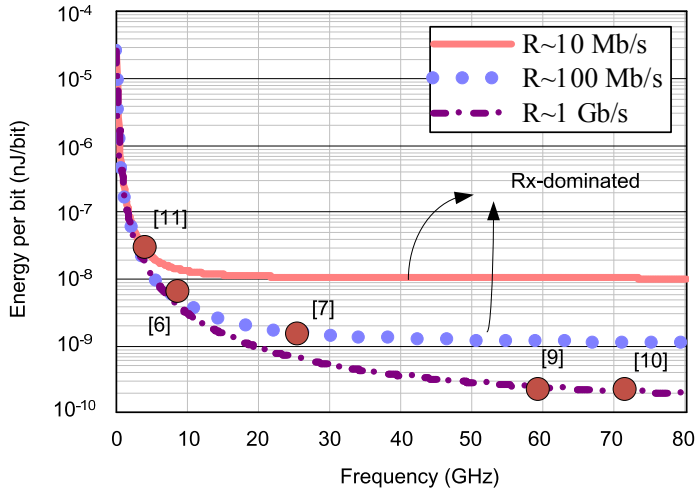


Fig. 3.5 Trend of the E_{bit} scaling with frequency for entire transceivers.

It can be seen that the down-scaling trend of the E_{bit} along with the frequency stops at a certain frequency. It means that under the existing technology level, the Rx power consumption increases very fast with frequency. When the frequency exceeds a limit, the Rx power becomes the dominating factor to the total power consumption of a transceiver. However, if the technology improves in the future, with larger f_T, lower parasitics level and better noise performance, the optimum frequency is expected to shift to higher frequency and eventually following the similar trends as the Tx E_{bit}. Besides, it can be seen from Fig. 3.5 that the E_{bit} is also decreased with larger data rate, which confirms the previous conclusion of the burst-mode transmission. For example, with 1 Gbps peak data rate, the overall E_{bit} achieves its optimum (minimum) from 60 GHz to about 80 GHz[13], which means in 65-nm CMOS technology, these mmW frequencies are the optimum carrier frequency candidates to achieve highly power efficient communication. The theoretical minimum E_{bit} is about 200 pJ/bit according to Fig. 3.5.

Due to the technology limitation, communication systems with higher data rate, e.g. beyond 1 Gbps are not included in this model, because their requirements are very different compared to other low-power wireless systems. These ultra-high data rate systems are normally not power optimized.

3.2.2 Optimum data rate for minimum P_{ave}

As discussed in the previous section, in order to achieve a better E_{bit}, the mmW and high-data-rate communication becomes the optimum choice. However, the system data rate cannot be chosen arbitrarily. The primary boundary condition is the Shannon-Hartley theorem, i.e. the maximum data rate should be compatible with the channel capacity limitation. Moreover, it also relates to the physical circuit performance and the communication method. For example, in a duty-cycled radio system, high data rate means more DC power should be used to generate higher SNR to maintain a certain BER level, but it also implies that the transmission time will be shorter and thus the low-power "sleep time" of the radio is longer. In short, an optimum data rate should exist, which leads to the minimum power consumption.

The optimum data rate is derived as follows. In a wireless front end, the input SNR (SNR_{in}) of the Rx is calculated as

$$SNR_{in} = \frac{P_{out} \cdot G_{TRx}}{PL \cdot KTB} \tag{3.25}$$

where P_{out} is the output power of the Tx and G_{TRx} is the product of the Tx and Rx antenna

[13] In fact, when the frequency goes further up, the results from the literature are not relevant due to very different design objectives, methodologies and system applications, and thus not included in this analysis.

gain. The output SNR (SNR_{out}) can be calculated as the SNR_{in} over the noise factor as

$$SNR_{out} = SNR_{in} / F = \frac{P_{out} \cdot G_{TRx}}{KTB \cdot PL \cdot F} \tag{3.26}$$

From the Shannon-Hartley theorem, the channel capacity (C) is

$$C = R \cdot \varepsilon = B \cdot \log_2(1 + SNR_{out}) \tag{3.27}$$

where ε is the unit-less ratio between data rate and capacity and it is larger than unity in practice. Rewriting (3.27), the SNR_{out} can be expressed as

$$SNR_{out} = 2^{\frac{\varepsilon \cdot R}{B}} - 1 \tag{3.28}$$

Equaling (3.28) to (3.26), it can be obtained that

$$2^{\frac{\varepsilon \cdot R}{B}} - 1 = \frac{P_{out} \cdot G_{TRx}}{KTB \cdot PL \cdot F} \tag{3.29}$$

As a result, the output power of the Tx can be expressed as

$$P_{out} = \frac{KTB \cdot PL \cdot F \cdot (2^{\frac{\varepsilon \cdot R}{B}} - 1)}{G_{TRx}} \tag{3.30}$$

and the DC power consumption of the Tx becomes

$$P_{Tx,DC} = \frac{P_{out}}{\eta_{Tx}} = \frac{KTB \cdot PL \cdot F \cdot (2^{\frac{\varepsilon \cdot R}{B}} - 1)}{\eta_{Tx} \cdot G_{TRx}} \tag{3.31}$$

In a generic rendezvous radio system, the power consumption can be mainly divided by four parts: Tx packet transmitting power ($P_{Tx,DC}$), Rx packet receiving power ($P_{Rx,DC}$) and the power consumed by the other pre-receiving activities like Rx listening, mode switching ($P_{pr,DC}$) and also includes the power dissipated in the idle mode ($P_{idle,DC}$). Regardless the power consumed by the packet overhead or other MAC-layer activities (which will be included into the optimization model in the next section), the total power consumption of the physical system can be calculated as

$$P_{DC,tot} = P_{Tx,DC} + P_{Rx,DC} + P_{pr,DC} + P_{idle,DC} \tag{3.32}$$

If observing the system for a sufficient-long time of T_{tot}, the total average power consumption ($P_{ave,tot}$) becomes

$$P_{ave,DC,tot} = [\frac{P_{out}}{\eta_{Tx}} \cdot T_{Tx} + P_{Rx,DC} \cdot T_{Rx} + P_{pr,DC} \cdot T_{pr} + P_{idle,DC} \cdot (T_{tot} - T_{Tx} - T_{Rx} - T_{pr})] \cdot \frac{1}{T_{tot}} \tag{3.33}$$

where T_{Tx}, T_{Rx}, T_{pr} are the effective transmitting, receiving and Rx pre-listening time (including the mode switching time here) respectively. T_{Tx} equals T_{Rx}, and can be

calculated as the packet length L_{pkt} over the data rate, i.e.

$$T_{Tx} = T_{Rx} = \frac{L_{pkt}}{R} \tag{3.34}$$

Substituting (3.31) and (3.34) into (3.33) and letting $\frac{\partial P_{ave,DC,tot}}{\partial R} = 0$, it can be seen that the solution becomes equivalent to finding the extreme value of the following equation:

$$\frac{\partial}{\partial R}\{[\frac{KTB \cdot PL \cdot F \cdot (2^{\frac{\varepsilon \cdot R}{B}} - 1)}{\eta \cdot G_{ant}} + P_{Rx,DC} - 2P_{idle,DC}] \cdot \frac{1}{R}\} = 0 \tag{3.35}$$

Solving (3.35), a transcendental equation can be obtained as

$$R = \frac{B}{\varepsilon} - \frac{B}{\varepsilon \cdot 2^{\frac{\varepsilon \cdot R}{B}}} + \frac{P_{Rx,DC} - 2P_{idle,DC}}{\frac{KTB \cdot PL \cdot F \cdot \varepsilon \cdot 2^{\frac{\varepsilon \cdot R}{B}}}{G_{TRx} \cdot \eta_{Tx} \cdot B}} \tag{3.36}$$

In order to find the solution, the data of a 60-GHz wireless system[14] is taken into (3.36), i.e. ε=2, B=7·10^9 Hz, $P_{Rx,DC}$=100 mW, $P_{idle,DC}$≈0, F=10, K=1.38·10^{-23} J/K, T=300 K, η_{Tx}=0.25. The left part of (3.36) is plotted as a linear function and the right part of (3.36) is plotted as an exponential equation of R. The cross points of them are therefore the solution of (3.36). Furthermore, if we take the antenna gain as a variable too, a series of solutions can be found, as shown in Fig. 3.6.

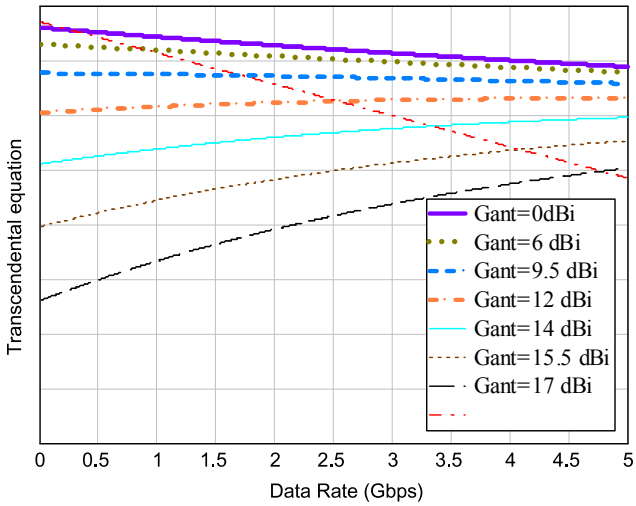


Fig. 3.6 The optimum data rate as a function of antenna gain for the wireless wire system.

[14] Relevant data are obtained from calculation, simulation and measurement results of the wireless wire system in Chapter 4 to 6.

In order to show the solutions of the transcendental equation clearly, Fig. 3.7 is plotted and the relation between the optimum data rate and antenna gain are shown.

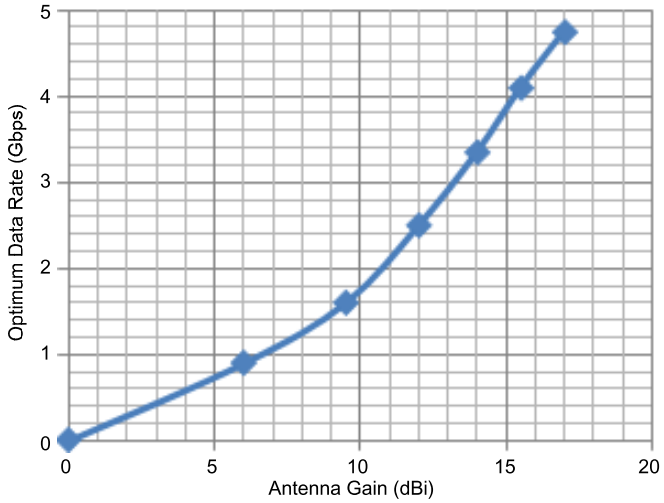


Fig. 3.7 Peak data rate vs. antenna gain.

It can be seen that in order to achieve a 1 to 1.5 Gbps data rate, 10-dBi antenna gain is required, which can be realized by, e.g. a 4-element phased array beamforming front-end. The average power consumption is plotted in Fig. 3.8 as a function of the peak data rate and the amount of total transmission data per day.

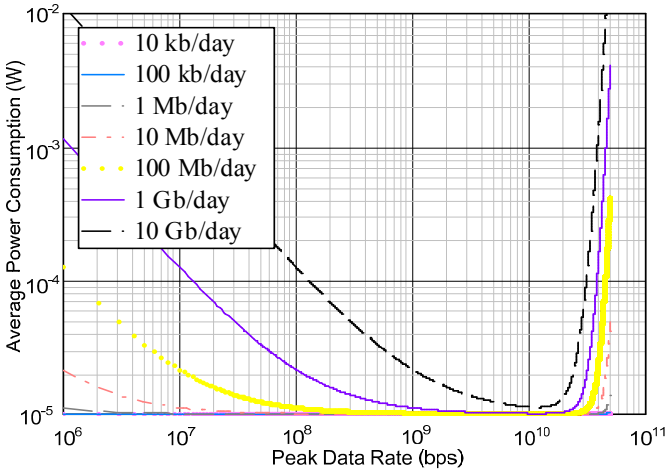


Fig. 3.8 Average power consumption as a function of data rate (G_{ant}=10 dBi).

It should be noticed that in this analysis, the interference is not considered yet. Very large interferers will be solved in the MAC-layer and network-layer optimization. For example, by adding a unique ID for each node, the blockers can be identified in the packet overhead and thus removed later on during the payload communication. Alternatively, a self-learning network can be used to enhance the nodes to adapt to the network topology and thereby to avoid miss and false alarms. More details about this will be given in

Section 3.4. In this part, only the low-level noise-like co-channel interference is taken into consideration. The direct result is that the required SNR will be increased according to Shannon's equation.

Recalling (3.27), the channel capacity becomes

$$C = B \cdot \log_2(1 + SNR_{out}) = B \cdot \log_2(1 + \frac{P_{s,in}}{KTB \cdot F + P_{interference}}) \tag{3.37}$$

where $P_{s,in}$ is the signal power at the input of the Rx and $P_{interference}$ is the total power of the noise-like co-channel interferences. The influence of the interference level to the channel capacity is plotted in Fig. 3.9.

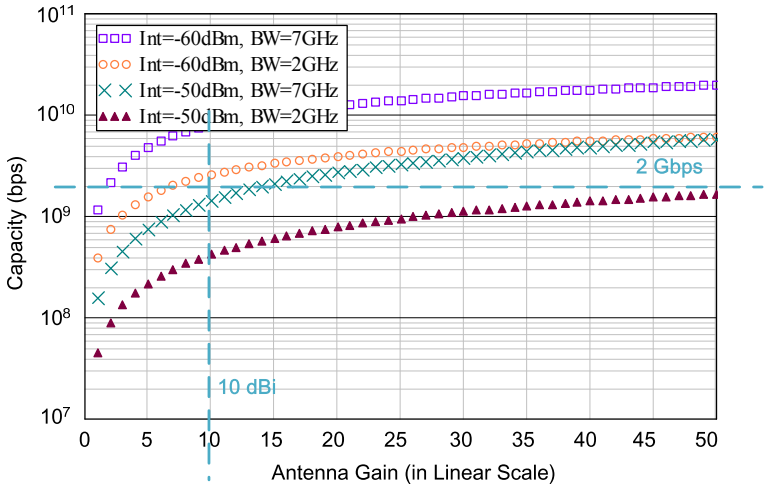


Fig. 3.9 Antenna gain vs. channel capacity under interference.

It can be seen from Fig. 3.9 that in a 60-GHz wireless system, in order to achieve a 1 Gbps data rate with about 10 dBi (10 in the linear scale as illustrated) total antenna gain (Tx plus Rx antenna gain), the interference level should not exceed -60 dBm. Otherwise the signal bandwidth must be very large, which can be done, for example, by using the spread-spectrum technology. For instance, in Fig. 3.9, when the interference level is -50 dBm, in order to achieve a 2-Gbps capacity, the signal bandwidth has to be extended to 7 GHz, as illustrated by the green crosses. On the other hand, however, large bandwidth would mean high noise floor (*KTB*), which will impair the Rx sensitivity. The possible solution is to keep a large signal bandwidth by sweeping the input signal slowly across a wide bandwidth while reducing the noise bandwidth by using a dynamic band-tracking filter[15]. However, such a filter is normally complicated and high power. When deployed at the first stage of an Rx, it may degrade the overall noise figure. Besides of that, the IF frequency in the Rx also becomes vague, which may cause problems in demodulation

[15] For more information, please refer to Section 4.2.2 and 4.3.3.

process. In this work, this problem is solved by using a passive mixer and low-pass filter at the output of the Rx. Details will be discussed in Chapter 4 and designs and implementations will be presented in Chapter 5.

The relation among the signal bandwidth, antenna gain and the output power of the Tx is plotted in Fig. 3.10.

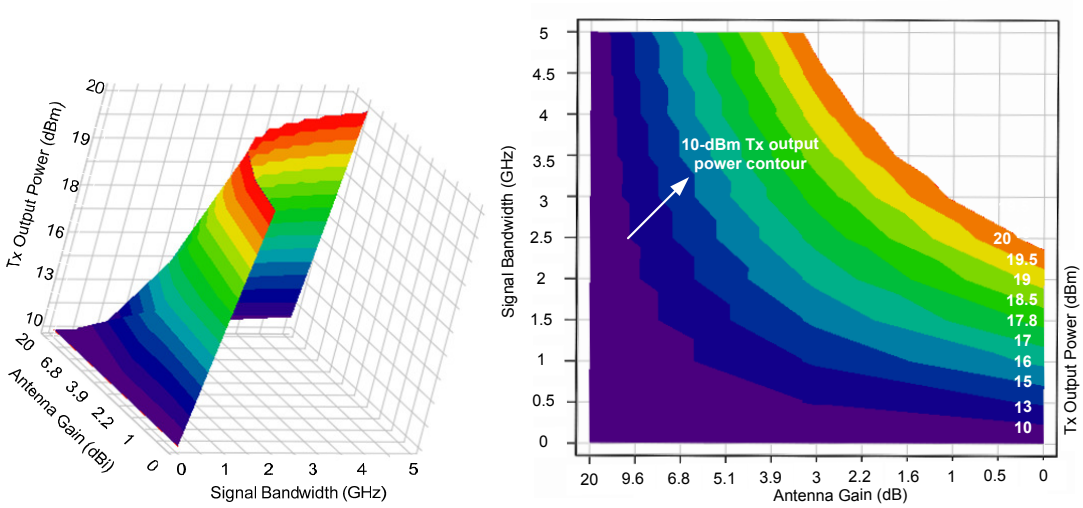


Fig. 3.10 (a) Tx Output as a function of the antenna gain and signal BW; (b) top view.

In order to observe the relation clearly, a top-down view of Fig. 3.10 (a) is plotted in Fig. 3.10 (b). With a 10 dBi antenna gain, the required EIRP of the Tx is about 10 dBm with 2-GHz signal bandwidth.

3.2.3 Optimum modulation scheme

If we substitute (3.13), (3.14) to (3.16), the minimum required output power of Tx can be modeled as[16]

$$P_{Tx,DC,\min} = \frac{KTB \cdot F \cdot (\frac{E_b}{N_0} \cdot \frac{R}{B}) \cdot PL}{\eta_{Tx} \cdot G_{Tx} \cdot G_{Rx}} = \frac{KT \cdot F \cdot (\frac{E_b}{N_0}) \cdot R \cdot PL}{\eta_{Tx} \cdot G_{Tx} \cdot G_{Rx}} \tag{3.38}$$

Given certain system and circuit related design constants in (3.38) like noise level, data rate, system bandwidth, PL and antenna gain, the Tx power consumption can be evaluated by a factor of $E_b/N_0/\eta_{Tx}$. It reveals the relation between the power efficiency and the modulation efficiency of a transceiver front end, and the smallest $E_b/N_0/\eta_{Tx}$ is preferable for low power systems. Referring to Fig. 3.3, in order to achieve a BER of 10^{-3}, the required E_b/N_0 of the BPSK and QPSK modulation schemes is about 7 dB while the

[16] In fact, it is equivalent to (3.23), but now we do not focus on frequency influence.

coherent and noncoherent OOK schemes require 10 dB and 11 dB E_b/N_0 respectively. However, the PSK modulation scheme has a stringent requirement of the system linearity while it is not very critical for the noncoherent OOK scheme, which can be simply demodulated by detecting the signal envelope by an energy detector. As a result, the transceiver efficiency of the latter can be higher. In a Tx front end, relaxed linearity requirement enables the use of non-linear and high efficiency PA. The theoretical efficiency of a PA (η_{PA}) can be calculated as [12]

$$\eta_{PA} = \frac{\alpha - \sin\alpha}{4(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2})} \tag{3.39}$$

where α is the conduction angle and it is in the range from 0 to 2π. The PA efficiency and normalized output power can be plotted in Fig. 3.11.

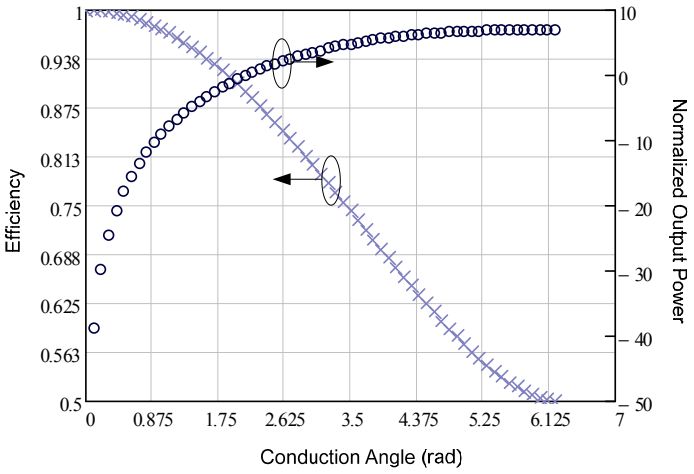


Fig. 3.11 The PA efficiency and output power as a function of the conduction angle.

It can be seen that the theoretical maximum efficiency of the linear class-A PA is 50% while it can reach 100% for the non-linear class-C PA. Theoretically, the electronics efficiency of the Tx can be improved almost by a factor 2 by using a non-linear PA (assuming the power of the Tx is dominated by the PA). On the Rx side, the circuit complexity and power consumption also become lower as well without the critical linearity requirement. For instance, the Rx can be simply a LNA cascaded with a signal squarer for OOK demodulation. For the PSK demodulation, however, the LNA, mixer and I/Q structure are normally compulsory. The overall effect if that the noncoherent OOK modulation scheme becomes more power-efficient. Besides the simple circuit structure, by using the OOK modulation, the requirement of high speed analog-to-digital convertors (ADC) of the Rx can be removed. The demodulated envelope signal can be digitalized by a comparator and sent to the baseband module directly. Consequently, non-coherent OOK modulation leads to lower circuit complexity and therefore be frequently used in low power systems.

3.2.4 Antenna gain

As discussed in Section 3.2.1, if keeping the antenna area non-scaled with frequency, the extra antenna gain at higher frequency is capable to compensate the "frequency-induced losses", e.g. the pathloss and extra noise. The essential idea of this argument is to find out the boundary value of the required antenna gain instead of offering the solution. In other words, by knowing the value, the needed extra gain can be realized by many other approaches. The most well-known method is using a multi-element antenna array to increase the EIRP in a certain direction which is the basis of the beamforming technique. The theoretical additional antenna gain can be interpreted as [13]

$$G_{tot} = \eta_{rad} \cdot \eta_{ap} \cdot \frac{4\pi \cdot n^2 A}{\lambda^2} = n^2 \tag{3.40}$$

where n is the number of antenna elements. It can be seen that by using the multi-element phased array structure, n^2 additional gain can be obtained (compared with the omni-directional antenna with unity antenna gain) and the antenna area can be kept on scaling down when increasing the frequency. For example, comparing the antenna gain between a 2.4-GHz antenna and a 60-GHz antenna with the same area, the extra antenna gain (G_{ext}) can be calculated as

$$G_{ext} = \frac{\eta_{rad} \cdot \eta_{ap} \cdot \frac{4\pi \cdot A \cdot 60^2}{c^2}}{\eta_{rad} \cdot \eta_{ap} \cdot \frac{4\pi \cdot A \cdot 2.4^2}{c^2}} = 625 \tag{3.41}$$

Theoretically speaking, a 625 times higher antenna gain is obtained. By using an 8-elemnt beamsteering antenna array, an extra 64 times higher array gain can be obtained, which allows shrinking the area of each antenna by a factor of 64 without impairing the total extra antenna gain. It can be illustrated as

$$G_{ext} = \frac{8^2 \cdot \eta_{rad} \cdot \eta_{ap} \cdot \frac{4\pi \cdot \frac{A}{8^2} \cdot 60^2}{c^2}}{\eta_{rad} \cdot \eta_{ap} \cdot \frac{4\pi \cdot A \cdot 2.4^2}{c^2}} = 625 \tag{3.42}$$

In fact, the advantage of the multi-antenna beamsteering technique does not only lie on the array gain, but also in the spatial filtering, i.e. reduction of the co-channel interference in the undesired directions. The antenna beamwidth (θ_b) of an n-path array can be estimated as [14]

$$\theta_b = \frac{\lambda}{(n-1) \cdot d} \tag{3.43}$$

where d is the distance between each antenna element. If d is fixed to half of the wavelength, the beamwidth can be linearly scaled down when increasing the antenna

number. With a narrow beamwidth, the nodes will receive much less co-channel interference in the undesired directions. Referring to Fig. 3.9, if keeping the SNR constant, the capacity will be improved under lower interference level. Besides, the output SNR of the Rx front end can be improved by *n*, which is very important for the link budget model. More details about the multi-element antenna array beamsteering technique will be discussed in Chapter 4 and 6 at the system and circuit levels.

3.3 PHY-MAC co-design towards minimum average power

According to previous discussions, some of the fundamental PHY parameters of the wireless system are identified for the optimum power performance, e.g. the choices of the communication frequency, data rate, modulation scheme and the beamsteering technique. However, the power consumed to control and regulate the communication is not yet considered. Possibilities for a low-power MAC protocol should also be investigated in order to achieve the truly overall optimum power performance.

3.3.1 Generic solutions for power management

A very effective method to reduce the average power consumption is to make the radio incontinuously operate and extend the idle or inactive period to the maximum extent. As long as the missed-alarms are avoided, the duty-cycle should be as low as possible. Two methods are widely used: duty cycling and adding an extra WuRx to monitor the communication, as illustrated in Fig. 3.12 (a) and (b), respectively.

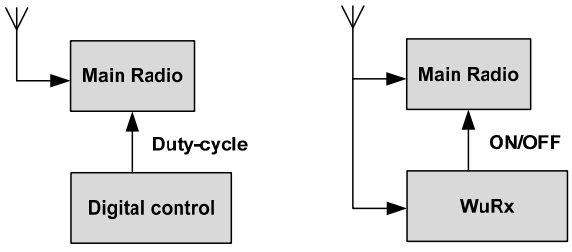


Fig. 3.12 Architectures of (a) duty-cycled radio and (b) the wake-up radio.

With the duty-cycled power management scheme, a power-efficient MAC protocol is crucial to monitor the channel and control the communication. Since it is unknown when the RF signal is sent, the Rx has to activate frequently during the idle period (or sleep period) and listen to the channel in order to switch the Rx on in time. The operation diagram is illustrated in Fig. 3.13 (a). If the event occurrence frequency is very low, a great amount of power will be used for listening activity, which can be even comparable with the packet communication power. The situation is even worse in the burst-mode

communication system where the data rate is high, and the main radio "on" time is very short while the sleep time is very long.

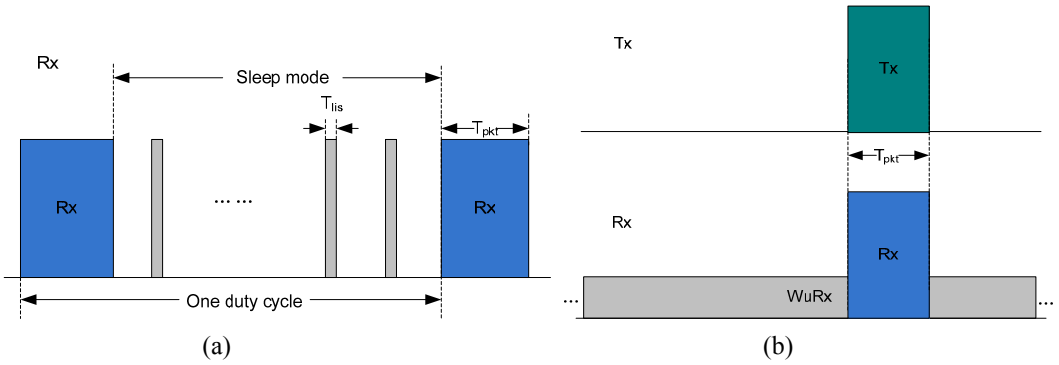


Fig. 3.13 Operation activity diagrams of (a) the duty-cycled Rx and (b) the wake-up radio.

The power management diagram of the wake-up radio is illustrated in Fig. 3.13 (b). The main radio Tx is event-triggered, i.e. it will activate only if there is data needed to be sent, and the Rx is switched on by the WuRx when the communication is identified. The WuRx is normally always-on in order to monitor the channel. Intuitively speaking, a significant amount of the power is consumed by the WuRx if it includes any active devices, e.g. amplifiers or active filters. The power cannot be neglectable if the WuRx works in a high frequency band. For example, according to the previous conclusion, the radio should work at the mmW frequency in order to optimize the E_{bit}. As a result, its power consumption can easily reach a level of several milliwatts. This power level is hardly compatible with a battery-based or energy scavenging based power supply system.

In order to minimize the overall average power consumption, the asynchronous duty-cycled wake-up power management scheme is proposed in the following.

3.3.2 Asynchronous duty-cycled wake-up scheme

The duty-cycled wake-up scheme is illustrated in Fig. 3.14. The essential idea is to make the WuRx work in a low duty-cycle mode to reduce its average power consumption. The main radio Tx is still event-triggered. Before sending the payload data, the Tx will first send a wake-up beacon to the WuRx. When the WuRx identifies and captures the wake-up information, it will switch the main radio Rx to the active mode. As long as the wake-up beacon is longer than at least one duty-cycle of the WuRx, the wake-up instruction can be recognized correctly. With using this scheme, the effective active time of the Tx and Rx achieves the minimum value, i.e. equal to the payload data packet length and the average power consumption of the WuRx is reduced significantly too. Besides, since the scheme is asynchronous without the accuracy requirement of the duty-cycle control module, the power of the always-on digital circuits can be neglectable. The effectiveness of this scheme is proved in Section 3.3.3.

In Fig. 3.14, $T_{b,Tx}$ is the time duration of the wake-up beacon from the Tx, T_{pkt} is the payload transmission time, T_{wu} is the "on" time of WuRx, T_{DC} is the total time duration of a duty cycle and T_{sleep} is the idle time when the WuRx is switched off. In order to avoid the missed alarms the following inequality should be satisfied: $T_{b,Tx} \geq T_{DC}+2 \cdot T_{wu}$ where $T_{DC}=T_{wu}+T_{sleep}$. The WuRx activates periodically with the time duration T_{wu} every T_{DC}. Since the system does not need to be synchronized, $T_{b,Tx}$ should be long enough to meet at least two T_{wu}'s. Intuitively speaking, long wake-up beacons will increase the power consumption of the Tx. However, it can also reduce the missed-alarms and thus the power of the wake-up beacon. As a result, an optimum T_{DC} exists, which leads to the minimum average power consumption.

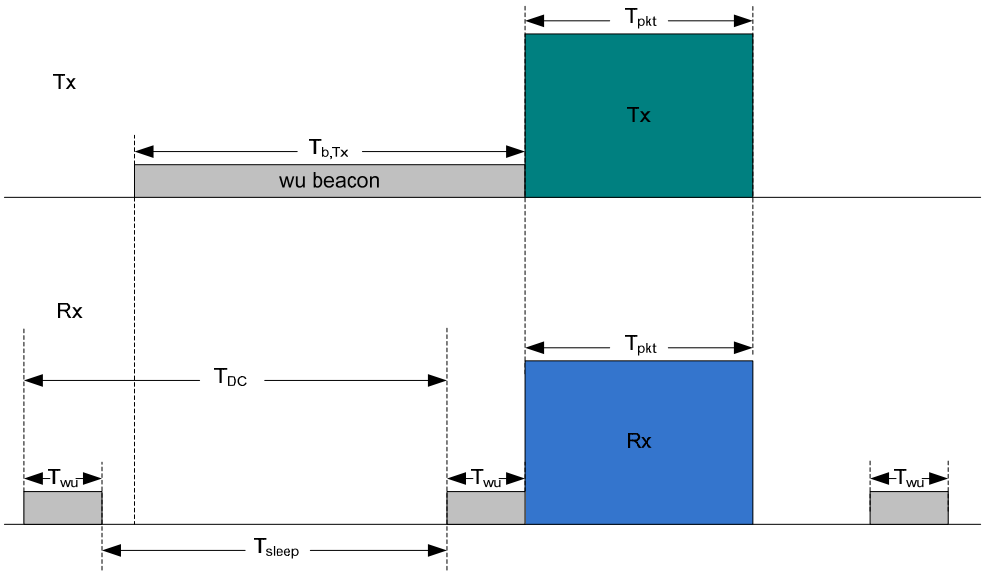


Fig. 3.14 The asynchronous duty-cycled wake-up scheme.

The total average power of the Tx and the WuRx pair is calculated as

$$P_{ave,DC} = P_{WuRx,DC} \cdot \frac{T_{wu}}{T_{DC}} + P_{b,Tx,DC} \cdot \frac{T_{b,Tx} \cdot n_{event}}{T_{tot}} \tag{3.44}$$

where $P_{WuRx,DC}$ is the power consumption of WuRx, $P_{b,Tx,DC}$ is the power consumption of the Tx to send the wake-up beacon (it is identical to $P_{Tx,DC}$ when re-using the Tx to send the beacon), n_{event} is the number of events of the Tx and T_{tot} is the total observing time. Differentiating (3.44) and letting $dP_{ave,DC}/dT_{DC}$ equal 0, the optimum T_{DC} becomes

$$T_{DC} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot T_{tot}}{P_{Tx,DC} \cdot n_{event}}} \tag{3.45}$$

Replacing $T_{b,Tx}$ in (3.44) by $T_{DC}+2 \cdot T_{wu}$, the minimum power consumption of the Tx and the WuRx $P_{ave,DC,min}$ is then calculated as

$$P_{ave,DC,min} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot P_{Tx,DC} \cdot n_{event}}{T_{tot}}} + P_{Tx,DC} \cdot \frac{(T_{DC} + 2 \cdot T_{wu}) \cdot n_{event}}{T_{tot}} \tag{3.46}$$

From (3.46), it can be seen that in order to achieve the minimum average power consumption, variables like $P_{WuRx,DC}$, $P_{Tx,DC}$ and T_{wu} should be minimized simultaneously.

Similar to the main radio, the minimum value of $P_{Tx,DC,min}$ can be determined via the system link budget analysis as

$$P_{Tx,DC,min} = \frac{S_{WuRx} \cdot PL}{G_{ant} \cdot \eta_{Tx}} \tag{3.47}$$

where S_{WuRx} is the WuRx sensitivity. $P_{WuRx,DC}$ and T_{wu} can be optimized in the circuit-level designs and optimizations. n_{event}/T_{tot} is primarily determined by the applications, e.g. low or high traffic scenario. Consequently, $P_{ave,DC,min}$ becomes only a function of S_{WuRx}.

For example, if a 60-GHz WuRx consumes 20 μs to identify the wake-up beacon, and assuming $P_{WuRx,DC}$ and $P_{Tx,DC}$ to be 5 mW and 100 mW (same as the Tx power) respectively[17], the optimum T_{DC} becomes 13 ms with 1000 times/day event frequency according to (3.45). Substituting these values into (3.46), the minimum $P_{ave,DC}$ of the Tx (sending wake-up beacon only) and WuRx pair is about 8 μW [15].

3.3.3 Comparison of power management schemes

Some terms should be pre-defined: $P_{ave,DC,duty\text{-}cycled}$, $P_{ave,DC,wake\text{-}up}$ and $P_{ave,DC,cycled\text{-}wake}$ are the average power of the duty-cycled radio, wake-up radio and our asynchronous duty-cycled wake-up radio, respectively. $P_{peak,DC}$ is the peak power consumed by the main transceiver, which is assumed to be identical for the three types of radios. T_{data} is the total data transmission time in the total observing time T_{tot}. T_{other} is the summation of synchronization, listening and guarding time for the duty-cycled radio. m is the number of cycles within T_{tot} and $m=T_{tot}/T_{DC}$. The following equations can be written as

$$P_{ave,DC,duty-cycled} = P_{peak,DC} \cdot \frac{T_{data}}{T_{tot}} + P_{peak,DC} \cdot \frac{T_{other} \cdot m}{T_{tot}} \tag{3.48}$$

$$P_{ave,DC,wake-up} = P_{peak,DC} \cdot \frac{T_{data}}{T_{tot}} + P_{WuRx,DC} \tag{3.49}$$

$$P_{ave,DC,cycled-wake} = P_{peak,DC} \cdot \frac{T_{data}}{T_{tot}} + P_{WuRx,DC} \cdot \frac{T_{wu}}{T_{DC}} + P_{Tx,DC} \cdot \frac{n_{event} \cdot (2 \cdot T_{wu} + T_{DC})}{T_{tot}} \tag{3.50}$$

[17] These values are obtained from our analytical models and simulation and measurement results, which will be presented in Chapter 4 to 6.

Since the wake-up beacon is sent by the Tx, $P_{Tx,DC}$ can be conservatively assumed to be equal to $P_{peak,DC}/2$ if the Tx and Rx power is identical (or at least on the same order). However, in a more pessimistic estimation, we can further assume $P_{Tx,DC}$ is very close to $P_{peak,DC}$, i.e. the total power is dominated by the Tx. Furthermore, for mathematical deviation convenience, the power consumption between the main radio and WuRx can be modeled as $P_{peak,DC}=k\cdot P_{WuRx,DC}$. As a result, the comparison among (3.48), (3.49) and (3.50) becomes equivalent to the comparison among the following equations:

$$P_{ave,DC,duty\text{-}cycled,\Delta} = P_{WuRx,DC}\cdot k\cdot\frac{T_{other}}{T_{DC}} \tag{3.51}$$

$$P_{ave,DC,wake\text{-}up,\Delta} = P_{WuRx,DC} \tag{3.52}$$

$$P_{ave,DC,cycled\text{-}wake,\Delta} = P_{WuRx,DC}\cdot\left(\frac{T_{wu}}{T_{DC}}+\frac{k\cdot n_{event}}{m}+\frac{2\cdot k\cdot n_{event}\cdot T_{wu}}{m\cdot T_{DC}}\right) \tag{3.53}$$

It can be seen that $P_{ave,DC,duty\text{-}cycled,\Delta}$ is smaller than $P_{ave,DC,wake\text{-}up,\Delta}$ only if T_{other} is shorter than T_{DC}/k. Though k varies with the system frequency, sensitivity and technology, it can be estimated from existing designs that k is normally in a level up to 100 when the system frequency is in GHz level [16], [17]. Consequently, T_{other} should be shorter than 1% of T_{DC}, which has already been achieved by some energy-efficient MAC protocols like S-MAC or RMAC [18]. Therefore, the power level of (3.51) and (3.52) are comparable. In other words, $P_{ave,DC,cycled\text{-}wake,1}$ would be more power efficient only if

$$\frac{T_{wu}}{T_{DC}}+\frac{k\cdot n_{event}}{m}+\frac{2\cdot k\cdot n_{event}\cdot T_{wu}}{m\cdot T_{DC}} << 1 \tag{3.54}$$

which can be further interpreted as $T_{wu}/T_{DC}<<1$ and $k\cdot n_{event}/2m<<1$ simultaneously.

In a low-traffic mmW WPAN scenario, a reasonable average event frequency, i.e. n_{event} is 1000 events per day. Recalling (3.45), it can be obtained that

$$\frac{T_{wu}}{T_{DC}} = \sqrt{\frac{n_{event}\cdot k\cdot T_{wu}}{2\cdot T_{tot}}} \tag{3.55}$$

where k is 100 and T_{tot} is 24 hours, i.e. 86400 seconds. Taking T_{wu} of 20 μs as we calculated for the 60-GHz radio, T_{wu}/T_{DC} turns out to be in an order of 10^{-3}, which is much smaller than 1. Likewise, $k\cdot n_{event}/m$ can be estimated by $(k\cdot n_{event})/(T_{tot}/T_{DC})$, which is in an order of 10^{-2} and much smaller than 1. Consequently, it is shown that

$$P_{ave,DC,cycled-wake,\Delta} = P_{WuRx,DC}\cdot\left(\frac{T_{wu}}{T_{DC}}+\frac{k\cdot n_{event}}{m}+\frac{2\cdot k\cdot n_{event}\cdot T_{wu}}{m\cdot T_{DC}}\right) \le P_{WuRx,DC}\cdot 0.1 \tag{3.56}$$

which indicates that the average power consumption of the duty-cycled WuRx is about 10% of that of a low-duty-cycled and normal wake-up protocols.

3.3.4 Duty-cycled wake-up with CSMA/CA medium access

In the asynchronous wake-up scheme, it may happen that two or more Txs want to wake up one WuRx and use the same channel for communication. Likewise, it also possible that during a transmission process, other undesired neighbor Tx nodes send a strong interference. If the interference has a comparable power level, it will be very difficult to maintain the correct communication. In order to avoid the problems, the carrier sense multiple access with collision avoidance (CSMA/CA) method should be used. CSMA is a probabilistic Media Access Control (MAC) protocol in which a node verifies the absence of other traffic before transmitting on a shared transmission medium. Furthermore, with the collision avoidance (CA) mechanism, if the channel is sensed busy before transmission then the transmission is deferred for a "random" interval. This reduces the probability of collisions on the channel and the nodes that over-hear each other will not transmit simultaneously. In order to detect the effectiveness of the connectivity in some critical applications, the acknowledgement (ACK) based feedback loop can be added, as shown in Fig. 3.15. In this case, the Tx of the destination node will send a very short ACK to the source node. A very short time slot should be reserved for the ACK transmitting and receiving [19].

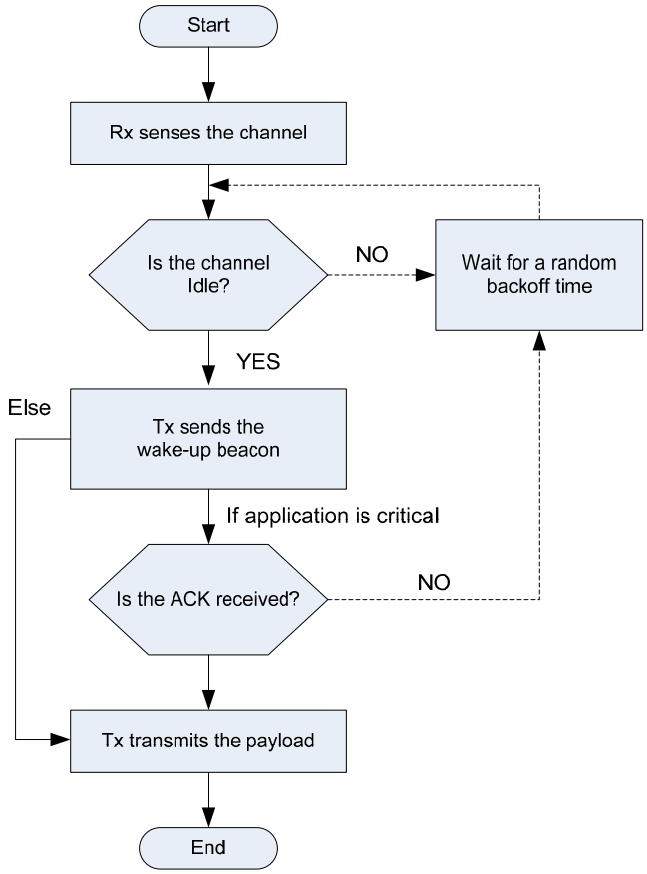


Fig. 3.15 CSMA/CA based asynchronous wake-up flow.

Since the channel sensing frequency is identical to the event occurrence frequency in the source node, the average power consumed for the sensing activity will become very low in the low-traffic application scenario. For example, before transmitting the effective data packet, the Rx of the source node consumes 100 mW power to sense the channel for 1 ms time duration assuming the network connectivity varies relatively slowly comparing to the communication speed. With an event frequency of 1000 events/day, the additional power dissipated for the channel sensing activity is only about 1 μW. The updated power management model is shown in Fig. 3.16. The power consumed by the ACK sending and receiving is ultra-low because the ACK is normally very short, so the average ACK consumed power can be neglected compared to the payload communication power.

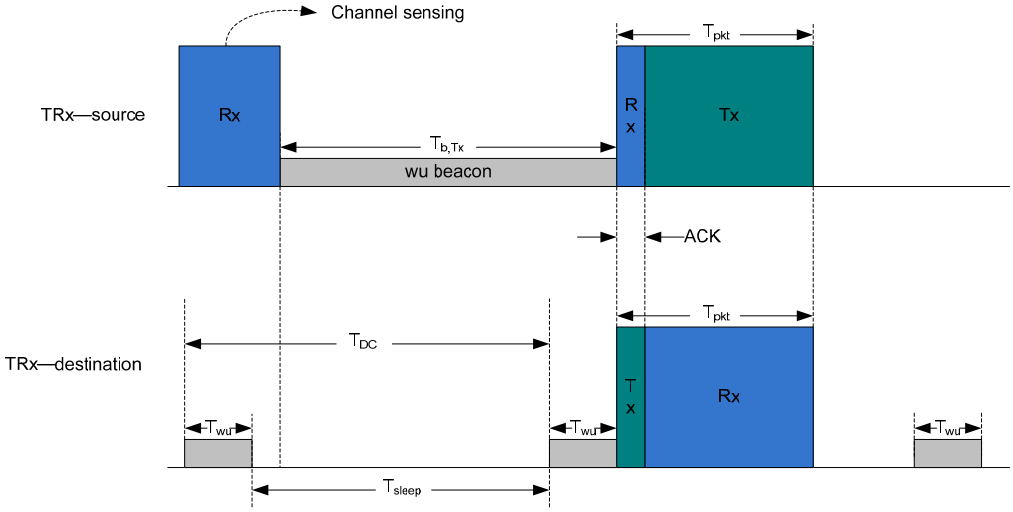


Fig. 3.16 Asynchronous duty-cycled wake-up with CSMA with ACK.

Referring to the 60-GHz WuRx example discussed in Section 3.3.2, with CSMA/CA included, the overall average power consumption of the WuRx, channel sensing and wake-up beacon generation is about 9 μW.

3.4 MAC-Network co-design: directional communication in the multi-hop and ad hoc networks

In previous sections, the PHY-layer parameters and potential MAC approach are discussed towards the low power performance. However, the optimization only targets at one communication pair. In this section, some basic network-related issues are discussed.

3.4.1 The ad hoc network: impact of interference on capacity

In a wireless ad hoc network, each node participates in routing by forwarding data for

other nodes, and the determination of which nodes forward data is made dynamically based on the network connectivity [20], as shown in Fig. 3.17. By re-using the spectrum resource, the system capacity is improved and thus it is widely used in modern wireless networks. However, the network capacity of this type of network configuration is sensitive to variations of the network size, nodes density and traffic per node. As a result, co-channel interferences may impair the capacity of communication nodes and degrade the performance of the entire network [21]. The influence of co-channel interferences on network capacity of a multi-hop and ad hoc network is analyzed as follows[18].

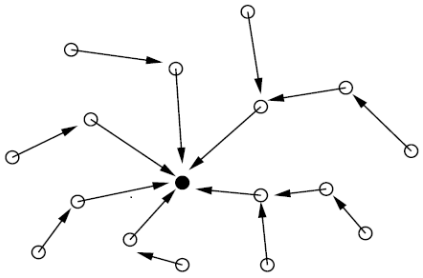


Fig. 3.17 A simple ad hoc network with multi-hop routing.

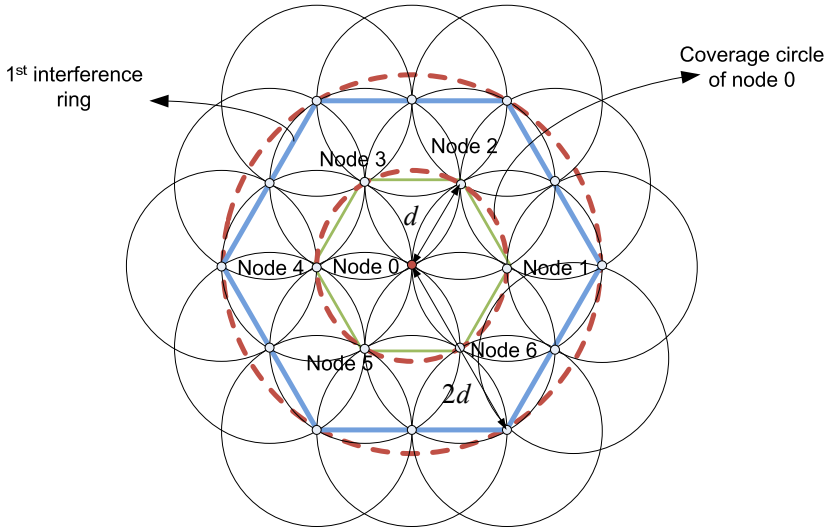


Fig. 3.18 Interferers prediction with CSMA/CA in a multi-hop network [22].

As discussed in the last section, by using the CSMA/CA medium access method, when one node is transmitting, all other nodes within its coverage area will not be in the transmit mode and therefore have no contribution to co-channel interference power. For

[18] More generic analyses can be found in [22]. In this section, the network model is discussed only in our particular application scenario.

instance, in Fig. 3.18, node 0 has a coverage circle with radius d, where d is the distance by which the Tx output power attenuates to a level just below the Rx sensitivity. When node 0 is transmitting, all other nodes within the coverage circle (i.e. node 1 to 6) do not transmit simultaneously, and interference can only be generated by nodes outside of the coverage circle [22]. By using the multi-hop routing method, when node 0 wants to communicate with any node outside its coverage circle, it has to wake up one of the nodes from node 1 to 6 and use it as the relay node[19]. Besides, homogeneity is assumed for this ad hoc network. In such a situation, all nodes emit the same output power in transmit mode and have the same transmission probability. It should be noticed that this model assumes a uniform distribution of nodes over a two-dimensional area with limited size. The actual situation is that inside the coverage area of the center node, there may also be many other nodes. However, it has been proved in [22] that this hexagonal lattice distribution approximation enables the estimation of network carrier-to-interference ratio (*CIR*) without requiring extra information of mobility and movement pattern of nodes in the network.

The upper bound of the interference mean power (I_{max}) in an ad-hoc network with CSMA/CA medium access method is

$$I_{max} = 12 \cdot q \cdot P_{out} \cdot (\frac{\lambda}{4\pi \cdot 2 \cdot d})^2 \tag{3.57}$$

where q is the probability of transmission per node, P_{out} is Tx output power (in fact, it is the EIRP), d is the radius of the hexagonal ring within the coverage area[20] and free-space LoS communication is assumed (pathloss index is taken as 2). Furthermore, in this calculation, only the first interference ring (the nearest one with a distance of $2d$ to the center node) is considered, because in our high frequency and personal communication scenario, high pathloss would attenuate the interference power if it is far from coverage area of the center node[21]. The expected wanted signal power from one of the nodes within the coverage area can be calculated as

$$P_{in,Rx} = P_{out} \cdot (\frac{\lambda}{4\pi \cdot d})^2 \tag{3.58}$$

The expected *CIR* of node 0 becomes

---

[19] In this model, all the relay nodes are placed in the edge of coverage circle while in reality these nodes would locate within the circle as well. Besides, we assume a node degree (number of nodes within its coverage area) of six in the model.

[20] With energy-efficient multi-hop routing method, only one hexagonal ring is assumed within the coverage area of a node. In other words, shortest distance should be kept for each hop.

[21] In fact, the antenna directivity also helps to alleviate the interference level significantly. However, this calculation only targets the worst case without taking in account the spatial filtering function.

$$CIR = \frac{P_{out} \cdot (\frac{\lambda}{4\pi \cdot d})^2}{12 \cdot q \cdot P_{out} \cdot (\frac{\lambda}{4\pi \cdot 2 \cdot d})^2} = \frac{1}{3 \cdot q} \tag{3.59}$$

It is shown in [22] that the transmission probability per node q (including its own traffic and the relay traffics) can be assumed to have a Poisson distribution as

$$q = 1 - e^{-\xi \cdot E[h]} \tag{3.60}$$

where ξ is the mean arrival rate of new packet per node (node's own traffic) and $E[h]$ is the expectation of number of hops. In our CSMA-CA based network, it can be that

$$E[h] = 0.53 \cdot N^{0.5} \tag{3.61}$$

where N is the total number of nodes in the network. Substituting (3.60) and (3.61) into (3.59), at the same output power, the CIR (the channel noise is not included) can be plotted as in Fig. 3.19.

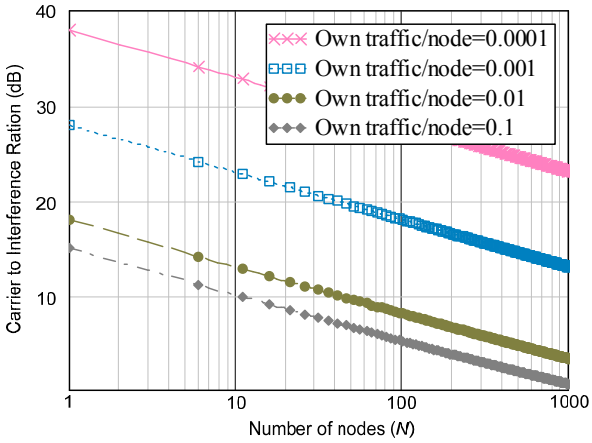


Fig. 3.19 Impact of the traffic on CIR in an ad hoc and multi-hop network.

It should be noticed that the hidden node problem[22] is not included in this model. However, it is already alleviated significantly by using the CSMA-CA method, because when the desired Rx node is occupied by other hidden nodes which cannot be sensed by the Tx before transmission, there would be no ACK back. In other cases, for instance, when suddenly encountering a high-power blocker during a transmission and the Rx is desensitized, re-transmissions of packets are required.

Referring to (3.37) and taking the 60-GHz front end as the example, the overall signal to noise and interference ratio at the output of the front-end ($SINR_{out}$) can be estimated as

[22] The hidden node refers to the node just outside the coverage area and therefore invisible during channel sensing process, but it will produce strong interference during transmission process.

$$SINR_{out} = \frac{S_{in}}{\frac{1}{G_{bf}} \cdot N_{in} \cdot F + I_{in}} = \frac{1}{\frac{1}{G_{bf}} \cdot \frac{1}{SNR_{in}} \cdot F + \frac{1}{CIR_{in}}} \tag{3.62}$$

where G_{bf} is the SNR improvement factor from beamforming, and S_{in}, N_{in} and I_{in} are the input-referred signal, noise and interference power respectively. With 2 GHz signal bandwidth and 4-times (that corresponds to a 4-element phased array Rx in our wireless wire system) better G_{bf}, the network capacity is illustrated in Fig. 3.20.

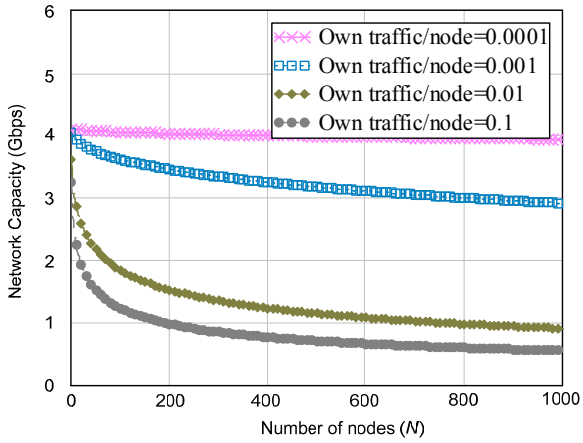


Fig. 3.20 The capacity of a 60-GHz system in the ad-hoc and multi-hop network.

Comparing to Fig. 3.9, it can be seen that in order to achieve Gbps high-speed communication in the ad hoc network, the maximum allowable interference level has a fixed boundary value. For example, if the own-traffic of each node is about 0.01 s^{-1}, the network has to be kept with a small to medium scale with at least less than 600 nodes. In the wireless wire low-traffic application scenario, the mean arrival packet rate (including its own and relay packets) is assumed as 1000 times/day, i.e. 0.0116 s^{-1}. The transmission probability per node can be obtained from (3.60) as shown in Fig. 3.21.

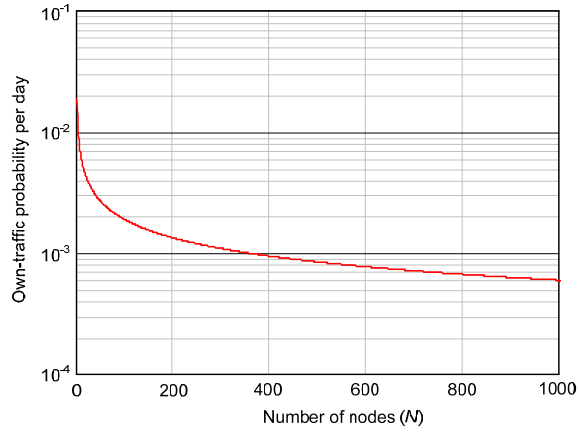


Fig. 3.21 Own-traffic probability of a node when total traffic is 0.0116.

If the total nodes number of the network is less than 600, the own-traffic probability of one node is about 0.008 according to the Poisson distribution. Referring to Fig. 3.20, the condition for the Gbps capacity is then guaranteed.

3.4.2 Multi-hop routing towards maximum network lifetime

In this section, the benefits of the multi-hop routing are further discussed. The simplified unit section of a homogeneous network model is shown in Fig. 3.22, i.e. with three nodes and all the nodes having equal chance to communication with each other with the same traffic generation behavior and the same priority. For example, node 1 communicates with 2, 2 with 3 and 1 with 3 and vice versa. Assuming all the Rxs are identical, i.e. has the same sensitivity and the Tx and Rx have the same antenna gain, the total power consumption of Fig. 3.22 (a) can be calculated as[23]

$$P_{tot,a,DC} = 2 \cdot \left(\frac{S \cdot PL_1}{G_{TRx} \cdot \eta_{Tx}} + \frac{S \cdot PL_2}{G_{TRx} \cdot \eta_{Tx}} + \frac{S \cdot PL_D}{G_{TRx} \cdot \eta_{Tx}}\right) \tag{3.63}$$

or

$$P_{tot,a,DC} = 2 \cdot \left(\frac{S \cdot (\frac{4\pi}{\lambda})^2 \cdot d_1^{\,2}}{G_{TRx} \cdot \eta_{Tx}} + \frac{S \cdot (\frac{4\pi}{\lambda})^2 \cdot d_2^{\,2}}{G_{TRx} \cdot \eta_{Tx}} + \frac{S \cdot (\frac{4\pi}{\lambda})^2 \cdot D^2}{G_{TRx} \cdot \eta_{Tx}}\right) \tag{3.64}$$

Similarly, the total power consumption of (b) is

$$P_{tot,b,DC} = 4 \cdot \left(\frac{S \cdot (\frac{4\pi}{\lambda})^2 \cdot d_1^{\,2}}{G_{TRx} \cdot \eta_{Tx}} + \frac{S \cdot (\frac{4\pi}{\lambda})^2 \cdot d_2^{\,2}}{G_{TRx} \cdot \eta_{Tx}}\right) \tag{3.65}$$

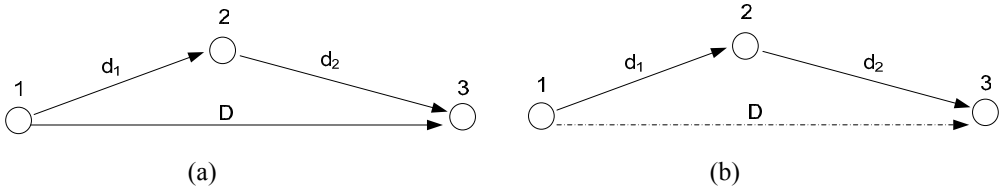


Fig. 3.22 (a) direct communication and (b) multi-hop routing.

Comparing (3.64) and (3.65), it can be seen that the $P_{tot,b}$ is lower than $P_{tot,a}$ only if

$$d_1^{\,2} + d_2^{\,2} < D^2 \tag{3.66}$$

i.e. D is sufficient long so that d_1 and d_2 form an obtuse angle.

---

[23] In this model, the Rx parameters (i.e. power, sensitivity and so on) are fixed and only the transmitting power is compared. The factor 2 in the multiplications refers to forward (e.g. 1 to 2) and backward (e.g. 2 to 1) transmissions with equal probability.

From (3.66), it can be concluded that the multi-hop routing is beneficial of saving total power of the entire network only when the long-distance communication happens frequently. This equation offers the boundary condition to choose the proper hopping algorithm. In fact, however, saving the total network power is not the primary benefit from multi-hop routing, but "saving the shortest wooden piece in the barrel[24]" is the most important function. For example, according the Frii's equation, the free-space pathloss of the signal is linear to the square of the communication distance. Increasing the communication distance from 1 meter to 10 meters, the pathloss is increased by 100 times or 20 dB, which means the Tx power consumption will be increased by a factor of 100 and its output power must be increased by 20 dB in order to compensate the extra loss. This is even not feasible in some cases. In a 60-GHz Tx, sending more than 20 dBm output power is quite a challenge task, which normally consumes DC power at a level of several watts. If the nodes have to get involved in the long-distance communication frequently, they will be extremely power-hungry and easily run out of power very fast. The performance of the entire network will be impaired due to those weak chains.

Two 2-D hopping arrays are illustrated in Fig. 3.23 (a) and (b). Assuming all the sources, destinations and relay nodes meet the condition described in (3.66), e.g. $d_1^2+d_2^2<D_1^2$ even if the node 2 is the only intermediate node between node 1 and node n.

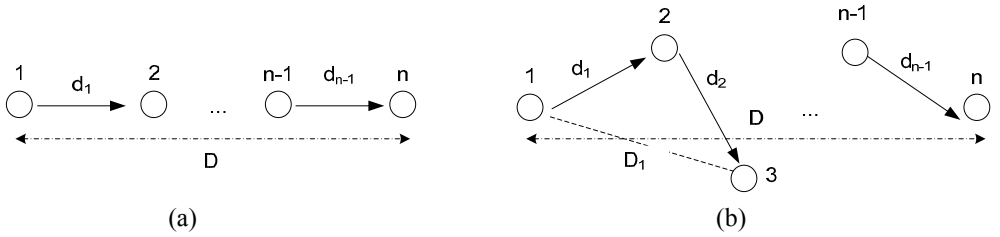


Fig. 3.23 (a) Linearly hopping array and (b) randomly hopping array.

If a data packet is transmitted from the node 1 to the node n with nodes 2 to n-1 as the relay nodes, the total power can be calculated as

$$P_{tot,DC} = \sum_{i=1}^{n-1} P_{Tx,DC,i} + (n-1)\cdot P_{Rx,DC} \tag{3.67}$$

where all the Rxs are assumed identical and the Txs can adjust their output power according to the distance. The power consumption of each Rx is $P_{DC,Rx}$. Re-write the Tx power consumption as

$$P_{tot,DC} = \sum_{i=1}^{n-1} (a\cdot d_i^{\,2}) + (n-1)\cdot P_{Rx,DC} \tag{3.68}$$

where $a=[S\cdot(4\pi/\lambda)^2]/(G_{TRx}\cdot\eta_{Tx})$. The simplest situation is illustrated in Fig. 3.23 (a). In

[24] Please refer to the wooden barrel theory in Section 1.3

this case, all relay nodes are aligned with the direct path from 1 to n, i.e.

$$\sum_{i=1}^{n-1} d_i = D \tag{3.69}$$

and the minimum power can be achieved when all nodes are evenly distributed, i.e.

$$d_1 = d_2 = \ldots = d_{n-1} \tag{3.70}$$

Comparing the power between the direct communication and the multi-hop routing, the following function can be obtained

$$f(n,\ D) = \frac{P_{Tx,DC}(D^2) + P_{Rx,DC}}{(n-1)\cdot P_{Tx,DC}((\frac{D}{n-1})^2) + (n-1)\cdot P_{Rx,DC}} \tag{3.71}$$

It can be seen that with a certain value of D, the multi-hop routing consumes less power only when $f(n,\ D)$ is larger than 1. Assuming the antenna gain is 1, the Rx power consumption is 100 mW and the electronics efficiency is 10%, this function of a 60-GHz linear hopping array can be plotted as

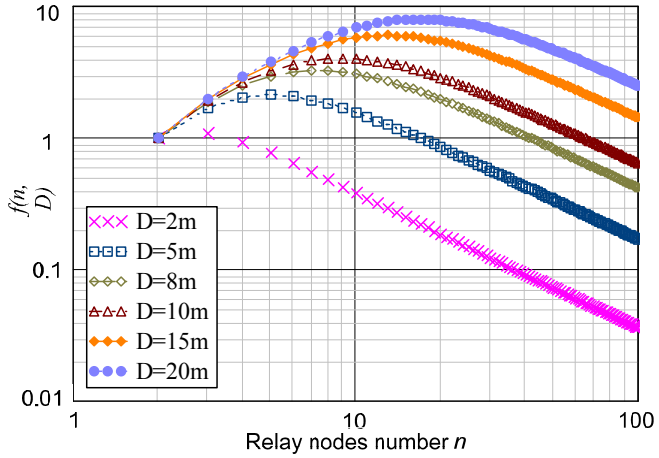


Fig. 3.24 Optimum relay nodes number in the linear array.

The data of the maximum distance and nodes number in Fig. 3.24 are listed in Table 3.1.

Table 3.1 The optimum relay nodes number

Maximum distance (m)	Nodes number	Hopping distance (m)
2	3	≈1
5	5	1
8	6	≈1
10	10	1
15	10	≈1
20	11	≈2

It can be seen that the optimum hopping distance is about 1 meter when the maximum distance is shorter than 20 meters. The situation described in Fig. 3.23 (b) is much more complicated. If the node keeps on choosing the nearest neighbor as the relay node, it is highly likely that the data is routed away and travels for a long way before it gets close to the final destination, i.e. node n. This situation should be avoided by the routing algorithm. The machine learning based network adaption would be a proper method. Even though some of the nodes in this application are moving with high mobility, if observing for a sufficiently long period, their moving patterns or routes are relatively fixed. For example, with a certain room configuration, the laptop is mainly used nearby the desk, sofa or the bed, and the most frequent communication peer is the mp4 player, which is mainly be used nearby the sofa. The paths connecting those points can be assigned with higher priority during the communication and thus fewer computations for routing or beamforming is needed and the connectivity is improved as well. As a result, the power model of Fig. 3.23 (b) can be simplified. Considering the topology described in Fig. 3.18 (b), only one situation is allowed in this case, i.e. data is routing from one inner hexagonal ring to one outer ring and the routings along a single ring are not allowed. In this way, most hops are kept to a linear straight line as long as there are relay nodes in between. Consequently, the total power optimization is quite close to the linear model illustrated in Fig. 3.18 (a).

3.5 Comparison by FOM

In the cross-layer optimization process, the energy efficiency is optimized by choosing the mmW frequency communication with a directional antenna, and the optimum data rate is evaluated accordingly. Besides, the average power consumption of the physical node is minimized by applying the asynchronous duty-cycled wake-up MAC scheme together with the multi-hop routing scheme. The packet overhead of the wireless system is further minimized by adapting the CSMA/CA method together with the asynchronous wake-up method. According to the FOM formula (3.8), the overall performance can be optimized. To conclude, the performance of different low-power wireless systems is compared by using the FOM at the end of this chapter (some of them are compared by $E_{ave,corr}$ stated in (3.7) due to incompletely reported data in the literature). Though some of the systems target at very different applications and have very different specifications, it would still be useful to evaluate them regarding the overall power efficiency and communication quality, which also indicates the overall design complexity. When the systems are used for the same applications, e.g. a Bluetooth WPAN and a 60-GHz WPAN, this comparison would become more fair and important. In Fig. 3.25, the FOMs of ZigBee, Bluetooth (IEEE 802.15.1), UWB (IEEE 802.15.3a) and this work are compared, and it shows that at the similar application scenario, this work achieves the best performance at the highest data rate. Qualitively speaking, the system power efficiency

and communication quality is improved while the system throughput is also increased, which verifies our theory.

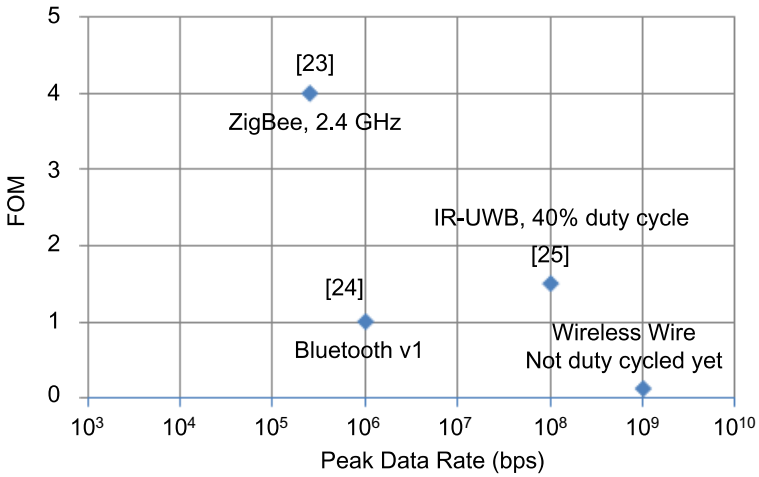


Fig. 3.25 Comparison of ULP systems by FOM or $E_{ave,corr}$.

References

[1] Xia Li *et al*., "An EFOM for Cross-Layer Optimization towards Low-Power and High-performance Wireless Networks," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, Paris, 2010.

[2] Behzad Razavi, "A 60-GHz CMOS Receiver Front-End," in *IEEE Journal of Solid-State Circuits* (*JSSC*), vol. 41, No. 1, pp. 17-22, 2006.

[3] Thomas H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits* (Second Edition), ISBN 7-121-01127-1, 2005.

[4] Chia-Hsin Wu *et al*., "A GPS/Galileo SoC with Adaptive In-Band Blocker Cancellation in 65nm CMOS," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[5] Shahram Abdollahi-Alibeik et al., "A 65nm Dual-Band 3-Stream 802.11n MIMO WLAN SoC," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[6] Domine Leenaerts *et al*., "A 65nm CMOS Inductorless Triple-Band-Group WiMedia UWB PHY," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2009.

[7] H. Hashemi *et al*., "A Fully Integrated 24-GHz 8-Path Phased-Array Receiver in Silicon," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2004.

[8] Kenichi Okada *et al*., "A 60GHz 16QAM/8PSK/QPSK/BPSK Direct-Conversion Transceiver for IEEE 802.15.3c," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[9] Sohrab Emami *et al*., "A 60GHz CMOS Phased-Array Transceiver Pair for Multi-Gb/s Wireless Communications," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[10] Mehdi Khanpour *et al*., "A Wideband W-Band Receiver Front-End in 65-nm CMOS," in IEEE *Journal of Solid-State Circuits* (*JSSC*), vol. 43, No. 8, pp. 1717-1730, 2008.

[11] Maja Vidojkovic et al., "A 2.4GHz ULP OOK Single-Chip Transceiver for Healthcare Applications," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011.

[12] Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*, Norwood MA: Artech House, 2006.

[13] Antenna gain, www.wikipedia.org, 2011.

[14] Ali M. Niknejad *et al*., *mm-Wave Silicon Technology 60 GHz and Beyond*, Springer, 2008.

[15] Xia Li et al., "A 60 GHz Ultra Low-Power Wake-up Radio," in *IEEE Radio and Wireless Symposium (RWS)*, 2011.

[16] Nathan M. Pletcher *et al*., "A 2GHz 52μW Wake-up Receiver with -72 dBm Sensitivity Using Uncertain-IF Architecture," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2008.

[17] Ajay Balankutty *et al*., "A 0.6-V Zero-IF/Low-IF Receiver With Integrated Fractional-N Synthesizer for 2.4-GHz ISM-Band Applications", in *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 45, NO. 3, 2010, pp. 538-553.

[18] Shu Du *et al*., "RMSC: A Routing Enhanced Duty-Cycle MAC Protocol for Wireless Sensor Networks," *in 26th IEEE International Conference on Computer*, pp.1478-1486, May 2007.

[19] Channel Sensing Multiple Access, www.wikipedia.org, 2011.

[20] ad hoc network, www.wikipedia.org, 2011.

[21] Olivier Dousse *et al*., "Impact of Interference on Connectivity in Ad Hoc Networks," in IEEE/ACM Transactions on Networking, vol. 13, no. 12, pp. 1724-1733, 2005.

[22] R. hekmat et al., "Interference in Wireless Multi-hop Ad-hoc Networks and its Effect on Network Capacity," in IEEE Journal of Wireless Networks, vol. 10, no. 4, 2004.

[23] ZigBee White Paper, in ZigBee™ Wireless Transceiver Engineering Options, Cirronet, Inc., 2005.

[24] Bluetooth, www.wikipedia.org, 2011.

[25] Marco Crepaldi *et al*., "An Ultra-Low-Power Interference-Robust IR-UWB Transceiver Chipset Using Self-Synchronizing OOK Modulation," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2010.

CHAPTER 4 SYSTEM DESIGN AND OPTIMIZATION

In the previous chapter, the power-efficient PHY-layer specifications, power management method and the network configuration are discussed. With the cooperation and coordination among different layers, the energy efficiency and the average power consumption of a wireless communication system can be optimized simultaneously. However, during the optimization process, the physical system is approximated like a "black box" with several inputs describing its features, and the information about how to realize such a system is not specified. In this chapter, the details about realization and optimization of the physical system itself towards maximum energy efficiency and minimum average power dissipation will be discussed thoroughly.

4.1 System architecture and power optimization

According to the proposed asynchronous duty-cycled wake-up scheme, the radio system can be divided into three hierarchical levels: the main radio for the payload communication, the duty-cycled wake-up receiver for the MAC control and the always-on digital module for the duty-cycle control of the WuRx. The complete system architecture and the average power are discussed in the following.

4.1.1 Wireless wire system structure

The wireless wire communication system is shown in Fig. 4.1. At the top of the architecture is the main radio. The antenna directivity and the required antenna gain are realized by using a phased array beamsteering front end. The Tx is event-triggered, as shown in the top left, and the Rx is controlled by the WuRx, as shown in the top right of the figure. When there is no task to be carried out, the Tx and Rx in the main radio are set in the sleep mode, i.e. with zero supply voltage and the power consumption is theoretically zero. When a data packet is ready to be transmitted, the Rx on the sender side will first sense the channel for a certain period. If the channel is identified to be available, the Tx will send a wake-up beacon to the WuRx on the desired Rx side. As long as the wake-up beacon is longer than one duty cycle of the WuRx, the WuRx will capture the wake-up instruction and turn the Rx on. In such a case, the average duty cycle of the main radio can be minimized. According to the analytical models discussed in Section 3.2.1, the multi-element beamforming main radio should be implemented at the mmW frequencies with Gbps data rate in order to achieve lower E_{bit}. The peak power consumption of such a front end is quite high. As a result, minimizing its active period will effectively reduce its average consumption. This is particularly efficient in our

low-traffic personal communication scenario.

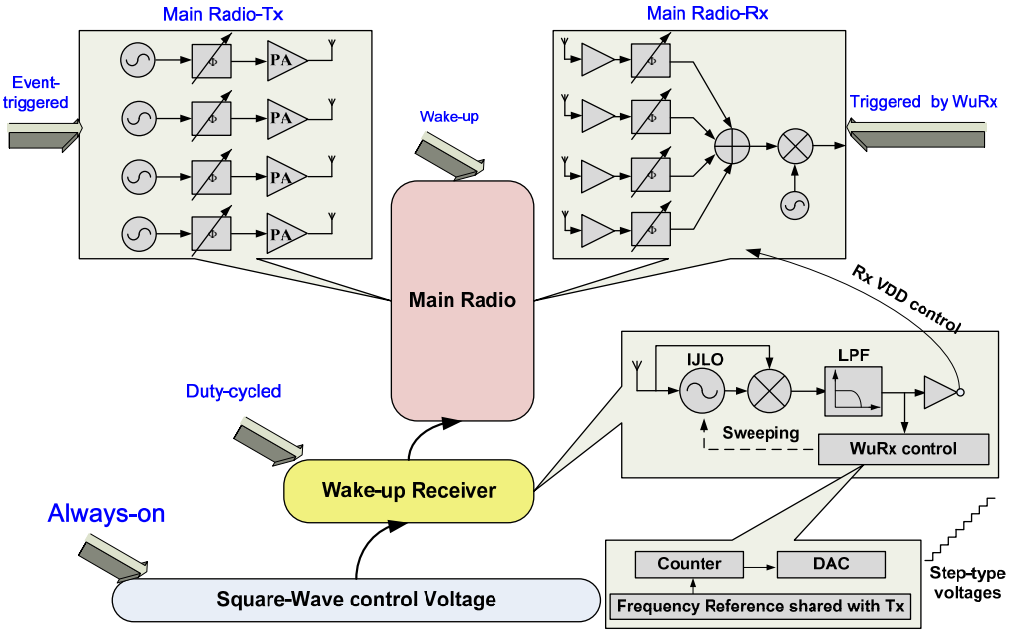


Fig. 4.1 The wireless wire front end.

The average power consumed by the MAC control module is strongly related to the duty cycle factor (DCF) of the WuRx. For example, if one duty cycle is too long and the DCF is too small, the required wake-up beacon must be sufficiently long as well to coincide with at least one active period of the WuRx. Therefore, it will result in more power consumption of the Tx. On the other hand, if the DCF is too large, the average power consumption of the WuRx will increase dramatically. As discussed in 3.3.2, an optimum value of the DCF exists as a function of the event frequency, the peak power of front end circuits and the settling time of the WuRx. The optimum DCF and the WuRx circuit structure will be investigated in this chapter, and the minimum average power consumed in the wake-up process of the WuRx and Tx peer will be identified. It should be noted that in the following discussions, a symmetrical link of the Tx and the Rx is assumed in an ad hoc scenario. If using a base station to send the wake-up beacon (e.g. a one-way link), only the WuRx power needs to be optimized.

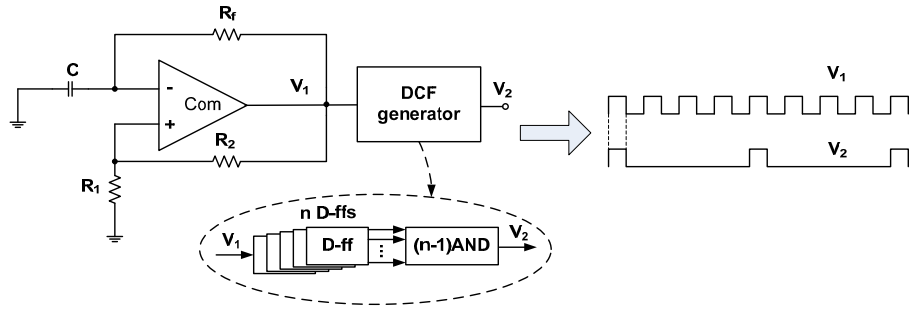


Fig. 4.2 Duty cycle control voltage generator.

The duty-cycle control voltage can be generated by the low-power digital circuits shown in Fig 4.2. For example, the clock signal can be generated from a comparator-based square-wave generator, and the DCF can be controlled by using a digital frequency divider. The output frequency of the oscillator can be calculated as

$$f_{out} = \frac{1}{2R_f \cdot C \cdot \ln(\frac{2R_1}{R_2} + 1)} \tag{4.1}$$

and the number of the D flip-flops are estimated as

$$n_{FF} = \log_2 \frac{1}{DCF} \tag{4.2}$$

Since the wake-up process is asynchronous without special accuracy requirements, the power consumption of such a circuit can be minimized by using the smallest digital devices in the library.

4.1.2 Average power consumption optimization

Since the digital duty-cycle control block is always on, its peak power is identical to the average power, which can be estimated by

$$P_{d,DC} = V_{DD}^2 \cdot C_L \cdot f \cdot n_T \tag{4.3}$$

where V_{DD} is the supply voltage, C_L is the load capacitor, f is the frequency and n_T is the total number of digital gates. For example, if the active time of the WuRx in one duty cycle is 20 μs (i.e. slightly larger than the worst-case settling time of the WuRx, which will be identified later in Section 4.3.2), the clock frequency becomes 50 kHz. The number of required D-FFs for the DCF generator is 10 if the DCF is 0.15% (i.e. 20 μs over 13 ms according to the calculation result from equation (4.2)). Assuming a 100 fF load capacitance, the total power consumption of the duty-cycle control voltages generator is about 0.5 μW. In fact, it is a bit over-estimated because for some gates of the DCF generator, the operation frequency is lower than the clock frequency. Consequently, the total power consumption of the duty-cycle control voltage generator is about 1 μW given the low-speed square-wave generator consumes another 0.5 μW power. Considering the leakage power and other potential power consumptions, we will leave 10-μW power budget for this block, which is the power floor of the entire wireless wire system.

In the case of a symmetrical link in the ad hoc scenario, the power consumption of the WuRx should be optimized together with the power consumption of the wake-up beacon generator, which is, in this case, realized by the main radio Tx. According to (3.44) in Section 3.3.2, the average power consumption of the WuRx and Tx peer is calculated as

$$P_{ave,WuRx,DC} = P_{WuRx,DC} \cdot \frac{T_{wu}}{T_{DC}} + P_{b,Tx,DC} \cdot \frac{T_{b,Tx} \cdot n_{event}}{T_{tot}} \tag{4.4}$$

where $P_{b,Tx,DC}$ is identical to $P_{Tx,DC}$ (i.e. the Tx is re-used to send the wake-up beacon).

Adding the channel sensing power consumed by the CSMA/CA process, the average power becomes

$$P_{ave,WuRx,DC} = P_{WuRx,DC} \cdot \frac{T_{wu}}{T_{DC}} + P_{Tx,DC} \cdot \frac{T_{Tx,DC} \cdot n_{event}}{T_{tot}} + P_{Rx,DC} \cdot \frac{T_{CSMA} \cdot n_{event}}{T_{tot}} \tag{4.5}$$

where T_{CSMA} is the time spent in channel sensing.

Since the channel sensing process is independent from the wake-up process, the optimum duty cycle of the WuRx is still identical to

$$T_{DC} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot T_{tot}}{P_{Tx,DC} \cdot n_{event}}} \tag{4.6}$$

by differentiating (4.5) with respect to T_{DC} and assuming $T_{b,Tx}$ as $T_{DC}+2 \cdot T_{wu}$.

The minimum average power consumption of the wake-up process including the CSMA/CA can be estimated as

$$P_{ave,WuRx,DC,\min} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot P_{Tx,DC} \cdot n_{event}}{T_{tot}}} + P_{Tx,DC} \cdot \frac{(T_{DC} + 2 \cdot T_{wu}) \cdot n_{event}}{T_{tot}} + P_{Rx,DC} \cdot \frac{T_{CSMA} \cdot n_{event}}{T_{tot}} \tag{4.7}$$

The average power consumption of the main radio Tx and Rx can be calculated as

$$P_{ave,main,DC,\min} = (P_{Rx,DC} + P_{Tx,DC}) \cdot \frac{n_{tot}}{R \cdot T_{tot}} \tag{4.8}$$

where n_{tot} is the total amount of payload data bits per T_{tot} and R is the peak data rate.

As a result, the overall average power consumption becomes

$$P_{ave,all,DC,min} = P_{d,DC} + P_{ave,WuRx,DC,min} + P_{ave,main,DC,min} \tag{4.9}$$

Taking the 60-GHz front end again as the example, if the WuRx consumes 20 μs to capture the wake-up beacon and start up the main radio, and $P_{WuRx,DC}$ and $P_{Tx,DC}$ as 5 mW and 100 mW (the circuits details will be presented in Chapter 5 and 6) respectively, the overall minimum average power consumption of the entire wireless wire system can be plotted as a function of the total data amount in Fig. 4.3.

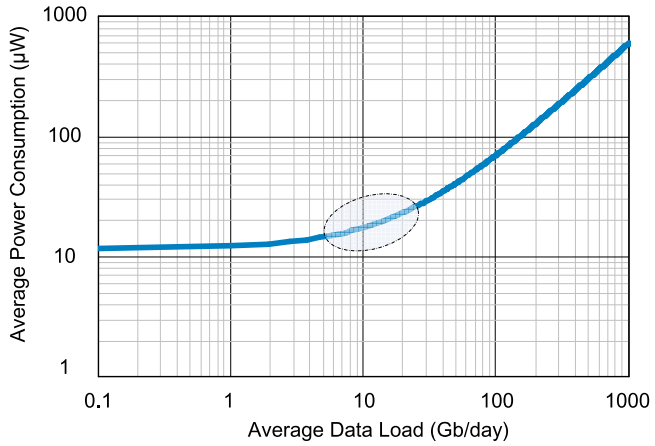


Fig. 4.3 The minimum average power consumption.

It can be seen that if the total data amount is about 10 Gb/day and event frequency is 1000 times/day, the overall average power consumption of the wireless wire system is about 20 μW.

4.1.3 Energy sources: harvesting, battery and combined

The suitability of the energy harvesting technologies is discussed as follows. In Table 4.1, the most well-known energy harvesting technologies are listed and the form factors are calculated in order to harvest 20 μW power [1-2]. Besides the transceiver power consumption, the baseband DSP, microprocessor and other relevant modules also consume significant amount of power, as shown in Fig. 4.4. Though the power consumptions of the baseband circuitry, processor and memory are beyond the scope of this thesis, their typical values are still shown in order to evaluate the feasibility of the energy harvesting technologies. In the future discussions, these power consumptions and related optimizations will not be considered anymore.

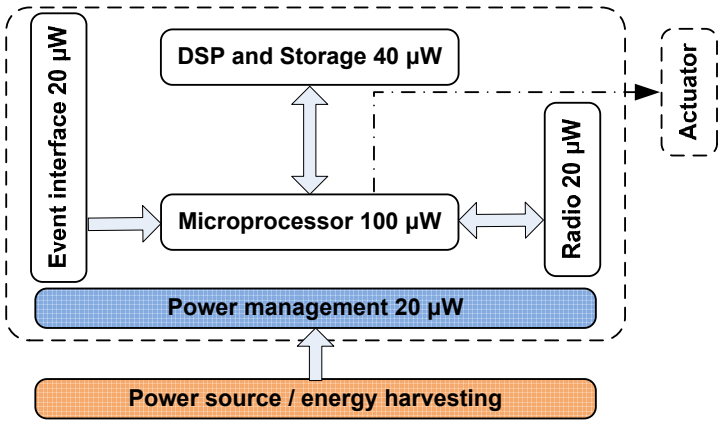


Fig. 4.4 The power budget of the entire node.

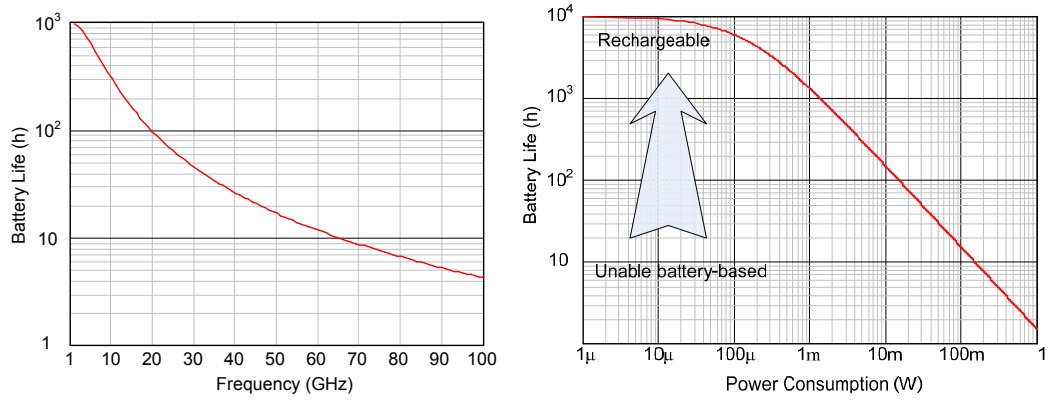
It can be seen that in the ideal case, at least 200 μW total power consumption is required from either the energy harvesting module or the battery. The characteristics of various well-known energy harvesting methods are listed in Table 4.1.

Table 4.1 Power densities of the energy harvesting methods [1] [3] [4]

Energy Harvesting Methods		**Power Density**	**Area/Volume for 200 μW**
Types	**Examples**		
Light	Solar cells (indoor)	<100 μW/cm^2	2 cm^2
Mechanical	Vibration	116 μW/cm^3	1.8 cm^3
	Piezoelectric	200 μW/cm^3	1 cm^3
	Electrostatic	100 μW/cm^3	2 cm^3
Thermal	Thermoelectric (10 °C gradient)	330 μW/cm^3	0.6 cm^3
RF	GSM	0.1 μW/cm^2	2000 cm^2
	Wi-Fi	0.01μW/cm^2	20000 cm^2
	Directional beam charging 2.4 GHz RF, 10 dBm output power for each path, 1 meter distance, 100 paths	20 μW/cm^2	1 cm^2
	Directional beam charging 60 GHz RF, 10 dBm output power for each path, 1 meter distance, 800 paths	20 μW/cm^2	1 cm^2
Human body	Combined method	20 μW/cm^2	10 cm^2

Considering the imperfections of the power supply like the discharge rate, relaxation effect, supply voltage variation and the diffusion rate of the active materials in the electrolyte, some margin has to be left to maintain the reliability of energy harvesting performance. For example, a bit larger area or volume of the energy harvesting devices should be reserved to collect more energy and compensate the power consumption and leakage of the scavenger itself [4]. According to Table 4.1, the thermal energy harvesting is the most efficient scavenger and it can easily achieve high power density with around 10 °C temperature gradient, which is quite common to be found in the living condition. Besides, the RF energy harvesting has received a lot of attention in the research community recently. It is beneficial with a high level of integration, i.e. it can be fully realized on-chip, but it is very hard to achieve a relatively high power density due to the high pathloss of the RF signal especially at high frequencies. Even with the beamforming technique where the equivalent isotropic radiation power density may be improved by a factor of 100, it is still quite low compared to other scavengers, as shown in table 4.1. In addition, the narrow pencil beam would complicate the connection, which will increase MAC power consumption. The high leakage level under modern CMOS technologies will also have strong influence and impair the harvesting efficiency.

Alternatively, a low-cost re-chargeable battery may be used as the energy source if its life time is sufficient to support the autonomous operation of a wireless system. For example, if a 1.5-V 1000 mAh rechargeable battery with 100 μA leakage current is used as the power supply, the battery life can be estimated by assuming certain circuit and system parameters. If the communication distance is 1 meter, Rx sensitivity is -60 dBm and the antenna gain is 1 dB (omni-directional), the battery life of an always-on wireless system can be calculated and plotted in Fig. 4.5 (a). It should be noted that the power consumption of the Rx is assumed identical to the Tx, which introduces some inaccuracy in the estimation results. However, they are normally in the same order, i.e. the Rx power also increases along with the operating frequency due to the higher noise factor and higher data rate. If we plot the battery life time as a function of the power consumption, as shown in Fig 4.5 (b), it can be seen that a months-long battery life can be achieved when the power consumption is reduced to less than 100 μW. With 20-μW power as obtained in our design (only the radio front-end is discussed at this moment), the battery life is about one year. Although it seems quite acceptable comparing to most of the mobile devices like, e.g. a cell phone that normally needs to be charged once per week, this battery life is still not good enough for autonomous operations. For example, if hundreds of wireless communication nodes are randomly deployed within a room, a battery life time of one year would mean that people may need to recharge one or two batteries per day if the batteries are dead randomly. Alternatively, people have to charge tens of batteries within one day when a group of batteries go empty simultaneously. That seems not an affordable task.



(a) (b)

Fig. 4.5 Battery life of (a) an always-on system and (b) as a function of the power.

Besides the above challenges, due to the present high leakage level of the low-cost batteries, a years-long life time is even not achievable. As a result, a combined power supply scheme can be used through the cooperation of the energy scavenger and the energy storage module. For example, during a long idle time in our system, the WuRx can be used for collecting the RF energy sent from, e.g. the base station (as discussed in

Appendix B). The high leakage power of the battery and part of the system power consumption can be compensated in this way, and thus a years-long autonomous operation can be achieved.

In the above discussion, the battery life is assumed to be determined only by the average power consumption of the load circuitry and the leakage current. However, it is not entirely true in fact. The instantaneous pulses of the required power of the load circuitry also has a significant influence, because when a large current is drawn out, it will cause a large voltage drop in the output resistance of the battery and eventually blackout the battery. In this work, the peak current can be up to 100 mA for the main radio (for Tx and Rx respectively), while the average current is just 20 μA, which means huge current spikes would be required when the system is active. As a result, a supercapacitor becomes quire necessary to smooth down those spikes and optimize the battery life further.

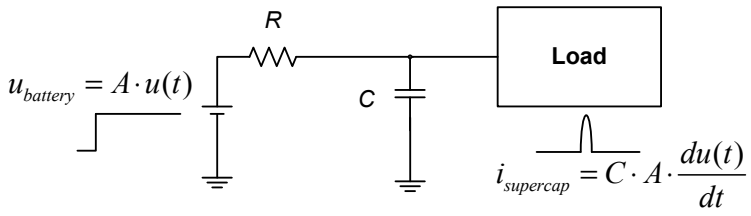


Fig. 4.6 Work principle of a supercapacitor.

It can be seen from Fig. 4.6 that by adding a big capacitor in parallel with the battery output resistance, the current spikes will be smoothed down due to its integrating effect. If the harvested energy is high enough, i.e. can meet the 20 μW requirement, the energy can be stored in the big capacitor.

According to [1], the size of the supercapacitor can be estimated by

$$C = 2 \cdot T_{\Delta} \cdot [P_{act} \cdot \alpha + P_{idle} \cdot (1 - \alpha)] \cdot \frac{1}{V_H^{\,2} - V_L^{\,2}} \tag{4.10}$$

where T_{Δ} is the time interval between two successive transmissions, P_{act} and P_{idle} are the peak powers in the active and idle mode respectively, α is the DCF, and V_H and V_L are maximum and minimum voltage of the battery respectively. Taking T_{Δ} as 86.4 s (1000 times/day with 10 s total payload transmission time per day, i.e. 10 Gbits and 1Gbps data rate), P_{act} as 200 mW (Tx power plus Rx power), P_{idle} is 10 μW, α is $1.15 \cdot 10^{-4}$, V_H is 1 V and V_L is 0.7 V (minimum allowable supply voltage), the capacitor is about 12 mF. Since the time intervals between two events are randomly distributed and sometimes could be very short, e.g. less than 10 s, a large margin has to be left, which increase the capacitor size further to the level 0.1 F.

4.2 Practical issues in the link budget model

In the link budget analysis, the RF system can be taken as a "black box" with several inputs to define its features. For example, on the Tx side, the implementation technology and the electronics efficiency are the most important factors and on the Rx side, the carrier frequency, communication data rate, the modulation scheme (E_b/N_0) and the BER are crucial. In other words, by knowing these inputs, a generic link model can be obtained without requiring more information about implementation details. Consequently, the system-level optimization would become more simple and straightforward.

As discussed in Sections 3.2.1 and 3.2.2, the wireless wire system is proposed in order to achieve the minimum E_{bit} and average power consumption. The most noticeable conclusion is that it is beneficial to use high-gain antennas for the directional communication in mmW frequency bands with Gbps-level communication data rate, as illustrated in Fig. 4.7.

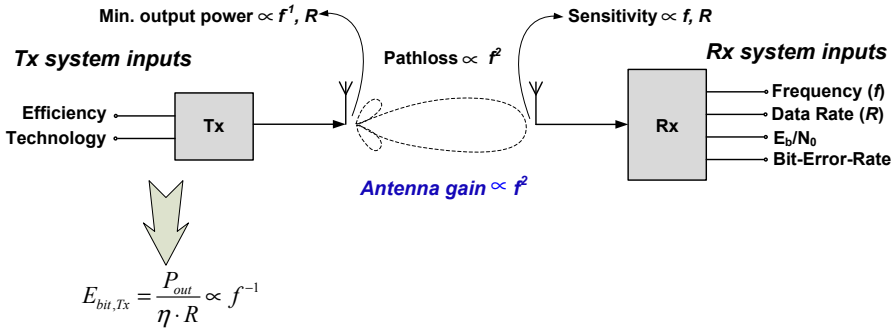


Fig. 4.7 The link model of the directional communication system.

In a generic link budget model, the sensitivity of the Rx (S) can be calculated by

$$S = KTB \cdot F \cdot \frac{E_b}{N_0} \cdot \frac{R}{B} \tag{4.11}$$

where KTB is the noise floor, F is the Rx noise factor, E_b/N_0 is the SNR per bit (determined by the modulation scheme), R is the data rate and B is the signal bandwidth. The minimum output power of the Tx is

$$P_{out,Tx,min} = S + PL - G_{TRx} \tag{4.12}$$

According to the discussion in Section 3.2.1, the essential idea of the wireless wire system is to use the additional antenna gain to compensate the high pathloss and noise at high frequency bands and thus achieve low-power communication. However, in reality, some challenges may be encountered during the implementation of the directional communication. For example, in Section 3.2.3, the phased array structure is proposed to

modulate the beam and increase the antenna directivity as well as the antenna gain. However, multiple RF paths (at least multiple phase shifters or time delays) must be used to control the beam pointing angle. The power consumption of the front end would increase in this case, which may reduce the benefit of the additional antenna gain. The practical related issues are discussed in the following.

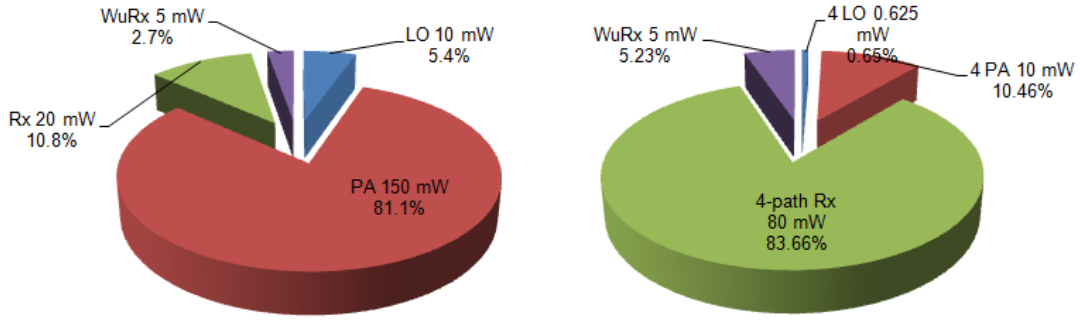
4.2.1 Antenna gain vs. power consumption of multi-path transceiver

As discussed above, with the same link budget requirements, the system-level power can be optimized without knowing the technical details of the circuitry. In this section, the validity of the circuit-related assumptions is discussed by using the state-of-the-art results in the literatures. Taking the 60-GHz low-power circuits as the examples, the power budget is shown in Table 4.2.

Table 4.2 The performance of 60-GHz front ends and RF blocks

Blocks	**Performance**	**Power consumption**
Oscillator [5]	0-dBm output power	10 mW
PA [6]	20% PAE, 15-dB power gain	150 mW (P_{in}=-15 dBm)
WuRx (this work)	-60-dBm sensitivity	5 mW
Rx (this work, one-path)	-60-dBm sensitivity	20 mW

In Fig. 4.8 (a), it is seen that the PA is the dominating factor in the system power when an omni-directional antenna is used, and the total peak power is 185 mW. If the phased array structure is used to produce the additional antenna gain, multiple active blocks will also be required, e.g. n PAs and phase shifters will be needed in an n-path front end. Can the system power still be saved in this way? This question can be discussed in two different situations. The first one is to fix the Rx circuits in each block and optimize the Tx power according the changes in the link model. The second one is to fix the link budget parameters and optimize the Tx and Rx circuit parameters towards lower power consumption. The effectiveness of these approaches is discussed in the following.



(a) (b)

Fig. 4.8 The power distribution of (a) an omni-directional FE and (b) a phased array FE.

A. Scaling the Tx power

In this sub-section, we will consider a case where an *n*-element Tx (*n*-times LO as well as PA) is used, and the output EIRP is still kept as P_{out} in order to maintain the same link budget requirement. The required output power of each Tx path can be reduced to P_{out}/n^2 if the antenna area is not scaled down. Since the PAE of the PA is unchanged by scaling the transistor size of the PA according to the input and output power levels, the LO output power and the power consumption of the PA is scaled down by n^2 for each path accordingly, which means the LO power consumption is also scaled down by n^2 for each path. The overall power consumption of a *n*-element Tx becomes

$$P_{Tx,DC,n\text{-}path} = n \cdot (\frac{1}{n^2} \cdot P_{LO,DC,1} + \frac{1}{n^2} \cdot P_{PA,DC,1}) \tag{4.13}$$

where $P_{LO,DC,1}$ and $P_{PA,DC,1}$ are the power consumption of the single LO and PA respectively. Obviously, on the Tx side, an infinite path number leads to minimum power consumption as long as the LO and PA can be (theoretically) unlimited down-scaled, i.e. using the minimum transistor (power) to generate the required power. When *n* approaches a very large number, the PA is not necessary required in the Tx anymore as the power generated directly from the LO will be sufficient to meet the link requirement.

On the Rx side, however, there are two opposite influences. First, with an *n*-element Rx, the power consumption of the Rx circuits is increased by a factor of *n* if we keep all the blocks and their performance unchanged (e.g. *NF*, linearity, etc.), compared to the single-path omni-directional Rx case. It should be noted that the number of antenna elements in the Rx is not necessarily identical to the one of the Tx. However, a low number of antenna elements leads to less array gain while a large number of antenna elements increases the frequency selectivity of the Rx and causes link difficulty as well. Consequently, the same number of antenna elements is chosen in both Tx and Rx for design simplicity in this analytical model. On the other hand, besides the increase of power consumption, the SNR is improved by a factor of *n* in the Rx too. The reason is that at the adder of the Rx path, *n*-times antenna gain is obtained for the received signal, because it is correlated while the noise (or interference) is not. With the same circuitry of each block, an *n*-times better SNR means that *n*-times less signal power is needed from the Tx. The updated power model of the Tx ($P_{Tx,DC,n\text{-}path,update}$) becomes

$$P_{Tx,DC,n\text{-}path,update} = \frac{1}{n} \cdot n \cdot (\frac{1}{n^2} \cdot P_{LO,DC,1} + \frac{1}{n^2} \cdot P_{PA,DC,1}) \tag{4.14}$$

while the power of the *n*-element Rx $P_{Rx,n\text{-}path}$ becomes

$$P_{Rx,DC,n-path} = n \cdot P_{Rx,DC,1} \tag{4.15}$$

where $P_{Rx,DC,1}$ is the power consumption of the single-path Rx.

Consequently, the overall power of the transceiver becomes

$$P_{TRx,DC,n-path,update} = \frac{1}{n^2} \cdot P_{LO,DC,1} + \frac{1}{n^2} \cdot P_{PA,DC,1} + n \cdot P_{Rx,DC,1} \tag{4.16}$$

Differentiating $P_{TRx,DC,n\text{-}path,update}$ by n, the optimum antenna number n_{opt} is found as

$$n_{opt} = \sqrt[3]{\frac{2 \cdot (P_{LO,DC,1} + P_{PA,DC,1})}{P_{Rx,DC,1}}} \tag{4.17}$$

Substituting (4.17) into (4.16), the minimum power consumption of an n-element phased array transceiver is obtained as

$$\begin{aligned} P_{TRx,DC,n-path,update,min} &= [\frac{2 \cdot (P_{LO,DC,1} + P_{PA,DC,1})}{P_{Rx,DC,1}}]^{\frac{3}{5}} \cdot (P_{LO,DC,1} + P_{PA,DC,1}) \\ &+ \sqrt[3]{\frac{2 \cdot (P_{LO,DC,1} + P_{PA,DC,1})}{P_{Rx,DC,1}}} \cdot P_{Rx,DC,1} \end{aligned} \tag{4.18}$$

Re-using the results shown in Table 4.2, it can be obtained that the overall power consumption of the transceiver is reduced to 95 mW, and the power distribution is illustrated in Fig. 4.8 (b), if n_{opt} is taken as 4 based on (4.17).

Even though some practical issues like the wake up or the link start-up power are not yet included, this derivation can still show the advantage of using the phased array beamforming front end, i.e. without changing the circuitry of each block in the transceiver, the peak power consumption of the main radio is reduced by about 50% compared to a omni-directional front end. It should be mentioned that this calculation only targets at indicating a boundary value of the power optimization rather than proving guidelines for the specific design and implementation of the relevant circuitry, which is also limited by the technology level. For example, a μW-level (also with low output power) 60-GHz LO would be quite a challenging task due to high parasitics of current technologies and low quality factor of the on-chip resonator. Besides, when concerning the application-related additional power consumption like the interference filtering (e.g. by using the smart antenna system), the total power will be higher.

B. Scaling the Rx circuit parameters

In this sub-section, we will scale the Rx circuits parameter while keep the Tx output EIRP constant (i.e. identical to the required P_{out} level to maintain the same range) to verify the power advantage of using an n-element phased array front end. Assuming in a two-stage cascaded Rx system, the noise factors of individual stage (F_1, F_2) are known and the total Rx gain (G_{tot}), noise factor (F_{tot}) as well as the total input-referred intercept point ($IIP3_{tot}$) are pre-determined by the link and application requirements, the minimum power consumption of the Rx can be expressed as [7]

$$P_{DC,tot,min} = IIP3_{tot} \cdot \frac{\kappa_1 \cdot (F_2 - 1) + 2 \cdot \sqrt{G_{tot} \cdot \kappa_1 \cdot \kappa_2 \cdot (F_{tot} - F_1) \cdot (F_2 - 1)} + \kappa_2 \cdot G_{tot} \cdot \sqrt{F_{tot} - F_1}}{F_{tot} - F_1} \tag{4.19}$$

where κ_1 and κ_2 are power-linearity factors of the first and second stages respectively. κ is an EFOM indicating the proportional trade-offs among gain (G), linearity ($IIP3$) and power consumption (P_{DC}) and it is defined as

$$\kappa = \frac{P_{DC}}{G \cdot IIP3} \tag{4.20}$$

For a given technology, operation frequency and circuit structure, κ can be treated as a design constant.

As discussed, the output SNR is improved by a factor of n in an n-element phased array Rx. It can be interpreted as n-times better tolerance of the Rx noise, i.e. n-times larger total noise factor is allowed, while the link parameters like the Rx sensitivity and Tx output power are kept unchanged. Consequently, the Tx power consumption stays the same as described in (4.13). Taking the influence into (4.19), we obtain

$$P_{DC,tot,min,nFtot} = IIP3_{tot} \cdot \frac{\kappa_1 \cdot (F_2 - 1) + 2 \cdot \sqrt{G_{tot} \cdot \kappa_1 \cdot \kappa_2 \cdot (n \cdot F_{tot} - F_1) \cdot (F_2 - 1)} + \kappa_2 \cdot G_{tot} \cdot \sqrt{n \cdot F_{tot} - F_1}}{n \cdot F_{tot} - F_1} \tag{4.21}$$

Taking the results of the state-of-the-art 65-nm CMOS designs of 60-GHz phased arrays [8] [9] into (4.21), i.e. κ_1 is 1, κ_2 is 0.4, F_1 is 10, F_2 is 100, F_{tot} is 20 and G_{tot} is 32, the ratio of the power change (ε_P) is calculated as

$$\begin{aligned} \varepsilon_P &= \frac{\dfrac{\kappa_1 \cdot (F_2 - 1) + 2 \cdot \sqrt{G_{tot} \cdot \kappa_1 \cdot \kappa_2 \cdot (F_{tot} - F_1) \cdot (F_2 - 1)} + \kappa_2 \cdot G_{tot} \cdot \sqrt{F_{tot} - F_1}}{F_{tot} - F_1}}{\dfrac{\kappa_1 \cdot (F_2 - 1) + 2 \cdot \sqrt{G_{tot} \cdot \kappa_1 \cdot \kappa_2 \cdot (n \cdot F_{tot} - F_1) \cdot (F_2 - 1)} + \kappa_2 \cdot G_{tot} \cdot \sqrt{n \cdot F_{tot} - F_1}}{n \cdot F_{tot} - F_1}} \\ &\approx \frac{36.5}{\dfrac{10}{2 \cdot n - 1} + \dfrac{84}{\sqrt{20 \cdot n - 10}}} \end{aligned} \tag{4.22}$$

Taking (4.22) into the total power consumption equation of the n-element Rx, i.e.

$$P_{TRx,DC,n\text{-}path,update2} = \frac{1}{n} \cdot P_{LO,DC,1} + \frac{1}{n} \cdot P_{PA,DC,1} + \frac{1}{\varepsilon_P} \cdot n \cdot P_{Rx,DC,1} \tag{4.23}$$

the updated power consumption is calculated by

$$P_{TRx,DC,n\text{-}path,update2} = \frac{1}{n} \cdot P_{LO,DC,1} + \frac{1}{n} \cdot P_{PA,DC,1} + \frac{\dfrac{10}{2 \cdot n - 1} + \dfrac{84}{\sqrt{20 \cdot n - 10}}}{36.5} \cdot n \cdot P_{Rx,DC,1} \tag{4.24}$$

This updated total power is plotted in Fig. 4.9 as a function of the antenna number.

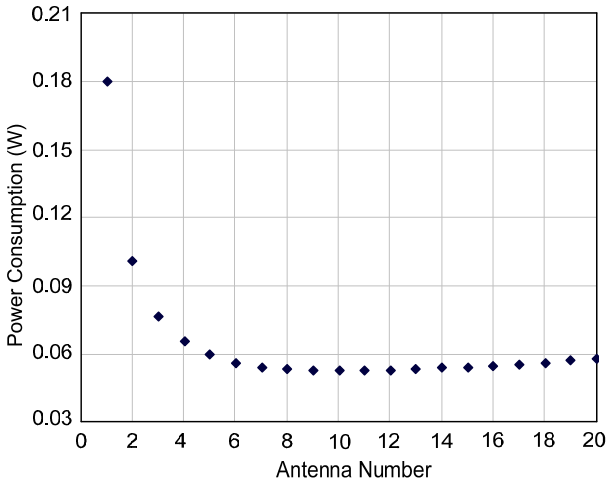


Fig. 4.9 The total power vs. antenna number when scaling the Rx circuit parameters.

In this situation, the power budget of the 4-element transceiver is illustrated in Fig. 4.10. It can be seen from (4.23) that when the Rx circuits consume less power, higher antenna number becomes more beneficial. In the extreme case, i.e. when the Rx ideally consumes zero power, infinite antenna number becomes the most power efficient choice, which is again compatible with the result of (4.13).

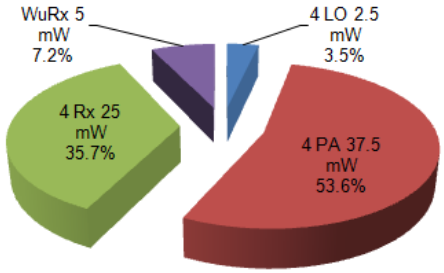


Fig. 4.10 Power budget of the 4-path transceiver array when scaling the Rx parameters.

Comparing the results of (4.24) with (4.18), it can be seen that the second optimization method is more power efficient. However, in practice, this power level may not be achievable due to technology limitations (e.g. minimal transistor size, NF_{min}). Besides, when implementing a multi-element Rx (phased array or true time delay array), the total power consumption will be increased in order to compensate the extra insertion losses, mismatches and couplings in the antenna, delay elements, adder, and so on.

4.2.2 System bandwidth discussion

Under the name “bandwidth”, there are actually several terms with very different physical meanings such as the half-power bandwidth (also known as the 3-dB bandwidth), noise

equivalent bandwidth, null-to-null bandwidth, fractional power containment bandwidth, and so on. Although it seems that bandwidth is a fully frequency-domain concept, it can also be observed by adding time variable as another dimension. This concept is discussed as follows.

Based on the above discussions, two different bandwidth concepts can be distinguished: the instantaneous signal bandwidth (BW_{inst}) and the total effective signal bandwidth (BW_{eff}). The former is the real signal bandwidth under certain time duration, and it determines the noise level of the system and the shape of a data pulse. The latter is the total effective bandwidth. By introducing these concepts, we can treat wideband signals as narrowband ones in some time slots. It should be noted that this is not generic and only valid for some particular modulation schemes and circuitries. In fact, there are two prerequisites that determine whether we can use this method: (1) the instantaneous bandwidth does not change fast. This is very similar to the inter-symbol interference (ISI) problem in the time-domain, i.e. when the symbol rate is high, time-domain signals will show more overlap with each other due to finite filter bandwidth and the ISI problem will eventually cause demodulation errors. (2) The circuits (which are virtually narrowband) are able to track the change of instantaneous signal bandwidth. However, a real-time dynamic tracking filter is normally complicated and it will introduce inaccuracies or loss. As a result, a new block that fulfills these two requirements is required.

Consider the operation principle of a spectrum analyzer: if we can "sweep" the BW_{eff} slowly by the step of BW_{inst} in the time duration L_{sweep}, it can be seen that the total sweeping time L becomes a new trade-off factor, i.e.

$$BW_{eff} = BW_{inst} \cdot \frac{L_{sweep}}{t_0} \tag{4.25}$$

where t_0 is the time duration of every BW_{inst}.

As part of the circuit latency, L_{sweep} should normally be very low especially for a low duty-cycle and burst-mode (high-data-rate and short active time) communication system. Thus, t_0 becomes a crucial optimization factor. In the ideal case, the minimum t_0 is identical to the settling time of the Rx. For instance, if a PLL is used in the Rx, t_0 should be at least adequate to start and stabilize the PLL output frequency for LO drivers. However, the settling process of the conventional PLL is normally very slow especially when operating at higher frequencies with constrained power consumption, which causes an unaffordable t_0 (very long overhead in front of the data packet) compared to payload length and deteriorates the effectiveness of the above mentioned bandwidth transformation techniques. As a result, the design of a fast-settled LO becomes very important. This concept is essential to the WuRx design and we will discuss the relevant details in the following sections.

4.3 IJLO based asynchronous duty-cycled WuRx

As discussed in the previous section, the settling time of the WuRx is very critical because it directly affects the feasibility of the bandwidth extension technique of the WuRx. A very long settling time of the WuRx will cause an unaffordable high latency of the system and it is power inefficient. Besides, it has also been discussed in Section 4.1.2 that the settling time of the WuRx (denoted as T_{wu} in Section 4.1.2) plays a very important role in minimizing the average power consumption of the entire system. However, T_{wu} is highly circuit-dependent and varies significantly with specific WuRx architectures, circuit designs and implementation technologies. In this section, the theory and design methodology of a low-power and fast-settling WuRx is provided.

The basic functions of the WuRx are: (i) receiving the wake-up beacon from the Tx and (ii) switch the main radio Rx on and off accordingly. Therefore, it is required to have the following features: (i) low-power; (ii) highly sensitive; (iii) fast self-settled when transiting from sleep to active mode; (iv) is able to re-use the transceiver resources, e.g. is able to operate at the same mmW band and use the main radio Tx to send the beacon; (v) is able to cover the entire system band and (vi) it should also be low-cost and with a high integration level.

Conventional low-power WuRxs tend to use the envelope detection structure, e.g. a low-biased diode detector. The received RF signal is multiplied by itself to obtain its envelope as the output signal. However, the sensitivity of the diode envelop detectors is usually poor due to their “quadratic” nature between the output and input signals. Besides, they are not suitable for very high frequency applications due to the poor sensitivity of the silicon-based on-chip diodes at mmW frequencies [10]. Some literature propose to add a low-noise amplifier in front of the diode to improve its sensitivity, but such a wideband mmW amplifier again requires high power consumption [11].

Besides the poor sensitivity, the conventional envelope detector is unable to identify the frequency information of the RF signal (i.e. normally no precise selectivity in frequency domain), which normally counts on a PLL in the LO path. However, the PLL is historically very power-hungry and slow especially when working at higher frequencies. For instance, a state-of-the-art 35 to 41.88 GHz PLL in 65-nm CMOS is reported recently with 80 mW power consumption that is comparable to the power consumption level of the front end [12]. Furthermore, the typical settling time of a PLL is typically as long as 40 to 300 μs [13], which means a long packet overhead will be required for settling the LO. It is particularly power-inefficient for the burst-mode low duty-cycled radios in which the payload data length is much shorter than that.

At this point, the following questions arise: “can we identify the frequency information by WuRx without using PLL”, “can we get rid of the quadratic feature of the envelope detection and have very good sensitivity” and “can we realize everything on chip with low power and low cost”? To answer the above questions, an injection-locked oscillator (IJLO) based WuRx is proposed, as shown in Fig. 4.11. The incoming RF signal is split into two branches. One is directly sent to the IJLO. When the IJLO is locked, it will generate a large constant-envelope output voltage as the LO driver for a passive mixer. The other path of the RF is fed into the mixer as the RF input. At the output of the mixer, a high DC voltage will appear at the locking condition and it will be used to switch on the main radio. When there is no RF coming or the IJLO is unlocked, the output of the passive mixer will be a low voltage at the DC and thus main Rx is kept at the “off” mode.

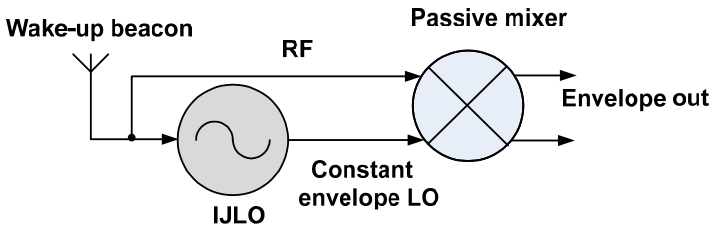


Fig. 4.11 An IJLO based WuRx.

With the IJLO, the “quadratic” behavior of the WuRx is replaced by a linear relation, and the sensitivity is improved significantly. Since the frequency information of the incoming RF signal has been identified during the locking process of the IJLO, it can be used in the main radio Rx and thus the PLL is not required anymore. Last but not least, the IJLO based WuRx can be fully implemented on chip without using any off-chip filter or high-Q resonator, so it can be sufficiently low cost and compact. One may come to a question that “what about the bandwidth, sensitivity and settling time of IJLO itself”? The basic principles of the IJLO-based WuRx are discussed in the following sections.

4.3.1 Basic theory of the IJLO

The conceptual model of a conventional IJLO is illustrated in Fig. 4.12 (a). In this model, a common-source transistor with an ideal positive feedback is used to approximate the oscillator core and an LC tank is used as the frequency selectivity (resonator). When an oscillator is under injection, e.g. by a small current injection I_{inj} directly to its resonator as shown in Fig. 4.12 (a), a phase shift will be created between its oscillation signal and the resonator path, i.e. between I_T and I_{osc}. In order to compensate the extra phase and maintain the 360° loop phase shift, the oscillator core will shift its free-running frequency to the injected one. In other words, the oscillator is locked to the injected signal [14].

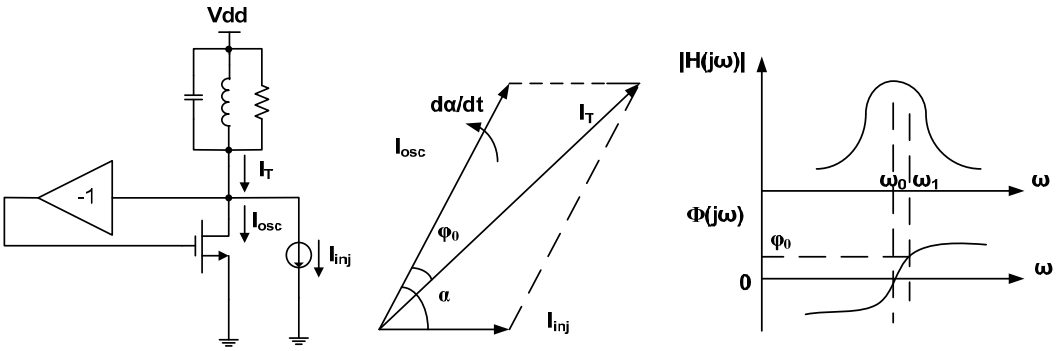


Fig. 4.12 (a) The conceptual IJLO model, (b) the phasor diagram at a given instant and (c) the open-loop characteristics of the resonator [10].

Locking range calculation of an IJLO can be found in [14] and [15]. In this section, we cite the locking range derivation methodology of [14] and use it as the foundation of our following studies of IJLO settling performance. It should be noted that this derivation flow is only valid for LC oscillators in which the quality factors of the resonators are explicitly defined. For other oscillator like ring oscillators, a very different methodology should be used when investigating its locking range and phase noise. More details can be found in 5.3.2. The derivation of the locking range can be explained using the phasor diagram as shown in Fig. 4.12 (b). The instantaneous phase difference between the oscillation current I_{osc} and resultant current I_T, i.e. φ_0 is directly related to the instantaneous phase difference between I_{osc} and I_{inj}, i.e. α as

$$\sin\varphi_0 = \frac{I_{inj}}{I_T}\cdot\sin\alpha \tag{4.26}$$

or

$$\sin\varphi_0 = \frac{I_{inj}\cdot\sin\alpha}{\sqrt{{I_{osc}}^2 + {I_{inj}}^2 + 2I_{osc}\cdot I_{inj}\cdot\cos\alpha}} \tag{4.27}$$

$\sin\varphi_0$ reaches its maximum value when I_T reaches its minimum value, i.e. perpendicular with I_{inj} as

$$\sin\varphi_{0,\max} = \frac{I_{inj}}{I_{osc}} \tag{4.28}$$

The phase shift of the LC tank in the vicinity of resonance is given by

$$\tan\varphi \approx \frac{2Q}{\omega_0}\cdot(\omega-\omega_0) \tag{4.29}$$

where ω_0 is the center frequency of the resonator and Q is the loaded quality factor of the resonator. The instantaneous phase φ_0 reaches its maximum value at the edge of the locking condition. From (4.28) and (4.29), it can be written that

$$\tan\varphi_{0,\max} \approx \frac{2Q}{\omega_0}(\omega_{L,1}-\omega_0) = \frac{I_{inj}}{I_T} = \frac{I_{inj}}{\sqrt{I^2_{osc}-I^2_{inj}}} \tag{4.30}$$

or

$$\omega_{L,1} = \omega_{L,max} - \omega_0 = \frac{\omega_0}{2Q}\cdot\frac{I_{inj}}{I_T} = \frac{\omega_0}{2Q}\cdot\frac{I_{inj}}{I_{osc}}\cdot\frac{1}{\sqrt{1-\frac{I^2_{inj}}{I^2_{osc}}}} \tag{4.31}$$

It is noted that $\omega_{L,1}$ is the one-side locking range (higher locking sideband) and the double-side locking range is twice of that.

Under small injection assumption, i.e. $I_{inj} << I_{osc}$ (this is a reasonable assumption for our mmW system), (4.31) can be approximated as

$$\omega_L = \frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}} \tag{4.32}$$

When the injection signal lies within the frequency range of ω_L, the IJLO will "capture" the injection frequency information and shifts its own free-running frequency to the injection frequency.

Besides proving the frequency-locking feature, the locking activity is also able to reduce the phase noise of the free-running oscillator. When the IJLO is completely locked to the injection signal, its phase noise performance will be a replica of the injection signal. This feature will be further discussed in Chapter 5 together with the details about IJLO circuit design.

The settling performance or the locking time of the IJLO is another critical parameter. As mentioned at the beginning of Section 4.3, the fast self-settling feature is required in order to achieve the power-efficient wake-up performance. The settling time of the IJLO is analyzed as follows.

In the well-known Adler's equation [15], the instantaneous phase change under injection of an IJLO can be modeled as

$$\frac{d\theta}{dt} = \omega_0 - \omega_{inj} - \omega_L\cdot\sin\theta \tag{4.33}$$

where θ is the instantaneous phase difference between the IJLO natural frequency ω_0 and the injection frequency ω_{inj}.

It can be seen that (4.33) should be solved under two difference conditions. Firstly, when $(\omega_0 - \omega_{inj})^2 > \omega_L{}^2$, i.e. the injection signal lies out of the locking range, the phase-varying time *t* is solved as

$$t = \frac{1}{\sqrt{(\omega_0 - \omega_{inj})^2 - \omega_L{}^2}} \cdot \tan^{-1}\left(\frac{-\omega_L + (\omega_0 - \omega_{inj}) \cdot \tan(\theta/2)}{\sqrt{(\omega_0 - \omega_{inj})^2 - \omega_L{}^2}}\right) + C_0 \tag{4.34}$$

where C_0 is the integration constant. This solution can be used to study the quasi-lock and fast-beat behaviors as shown in [12], which is very important for a PLL design. However, in our IJLO system, there is no low-frequency feedback or frequency beating processes. Therefore, we only focus on the locking phase of the IJLO.

A second solution of (4.33) is obtained under the condition $(\omega_0 - \omega_{inj})^2 < \omega_L{}^2$, i.e. when the locking condition is satisfied, the phase varying time *t* becomes

$$t = \frac{1}{\sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2}} \cdot \ln\left(\frac{-\omega_L - \sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2} + (\omega_0 - \omega_{inj}) \cdot \tan(\theta/2)}{-\omega_L + \sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2} + (\omega_0 - \omega_{inj}) \cdot \tan(\theta/2)}\right) + C_1 \tag{4.35}$$

where C_1 is the integration constant. When the IJLO goes into the steady state, the phase difference between the injected and the oscillation signals is calculated as

$$\theta_{ss} = \arcsin\left(\frac{\omega_0 - \omega_{inj}}{\omega_L}\right) \tag{4.36}$$

Substituting (4.36) into (4.35), the settling time of the IJLO becomes

$$t_{settling} = \frac{1}{\sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2}} \cdot \ln\left(\frac{\omega_L + \sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2} - (\omega_0 - \omega_{inj}) \cdot \tan\left(\frac{\arcsin\left(\frac{\omega_0 - \omega_{inj}}{\omega_L}\right)}{2}\right)}{\omega_L - \sqrt{\omega_L{}^2 - (\omega_0 - \omega_{inj})^2} - (\omega_0 - \omega_{inj}) \cdot \tan\left(\frac{\arcsin\left(\frac{\omega_0 - \omega_{inj}}{\omega_L}\right)}{2}\right)}\right) + C_1 \tag{4.37}$$

It is seen that at the initial condition, (4.37) is equivalent to

$$0 = a \cdot \ln\left(\frac{b + d \cdot \infty}{c + d \cdot \infty}\right) + C_1 \tag{4.38}$$

where *a*, *b*, *c* and *d* are non-zero values. As a result, the maximum value of C_1 is zero and in the most of the cases, it is a negative value. In this worst case design, C_1 is always taken as zero. It can be seen that the locking range of the IJLO should be sufficiently large to cover the entire system bandwidth. For example, in a 60-GHz radio, the locking range should at least identical the system bandwidth, i.e. 7 GHz (from 57 to 64 GHz). According to the linear relationship described in (4.32), a larger locking range can be achieved by injecting a higher power into the IJLO while fixing I_{osc} (in fact, the minimum allowable I_{osc} has already been pre-determined by the required power to the next stage as

well as the oscillation condition of the free-running IJLO and it cannot be reduced arbitrarily). However, this is inconsistent to our previous argument that the IJLO should also be highly sensitive to the injection signal in order to avoid an unrealistically high wake-up beacon power.

This problem can be solved by introducing the frequency-sweeping locking approach. Recalling the bandwidth discussion in Section 4.2.2, under a weak injection level, the single-step locking range (or understood as the instantaneous bandwidth) of an IJLO will be very small. We sweep the center frequency of the free-running IJLO with a certain frequency step (sweeping resolution, which should be covered by the actually locking range) crossing the total effective system bandwidth, and stop sweeping when the IJLO gets locked to the RF signal. It can be seen that in every time slot t_0 (here it is identical to the settling time of IJLO), the IJLO becomes a narrowband system, which eliminates the requirement of high injection power to support wide locking range in order to cover the total band. However, when observing in the system level, the frequency-sweeping IJLO achieves very good wideband performance, i.e. its total effective locking range is sufficiently large to cover the entire band. It should also be noted that the time duration of each sweep should be longer than the settling time of the IJLO. In such a way, the trade-off between the locking range and sensitivity is transformed into the trade-off between the circuit latency (total settling time of the IJLO) and the sensitivity while the total effective bandwidth (total locking range) can be maximized according to the system requirement. Keeping this concept, let us have a further discussion about the frequency-sweeping mechanism of the IJLO.

4.3.2 WuRx with frequency-sweeping injection locking mechanism

Substituting (4.32) into (4.37), with an injection current level I_{inj}, the single-step settling time (i.e. the actual locking range related to certain I_{inj} and I_{osc} levels) of an IJLO can be calculated as

$$t_{settling,single\text{-}step} = \frac{1}{\sqrt{(\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}})^2-(\omega_0-\omega_{inj})^2}}\cdot \ln(\frac{\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}}+\sqrt{(\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}})^2-(\omega_0-\omega_{inj})^2}-(\omega_0-\omega_{inj})\cdot\tan(\frac{\arcsin(\frac{\omega_0-\omega_{inj}}{(\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}})})}{2})}{\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}}-\sqrt{(\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}})^2-(\omega_0-\omega_{inj})^2}-(\omega_0-\omega_{inj})\cdot\tan(\frac{\arcsin(\frac{\omega_0-\omega_{inj}}{(\frac{\omega_0}{Q}\cdot\frac{I_{inj}}{I_{osc}})})}{2})}) \tag{4.39}$$

The frequency sweeping resolution is assumed as twice as large as the frequency

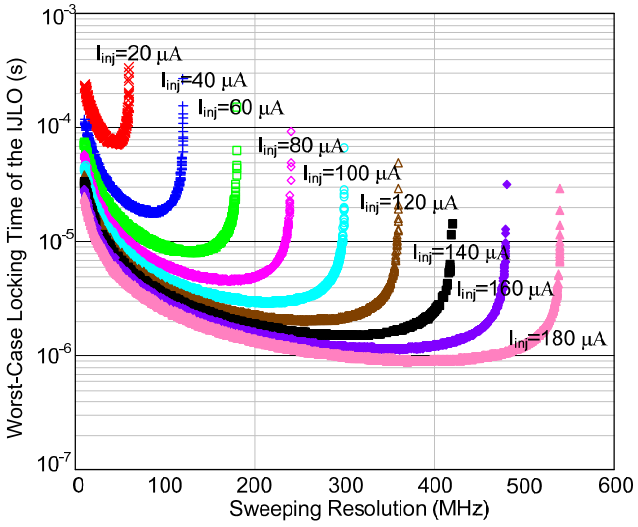
difference between the center frequency and the injection frequency, i.e.

$$\omega_{res} = 2 \cdot (\omega_0 - \omega_{inj}) \tag{4.40}$$

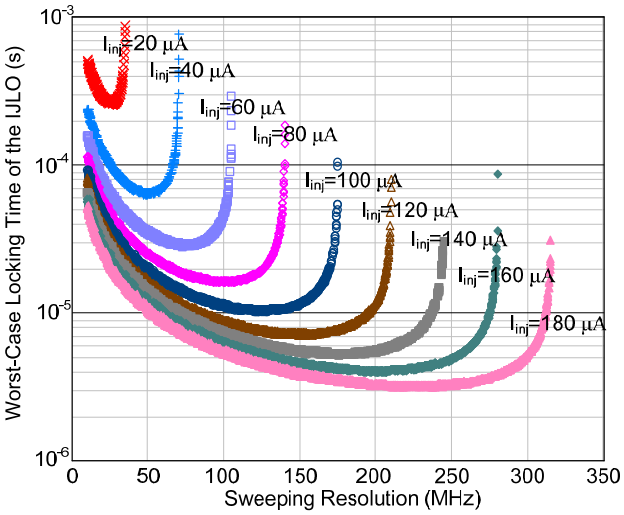
If we let the IJLO sweep the total system bandwidth BW_{tot} (in Hz), the worst-case settling time of the IJLO including the sweeping process can be calculated as

$$t_{settling,tot} = \frac{BW_{tot} \cdot 2\pi}{\omega_{res}} \cdot \frac{1}{\sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2}} \cdot \ln(\frac{\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}} + \sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2} - 2\omega_{res} \cdot \tan(\frac{\arcsin(\frac{2\omega_{res}}{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})})}{2})}{\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}} - \sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2} - 2\omega_{res} \cdot \tan(\frac{\arcsin(\frac{2\omega_{res}}{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})})}{2})}) \tag{4.41}$$

In can be seen that the total settling time calculated from (4.41) actually indicates the worst-case value, which happens when the injection signal frequency lies at one edge of the system band and the free-running frequency of the IJLO lies at the other edge. In such a case, the IJLO has to sweep the entire frequency band in order to find the injection signal and lock its frequency. At the best case when only one sweep is enough, the total settling time becomes identical to the single-step settling time. In the following analysis, the worst-case assumption will always be used. It can be seen that under the same injection level, when ω_{res} becomes larger, the number of sweeps in the worst-case will become smaller but the settling time of the IJLO will become longer. Consequently, an optimum sweep step should exist, which leads to the minimum worst-case settling time. Taking the 60-GHz ISM band and a 70-GHz (W-band) IJLO as examples, and assuming I_{osc} is 5 mA and the Q of the on-chip resonator is 10, the total worst-case locking times of them can be plotted in Fig. 4.13 (a) and (b).



(a)



(b)

Fig. 4.13 The worst-case settling times of (a) a 60-GHz IJLO with a 7-GHz total bandwidth and (b) a 70-GHz IJLO with a 10-GHz total bandwidth.

It can be seen that the total worst-case settling time of the IJLO is a function of the sweeping resolution (frequency step) and the injection current strength (sensitivity).

The settling performance of the frequency sweeping IJLO is compared with the conventional PLL as follows. The settling time of a PLL can be modeled as [1]

$$\begin{aligned} t_{settling,PLL} &= t_{lock-in} + t_{capture} \\ &= \sqrt{\frac{2 \cdot N_{mean}}{K_{VCO}}} \cdot \sqrt{\frac{C}{I_{CH}}} + \frac{2 \cdot |\Delta f|}{K_{VCO}} \cdot \frac{C}{I_{CH}} \end{aligned} \tag{4.42}$$

where K_{VCO} is the required VCO gain, C is the total capacitance of the RC loop filter, I_{CH} is the bias current of the charge pump Δf is the frequency jump after a change in the frequency division ratio (between reference and carrier frequency) N value and N_{mean} is the geometric mean value of N.

Assuming a 1-GHz off-chip resonator (e.g. a bulk acoustic wave resonator) is used as the frequency reference, the total settling time of the 60-GHz and 70-GHz PLLs can be plotted as a function of the bias current of the charge pump as shown in Fig. 4.14. It should be noted that the settling time of a PLL is directly related with the bandwidth of the LPF. As shown in (4.42), lower C will lead to lower settling time at the same I_{CH} level. However, it will cause a higher level of spurs, too [1]. In the following analyses, C is taken as 50 μF and 67.5 μF for the 60-GHz and 70-GHz PLLs respectively, as a trade-off between the power consumption and the level of spurs. In practice, a more advanced active LPF could be used in order to avoid such large capacitance values.

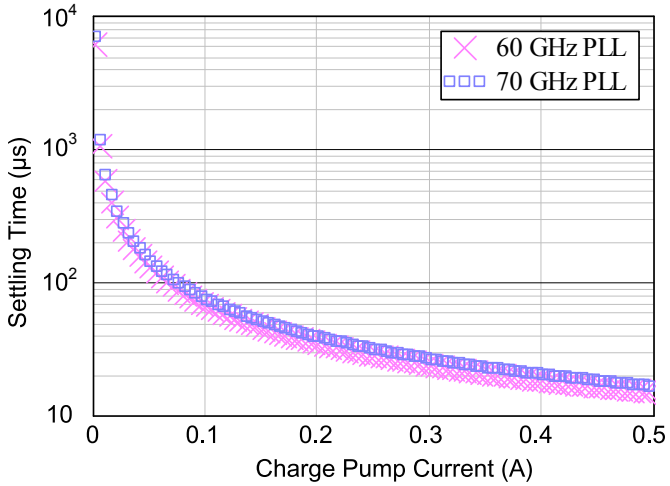


Fig. 4.14 Settling times of 60-GHz and 70-GHz PLLs as a function of bias currents of charge pumps.

The total power consumption of the PLL can be calculated as

$$\begin{aligned} P_{PLL} &= P_{PLL,lock-in} + P_{PLL,capture} \\ &= [N_{PFD} \cdot G_{PFD} + N_{mean} \cdot (2 \cdot n^2 \cdot k \cdot 8 + N_{div} \cdot G_{div})] \cdot C_{tech} \cdot V_{supply}{}^2 \cdot f_{ref} \\ &+ (C_{PFD} + C_{VCO} + C_{div}) \cdot V_{supply}{}^2 \cdot f_{capture,ave} + P_{CH} \end{aligned} \tag{4.43}$$

where N_{PFD} is the number of minimum size transistors in the P_{FD}, G_{PFD} is the average gate activity of the P_{FD}, k defines the voltage swing in each VCO cell, n is the number of stages of the VCO, N_{div} the number of minimum size transistors in the frequency divider, G_{div} is the average gate activity of the frequency divider, C_{tech} is the total capacitance, C_{PFD} is the product of N_{PFD}, C_{PFD} and C_{tech}, C_{VCO} is total capacitance of the VCO resonator, P_{CH} is the power consumption of the charge pump and $f_{capture,ave}$ is the average capture frequency. The power consumption of the PLL can be plotted in Fig. 4.15.

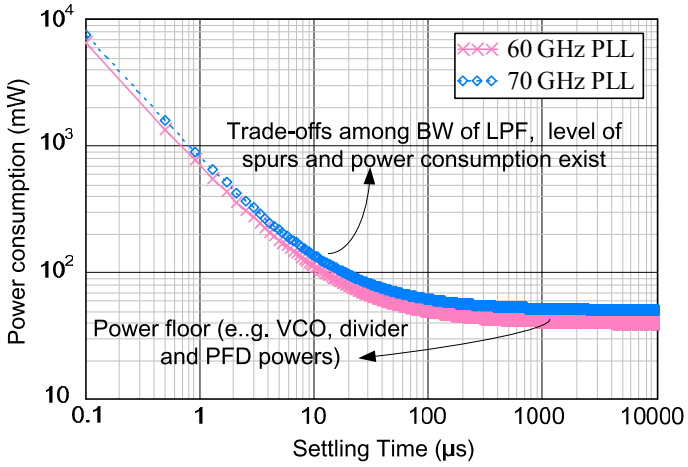


Fig. 4.15 Power consumption of 60-GHz and 70-GHz PLLs.

It can be seen that with similar power consumption level, the IJLO-based frequency capturing method takes at least 10-times less time for settling compared to a conventional PLL. Moreover, its sensitivity can be improved significantly through the frequency sweeping method. However, referring to Fig. 4.11, the overall sensitivity of the WuRx is determined not only by the IJLO, but also by the passive mixer as illustrated in Fig. 4.16.

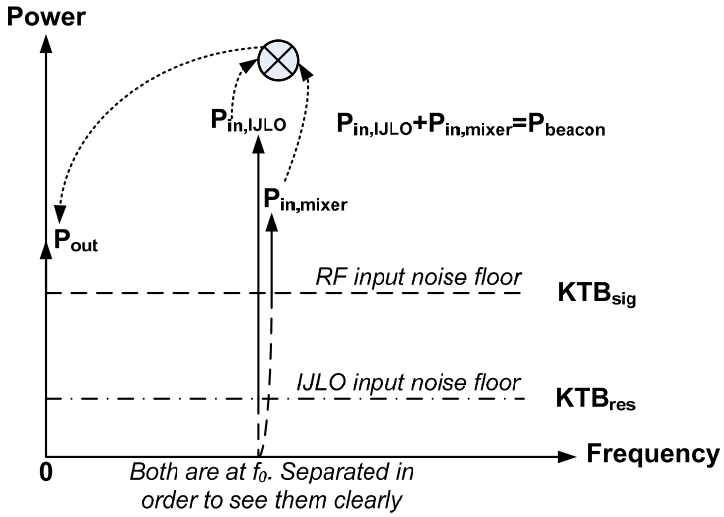


Fig. 4.16 Illustration of the sensitivity of the WuRx.

It can be seen that the noise floor of the mixer input, i.e. the equivalent RF port should be reduced simultaneously with the equivalent LO port. A possible solution for this issue is discussed in the following section.

4.3.3 Bandwidth extension and sensitivity enhancement of the WuRx

The most straightforward method to reduce the signal noise floor as well as to remove the blockers is to add a band-pass filter (BPF) in front of the equivalent RF path, as shown in Fig. 4.17.

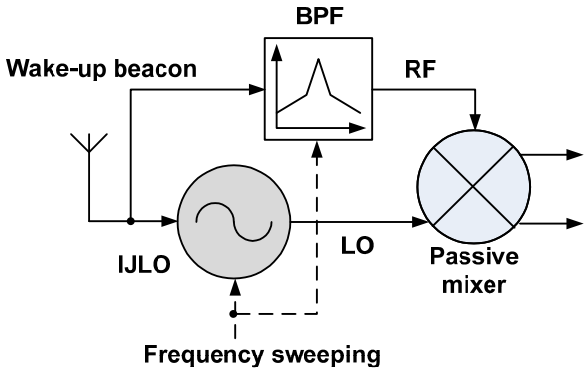


Fig. 4.17 A tunable BPF-first WuRx.

In order to minimize the noise bandwidth further, the BPF can be designed to be dynamically tunable. While sweeping the center frequency of the IJLO, the center frequency of the BPF is also swept with the same frequency step. In this way, the IJLO input noise floor and the RF input noise floor can be reduced simultaneously without affecting the locking condition. However, this BPF-first WuRx architecture has a poor noise factor penalty, because the BPF should be passive in order to provide the low-power performance of the WuRx.

Thanks to the impedance transparency feature of the passive mixer, this problem can be easily solved. Instead of adding a tunable BPF in front of the RF path, a simple non-tunable narrow-band LPF can be placed at the output of the passive mixer, and it has an equivalent effect as the BPF-first scheme [16], but the system noise factor is not impaired in this way. In order to enhance the RF gain, an output buffer amplifier can be used, which is able to amplify the output signal as well as to act like a LPF.

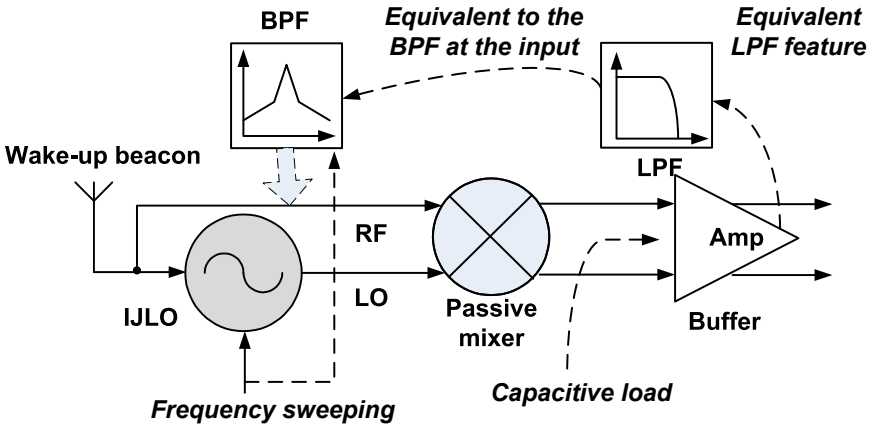


Fig. 4.18 The WuRx with an LPF at its output.

Since the input impedance of the buffer amplifier is capacitive, it behaves as a high-impedance at lower frequencies and low-impedance at higher frequencies. 'Seen' from the RF input, it yields a frequency response similar to that of a high-Q BPF due to frequency translation, which reduces the noise bandwidth of the mixer input dramatically and improves the WuRx sensitivity accordingly. More circuit-related details will be discussed in the Chapter 5.

4.3.4 Summarized design considerations and methods

The proposed WuRx design methods and their functions are summarized in Table. 4.3. It can be seen that by using these methods, the WuRx is able to achieve high sensitivity, wide bandwidth and low latency simultaneously.

Table 4.3 Summary of the WuRx performance enhancement methods

Methods	**Functionalities**
Injection-locking oscillator	Remove the quadratic behavior in self-mixing envelope detection
	Sensitivity improvement for the envelope detection
	Fast settling process
Frequency sweeping	Bandwidth extension
	Sensitivity improvement for the IJLO
	Trade-off between total latency and sensitivity
Passive mixer and output LPF	Bandwidth extension
	Sensitivity improvement for the passive mixer

4.3.5 Interference tolerance: pseudo capture effect

As discussed in Section 4.3.4, high power blocker signals can be easily suppressed in the WuRx due to the equivalent BPF effect of the output LPF. However, when the co-channel interference lies within the single-step locking range of the IJLO and with relatively high power, what will happen in the IJLO-based WuRx?

The conventional FM receiver is less sensitive to co-channel interferences due to the well-know "capture effect". The basic reason is that when a frequency modulated interference signal appears at the FM receiver with a very close-in frequency as the desired signal, it can be suppressed by the FM demodulator to a very high degree if it is sufficiently weak. The background theory is explained in [17]. Since the wake-up beacon in our work is pure RF carrier without any modulation applied, the WuRx will, in principle, not be capable to capture the desired RF signal while suppressing the co-channel interference. However, there is a similar effect happening in the IJLO-based WuRx and we call it "the pseudo capture effect".

Suppose, the desired signal and the interferer are expressed as

$$S_d(t) = A_d \cdot \cos(\varphi_d) = A_d \cdot \cos(\omega_d \cdot t) \tag{4.44}$$

$$S_i(t) = A_i \cdot \cos(\varphi_i) = A_i \cdot \cos(\omega_i \cdot t) \tag{4.45}$$

where A_d and A_i are their amplitudes, φ_d and φ_i are their phases and ω_d and ω_i are their angular frequencies.

The amplitude of the resultant signal becomes

$$A_R = \sqrt{A_d{}^2 + A_i{}^2 + 2 \cdot A_d \cdot A_i \cdot \cos \Delta\varphi} \tag{4.46}$$

and the resulting phase as

$$\psi_R = \varphi_d + \tan^{-1}\left(\frac{r_A \cdot \sin\Delta\varphi}{1 + r_A \cdot \cos\Delta\varphi}\right) \tag{4.47}$$

where $\Delta\varphi$ is the difference between φ_d and φ_i, and r_A is the ratio between A_i and A_d. ψ_R can be further expanded in a Taylor's series as

$$\psi_R = \varphi_d - \sum_{n=1}^{\infty} \frac{(-r_A)^n}{n} \cdot \sin n\Delta\varphi \tag{4.48}$$

The instantaneous frequency of the resultant signal can be calculated by differentiating ψ_R as

$$\frac{d\psi_R}{dt} = \omega_d - \sum_{n=1}^{\infty} (-r_A)^n \cdot (\omega_d - \omega_i) \cdot \cos[n \cdot (\omega_d - \omega_i) \cdot t] \tag{4.49}$$

Due to the inter-modulation products of ω_d and ω_i, the injection frequency experiences a "jitter" in both amplitude and phase. As a result, the angular frequency is also not stable. Since r_A<<1, the factor $(-r_A)^n$ rapidly approaches 0 for n larger than 1. Consequently, the instantaneous angular frequency from (4.49) can be approximately simplified as

$$\omega_{inst} = \omega_d + r_A \cdot (\omega_d - \omega_i) \cdot \cos[(\omega_d - \omega_i) \cdot t] \tag{4.50}$$

When the frequency difference is small, e.g. less than one tuning step of the IJLO (50 to 100 MHz), the variation is very slow. According to the theories of the IJLO, the oscillator core is able to adjust its oscillation frequency along with the instantaneous injection frequency as described in (4.50) and thus produce the same frequency pattern at its output. The frequency variation is plotted in Fig. 4.19.

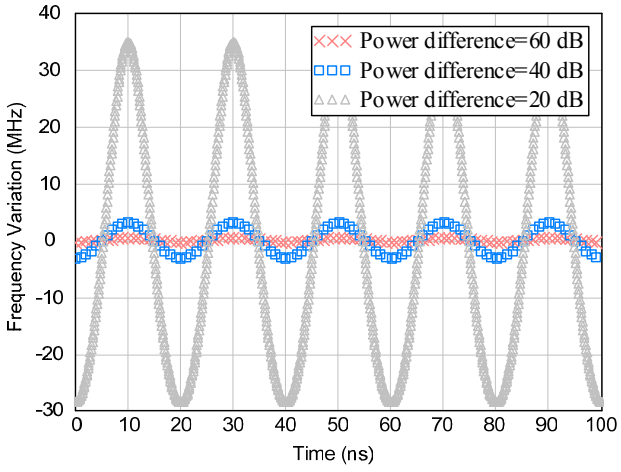


Fig. 4.19 The IJLO frequency variation under the influence of the interferer.

On the other hand, the passive mixer is a highly linear circuit, and thus the simple superposition of the two components can be assumed. When multiplying them with the instantaneous frequency described in (4.50), the output frequencies of the passive mixer

turns out to be composed of two components, i.e.

$$\omega_{out,d} = r_A \cdot (\omega_d - \omega_i) \cdot \cos[(\omega_d - \omega_i) \cdot t] \tag{4.51}$$

$$\omega_{out,i} = (\omega_d - \omega_i) \cdot \{1 + r_A \cdot \cos[(\omega_d - \omega_i) \cdot t]\} \tag{4.52}$$

Since r_A<<1 (e.g. r_A equals 0.01 and a 40-dB power difference exists between $s_d(t)$ and $s_i(t)$ as illustrated in Fig. 4.19), it can be seen that the output component produced by the desired signal, i.e. the $\omega_{out,d}$ component is very close to DC, and the output component produced by the interferer, i.e. the $\omega_{out,i}$ component can be further approximated as

$$\omega_{out,i} \approx \omega_d - \omega_i \tag{4.53}$$

As long as pass band of the output LPF is less that ω_d-ω_i, the output component produced by the interferer will be filtered out and therefore the influence of the interference signal will be suppressed. This is very close to the "capture effect" in conventional FM radios.

In fact, however, two different phases should be considered separately, i.e. the sweeping phase and the lock-in phase. In the sweeping phase, it is likely that the IJLO "meets" the interference frequency first. In this situation, the interference power level should be sufficiently low to prevent the IJLO to lock to it within one tuning step. Otherwise, a missed-alarm will happen, because if the IJLO already gets locked to the interference, the sweeping mechanism will be stopped and the desired signal will not be detected any more. With the assumption r_A<<1, the single-step locking time of the interference signal will be much longer than the one of the desired signal, and thus the problem described above will not exist. When r_A is close to 1 or even larger, the WuRx will be fully blocked and this problem should be prevented in the MAC and network layers instead. For instance, by using the CSMA-CA MAC protocol, high-power blockers in the channel are avoided in advance. If a blocker is not identified before beginning of the communication process, re-transmissions will be required with a random time period later.

4.3 Main radio power optimization

In Section 4.2.1, the theoretical minimum achievable power consumption of the transceiver has been derived as a function of the optimum antenna number. The basic idea is that with more antenna elements, a better beam directivity and larger antenna gain can be obtained to combat the loss and thus save the active power. On the other hand, more power would be dissipated in such a beamforming front end, i.e. consumed by multiple active devices in the array. In a particular situation, an optimum antenna number exists, which leads to the minimum overall power consumption. However, this power model only considers the physical circuitry when transmitting the payload of a packet. In reality, some more pre-communication processes should be also included in the model, because

they will affect the total power consumption significantly. For example, the neighborhood discovery process would be quite problematic in the directional communication system.

One of the most straightforward methods for a node to "get acquainted" with its neighbors is to know their positions in advance and then store the most frequent communication paths in its memory. Before communicating, the nodes will select the most possible directions to steer its beam. This technique is well described in several machine-learning networks. In this case, the training process of the entire network would consume a large portion of power.

The localization process of the nodes can be done by using two base stations with the angle-of-arrival algorithm (AoA). As illustrated in Fig. 4.20, by knowing the positions of two base stations A (x_1, y_1) and B (x_2, y_2), the position of the node P (x_p, y_p) can then by calculated by (4.54) to (4.58).

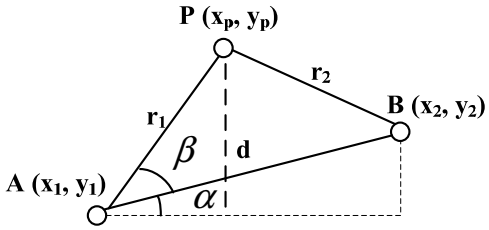


Fig. 4.20 AoA algorithm.

$$\alpha = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \tag{4.54}$$

$$\beta = \arccos\left(\frac{r_1^2 - r_2^2 + d^2}{2r_1 \cdot d}\right) \tag{4.55}$$

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4.56}$$

$$x_p = x_1 + r_1 \cos(\alpha \pm \beta) \tag{4.57}$$

$$y_p = y_1 + r_1 \sin(\alpha \pm \beta) \tag{4.58}$$

If the network varies very slowly (the information of x_p and y_p is kept unchanged or has regular moving patterns), the pre-positioning approach will be very effective to reduce the MAC complexity and therefore save the overall power consumption. However, when the nodes are moving (pseudo) randomly, the beam scanning process will be necessary for the neighborhood discovery. This is particularly important for the Tx to send the wake-up calls.

4.4.1 Optimized antenna number—beam scanning related

In the wake-up process, a directional to omni-directional (DO) link is established, i.e. the antenna of the WuRx is omni-directional while the wake-up beacon sent by the Tx is directional. In order to find the desired WuRx, the Tx will scan the environment with its directional beam (the wake-up beacon) until the correct WuRx is identified, as shown in Fig. 4.21. Since the beam scanning process is complex and time-consuming, the overall power model need to be updated accordingly.

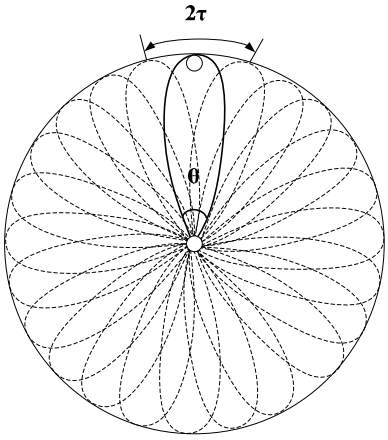


Fig. 4.21 Beam scanning process.

Denoting the beam width as θ and the scanning time step as τ, the single-cycle scanning time t_{scan} is

$$t_{scan} = 2\tau \cdot \frac{360}{\theta} \tag{4.59}$$

where an additional time period τ is reserved as the acknowledgement receiving period. The relationship between the expected number of scanning and the neighborhood node degree (the number of the neighbor nodes) as well as the antenna beam width is thoroughly analyzed and mathematically calculated in [18]. Using its results, the number of scanning can be approximately modeled as a function of the antenna beam width (in degree). In such a case, the neighbor node degree is fixed as 6 as an assumption.

$$f(\theta) = 10 \cdot \theta - 110 \tag{4.60}$$

If the antenna elements in an array are evenly distributed with a distance $\lambda/2$ between them, i.e. the half wave length, the beam width is approximately

$$\theta = \frac{2}{n-1} \cdot \frac{180}{\pi} \tag{4.61}$$

where n is the antenna number.

Substituting (4.61) into (4.60), the minimum number of scanning becomes

$$N_{scan} = \frac{3600}{(n-1)\cdot\pi} - 110 \tag{4.62}$$

which is plotted in Fig. 4.22.

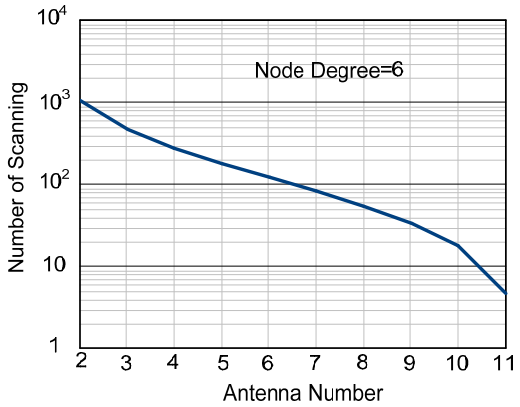


Fig. 4.22 The number of scanning as a function of the antenna number.

Assuming in the worst case all the nodes moving randomly and rapidly, and each communication starting with neighborhood discovery, i.e. the beam scanning, the overall power model including the beam scanning process can be extended as

$$\begin{aligned} P_{DC,tot} &= P_{Rx,DC}\cdot\frac{T_{data}}{T_{tot}} + P_{Tx,DC}\cdot\frac{T_{data}}{T_{tot}} + P_{WuRx,DC}\cdot DC_{WuRx} \\ &+ P_{Tx,DC}\cdot\frac{n_{event}\cdot T_{beacon}}{T_{tot}} + P_{Tx,DC}\cdot\frac{N_{scan}\cdot n_{event}\cdot t_{scan}}{T_{tot}} \end{aligned} \tag{4.63}$$

where $P_{Rx,DC}$, $P_{Tx,DC}$, $P_{WuRx,DC}$ are the peak power consumption of the Rx, Tx and WuRx respectively, T_{data} is the total payload data length in the total observing time T_{tot}, DC_{WuRx} is the optimum duty-cycle factor of the WuRx and n_{event} is the average event number per T_{tot}. Substituting (4.14), (4.15), (4.59) and (4.62) into (4.63), the overall power model is expressed as a function of the antenna number, the application-related factor n_{event}, the average event occurrence cycle T_{dada}/T_{tot} and the physical circuits power.

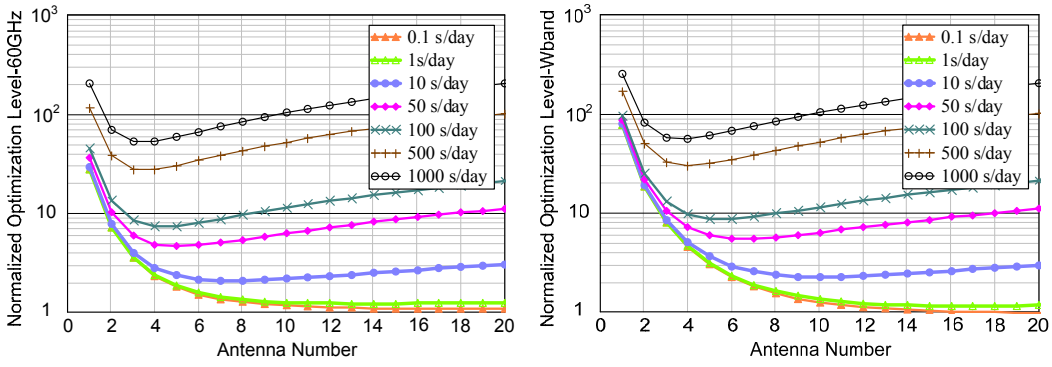
$$\begin{aligned} P_{DC,tot} &= (n\cdot P_{Rx,DC,1})\cdot\frac{T_{data}}{T_{tot}} + (\frac{1}{n^2}\cdot P_{LO,DC,1} + \frac{1}{n^2}\cdot P_{PA,DC,1})\cdot\frac{T_{data}}{T_{tot}} \\ &+ P_{WuRx,DC}\cdot DC_{WuRx} + (\frac{1}{n^2}\cdot P_{LO,DC,1} + \frac{1}{n^2}\cdot P_{PA,DC,1})\cdot\frac{n_{event}\cdot T_{beacon}}{T_{tot}} \\ &+ (\frac{1}{n^2}\cdot P_{LO,DC,1} + \frac{1}{n^2}\cdot P_{PA,DC,1})\cdot\frac{[\frac{3600}{(n-1)\cdot\pi} - 110]\cdot n_{event}\cdot(2\tau\cdot\frac{360}{\frac{2}{n-1}\cdot\frac{180}{\pi}})}{T_{tot}} \end{aligned} \tag{4.64}$$

All the parameters in (4.64) can be optimized independently. In order to reduce the dimension of the function and make the power optimization straightforward, the average event occurrence frequency k is defined as a prerequisite as 1000 times/day in this wireless wire applications, e.g. high-speed downloading or wireless point-to-point communication. Consequently, the power optimization only lies in the phase array configuration (i.e. n and θ) and physical circuitry design (i.e. circuits power, T_{beacon}, $P_{WuRx,DC}$, etc.).

Referring to the analyses in Section 4.1.2, the DCF of the WuRx can be estimated as

$$T_{DC} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot T_{tot}}{P_{Tx,DC} \cdot n_{event}}} \tag{4.6}$$

and the minimum T_{beacon} is identical to T_{DC} +2·T_{wu}. Taking the data from Table 4.2 and the worst-case settling time data from Fig. 4.13 (a) and (b), the average power levels of the 60-GHz and W-band transceivers can then be calculated as a function of the antenna number as well as the average event frequency, as shown in Fig. 4.23 (a) and (b), respectively.



(a) (b)

Fig. 4.23 The normalized optimized average power level as a function of the antenna number and the average event frequency of (a) the 60-GHz system and (b) the W-Band system.

It can be seen that in both cases, the optimum antenna number of the main radio decreases approximately from 11 to 3 when the total event frequency increases from 0.1 s/day (100 Mb at 1 Gbps data rate) to 1000 s/day (1000 Gb at 1 Gbps data rate). The reason is that while the active of the main radio increases, the receiving power becomes more and more dominating. Though the wake-up and the beam scanning powers get reduced with narrower beamwidth, the payload communication power will eventually becomes the dominant factor in the overall power budget, and thus less antenna (active path) number is preferred for the power's concern. In a normal wireless peer-to-peer communication network, an average antenna number, e.g. from 4 to 8, would be a practical choice. This is also helpful to alleviate the frequency selectivity (due to in-band

SNR variation) of a phased array front end. This point will be discussed in details in Chapter 6.

4.4.2 Antenna number discussion considering phase errors

Besides the power consumption of the electronic circuitry, the beam-pointing angle of the array also influences the link model. If the beam-pointing resolution is not sufficiently fine, it may happen that the beams of the Tx and Rx cannot point to each other accurately and thus experience less antenna gain. In the extreme case, there would be a "blind area" where no beam points to. Furthermore, if there is error and mismatch in each path, the final beam-pointing angle is also affected. Consequently, we can either design for the worst-case and left some margin for the receiver SNR, or we need to reserve a sufficiently large beam overlap area to avoid the antenna gain reduction or the pointing-blind area (i.e. avoiding severe SNR degradation in the Rx).

In a multi-element phased array (the reasons and possibilities of using a phased array instead of a true time delay will be specified in Chapter 6) receiver, the beam-pointing angle θ_m can be calculated by [19]

$$\theta_m = \sin^{-1}(\frac{c \cdot \Delta\tau}{d}) \tag{4.65}$$

where c is the speed of light, d is the distance between two antenna elements and $\Delta\tau$ is the required time delay that is calculated as

$$\Delta\tau = \frac{d \cdot \sin\theta_{in}}{c} \tag{4.66}$$

where θ_{in} the incoming beam angle. In a narrow-band phased array transceiver, the phase shifter can be used to replace the true time delay element and it holds that

$$\Delta\phi = \omega_0 \cdot \Delta\tau \tag{4.67}$$

where ω_0 is the center frequency.

Assuming each phase path has a phase error following Gaussian distribution as $N(0, \delta_{path}^2)$ and the overall resultant phase variance becomes $N(0, \delta_{beam}^2)$, δ_{beam} can then be estimated by

$$\delta_{beam}^2 = \frac{12 \cdot \delta_{path}^2}{\pi^2 \cdot \cos\theta_m^2 \cdot (n-1) \cdot n^2 \cdot (n+1)} \tag{4.68}$$

where n is the antenna number. In a 4-element phased array receiver with 85°[25]

[25] It can be see that 90°angle of incidence leads to infinite phase error in this model. However, due to relatively large beamwidth of a 4-element phased array, the worst case scenario is estimated by 85° and 5° phase pointing error will not degrade antenna gain significantly (see Fig. 6.8).

beam-pointing angle, for example, the overall phase error can be plotted in Fig. 4.24.

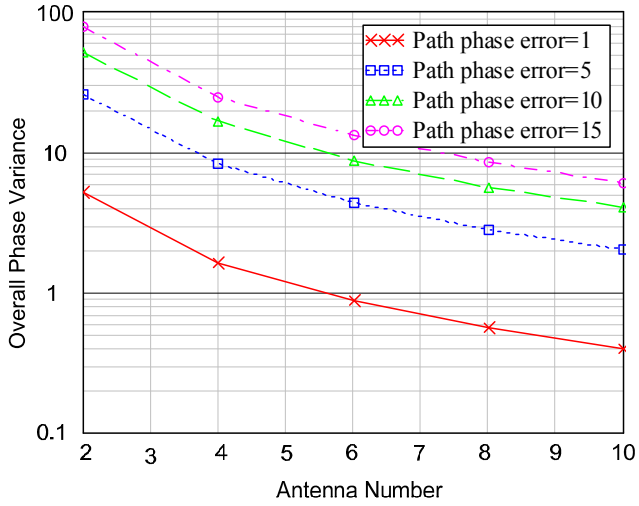


Fig. 4.24 Overall phase error when errors are present in each path.

It can be seen that the overall beam-pointing error is a function of phase error of each path and the number of elements. For example, when each path has a phase error with variance δ_{path} of 5°, the overall phase variance is about 9° for a 4-element phased array. In a 4-element phased array receiver with 38.2° beam width (assuming antenna distance is identical to half wavelength), a 9° phase pointing error will not cause serious SNR degradation and thus is acceptable (more details will be discussed in Chapter 6). It can be seen from Fig. 4.24 that a 10° phase error in each path may lead to a 17° beam pointing error which is identical to half of the beamwidth. In such a case, the antenna gain will be impaired significantly and therefore is not acceptable.

4.5 Conclusions

In this chapter, the complete wireless wire communication system architecture is proposed and its power model, including the wake-up and the beam scanning processes are discussed. An IJLO-based WuRx is proposed and analyzed, which helps to reduce the system average power consumption significantly. The optimum antenna number of the main radio is also derived considering the power consumption of the circuits and the practical beam-pointing issue.

References

[1] Emanuele Lopelli, *Transceiver Architecture and Sub-mW Fast Frequency-Hopping Synthesizers for Ultra-Low Power WSNs*, PhD dissertations, Eindhoven University of Technology, 2010.

[2] Pellerano, S., *et al*., "A mm-Wave Power-Harvesting RFID Tag in 90 nm CMOS," in *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8, pp. 1627-1637, Aug. 2010.

[3] Energy harvesting, www.wikipedia.org, 2011.

[4] Costis Kompis and Simon Aliwell, *Energy Harvesting Technologies to Enable Remote and Wireless Sensing*, in Sensors and Instrumentation KTN report, June 2008.

[5] Lianming Li, et al., "Design and Analysis of a 90 nm mm-Wave Oscillator Using Inductive-Division LC Tank," in *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 44, no. 7, pp. 1950-1958, July, 2010.

[6] Pallotta, A., *et al*., "Millimeter-wave 14dBm CMOS power amplifier with input-output distributed transformers", in *IEEE Custom Integrated Circuits Conference*, 2010.

[7] Peter Baltus, *Minimum Power Design of RF Front Ends*, PhD dissertation, Eindhoven University of Technology, Eindhoven, 2004.

[8] Kenichi Okada *et al*., "A 60GHz 16QAM/8PSK/QPSK/BPSK Direct-Conversion Transceiver for IEEE 802.15.3c," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2011.

[9] Maryam Tabesh *et al*., "A 65nm CMOS 4-Element Sub-34mW/Element 60GHz Phased-Array Transceiver," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2011

[10] Philippe Le-Huy, Sebastien Roy, "Low-Power Wake-up Radio for Wireless Sensor Networks," in *Springer Mobile Networks and Applications*, vol. 15, NO. 2, April 2010, pp. 226-236.

[11] Salvatore Drago *et al.*, "Impulse-Based Scheme for Crystal-Less ULP Radios," in *IEEE Transaction on Circuits and Systems—I: Regular papers*, vol. 56, NO. 5, May 2009.

[12] Dennis Michael Fischette *et al.*, "A 45nm SOI-CMOS Dual-PLL Processor Clock System for Multi-Protocol I/O," in *Proc. 2010 IEEE ISSCC*, San Francisco, 2010.

[13] Changhua Cap, Yanping Ding, Kenneth K. O, "A 50-GHz Phase-Locked Loop in 0.13-μm CMOS," in *IEEE Journal of Solid-State Circuits*, vol. 42, NO. 8, 2007, pp. 1649-1656.

[14] Behzad Razavi, "A 60-GHz CMOS Receiver Front-End," in *IEEE Journal of Solid-State Circuits*, vol. 41, No. 41, pp. 17-22, Jan. 2006.

[15] R. Adler, "A Study of Locking Phenomena in Oscillators," in *proceedings of the IEEE*, vol. 61, no. 10, October, 1973.

[16] J. Borremans *et al.* "A sub-3dB NF Voltage-Sampling Front-End with +18dBm IIP3 and -2dBm Blocker Compression Point," in *IEEE European Solid-State Circuits Conference*, Sep. 2010.

[17] K. Leentvaar and J. Flint, "The Capture Effect in FM Receivers," in *IEEE Transactions on Communications*, pp. vol. 24, no.5, 531-539, 1976.

[18] Xueli An, *Medium Access Control and Network Layer Design for 60 GHz Wireless Personal Area networks*, PhD dissertation, Delft University of Technology, 2010.

[19] Ali M. Niknejad *et al*., *mm-Wave Silicon Technology 60 GHz and Beyond*, Springer, New York, 2008.

CHAPTER 5 MMW LOW-POWER WAKE-UP RECEIVER

One of the most important intentions of a rendezvous scheme is to save power. Intuitively speaking, letting the radio operate incontinuously helps to reduce the average power consumption, but it is limited by application requirements and will affect the communication quality, reliability as well as the efficiency significantly.

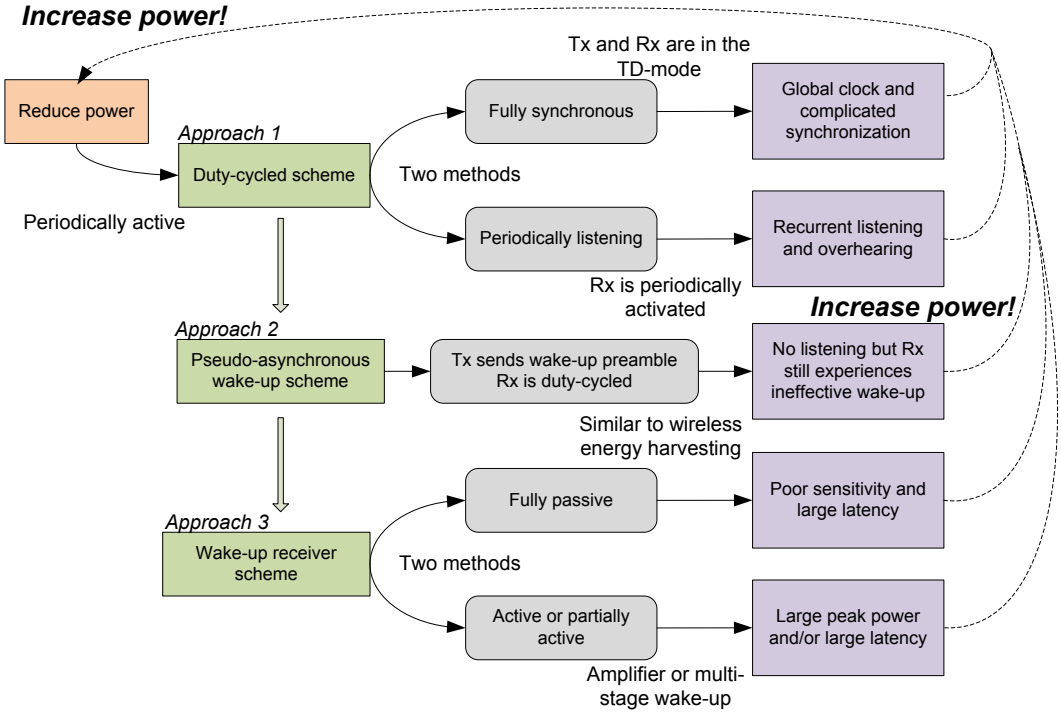


Fig. 5.1 Rendezvous schemes.

For instance, in a duty-cycled power management scheme (i.e. approach 1 in Fig. 5.1), either an accurate clock is required to synchronize the Tx and Rx to operate in the TD-mode, or the periodically listening activity is required for the Rx. In order to reduce the power consumed in the synchronization or the Rx listening process, a pseudo-asynchronous scheme (approach 2) is developed. The Tx becomes event-driven, i.e. it is switched on only if an event is identified. The Rx still works in the duty-cycled pattern. Before sending the payload data, the Tx will send a wake-up beacon to the Rx. As long as the wake-up beacon is sufficiently long to coincide with at least one duty cycle of the Rx, the communication can be built up, and no stringent synchronization process is required. However, the Rx is still activated frequently in the idle period where no data transmission is carried out, and thus a large amount of power is wasted. As a

result, the fully asynchronous wake-up scheme (approach 3) is further proposed, which is believed to be more energy efficient especially when the event frequency is relatively low like in the WPAN applications. In such a case, communications can be monitored and detected by a low-power WuRx, and the Rx of the main radio thus can be always switched off until it is waked up by the WuRx while the Tx is still event-driven.

In the approach 3, however, some potential problems may impair its benefits such as the low average power consumption or the short packet overhead. For example, the stand-by power of the WuRx may not be sufficiently low if it intends to share the same channel with the main radio, i.e. the WuRx operates at high frequencies as well. It may also happen that the wake-up process of the WuRx is too long, which causes high communication latency and eventually will either increase the output power of the Tx (and consequently the Tx power consumption) or decrease the communication range. The trade-offs among the power, range and latency always exist, which makes the choice of the asynchronous wake-up scheme controversial in practice.

As discussed in Chapter 4, a new type of wake-up radio system, i.e. the IJLO-based asynchronous duty-cycled wake-up radio system, is proposed to solve the above problems. Its feasibility will be discussed quantitatively in the following sections.

5.1 General introduction of wake-up receivers

A brief review and comparison of the conventional WuRxs is given in Sections 5.1.1 and 2.4, respectively. In this section, the very general and fundamental classifications, features and trade-offs of the WuRx are summarized, which are able to describe the nature of WuRx architectures and analyze their pros and cons at a higher level. Furthermore, the challenges and possible solutions of the mmW WuRx are discussed in Section 5.1.2 as well.

5.1.1 Classifications and trade-offs

General classifications and trade-offs of the WuRxs are shown in Fig. 5.2. From the perspective of the power sources, WuRxs can be either passive or active, and the potential major trade-off of them is between the sensitivity and the power consumption. Besides, WuRxs can share the same channel with main radio or have a separated channel, e.g. by using a low-frequency WuRx to wake up the high-frequency main Rx. The resulting main trade-off is between the power consumption and the cost. Furthermore, the wake up signal can be a simple wake up tone, or an ID-based wake up bits, and the most important trade-off is between the power (range) and the reliability [1].

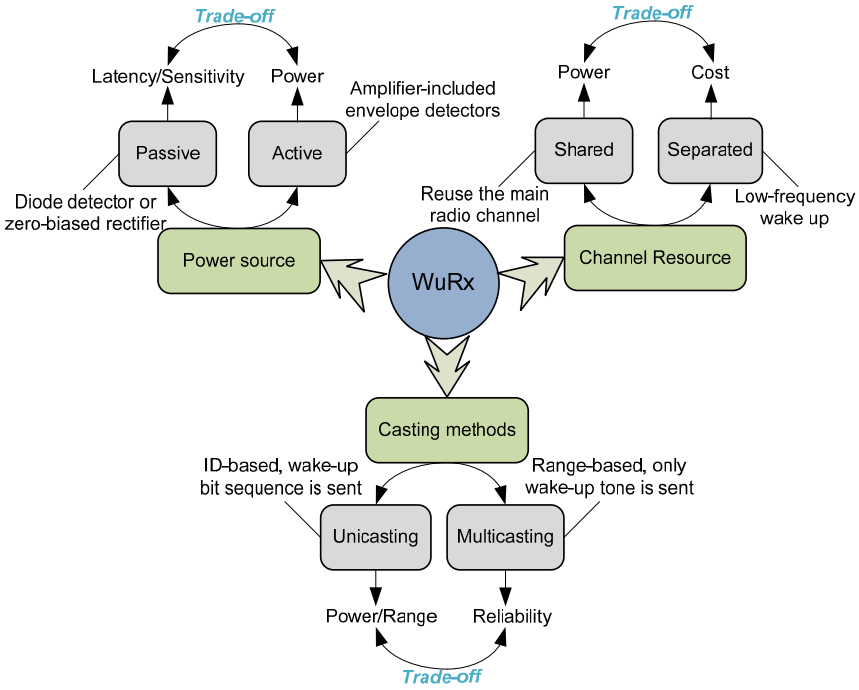


Fig. 5.2 Classifications and trade-offs of WuRxs.

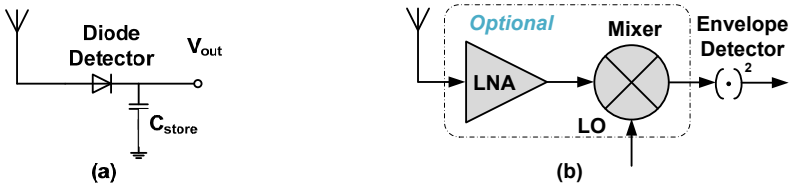


Fig. 5.3 Conceptual models of (a) a passive diode detector and (b) an active envelope detector.

From the power source perspective, the WuRxs can be classified as passive and (partially) active. Passive WuRxs like the diode detector shown in Fig. 5.3 (a) are widely used because of their low power consumption, but their sensitivity is normally poor due to the square law of diodes, i.e. the amplitude of the DC output signal of a diode is proportional to the square of the amplitude of the input signal. Some types of diodes (e.g. the Schottky or Germanium diodes) may have better sensitivity than others like the Silicon diodes due to their higher V-I transform efficiency, but the square law always exists. The only difference is how fast the slope would be. In [2], a radio-triggered WuRx is proposed which is able to improve the diode sensitivity by using a multi-stage rectifier to accumulate the charges to reach the threshold voltage. However, the latency of such circuits, i.e. the time consumed for the charge accumulation is very large. For example, in order to reach a 0.5-V threshold voltage (ΔV) of a CMOS switch, assuming the capacitor used to store energy is 10 fF (minimum capacitance) and the output current of the diode

(I_{out}) is about 1.7 pA[26], the total time consumed for a 5-stage change pump (ΔT) equals to

$$\Delta T = \frac{5 \cdot 0.1 \cdot 10 \cdot 10^{-15}}{1.7 \cdot 10^{-12}} \approx 3\ ms \tag{5.1}$$

In a high-speed WPAN system, e.g. sending a 1-Mb data packet with a 1-Gbps data rate takes about 1 ms, but according to (5.1), the wake up process takes 3-times longer time. This is not a very efficient design. Moreover, in order to achieve a better sensitivity of the WuRx, the input voltage of the diode will be much smaller than the value we have assumed. Therefore, the total wake up time will become even longer.

In contrast to the passive detectors, the active WuRxs (as illustrated in Fig, 5.3 (b)) normally have a relatively better sensitivity and a reasonable latency, but their standby power may cause problems especially when the system works at high frequencies, which may impair the system mobility. Besides the high standby power, the sensitivity of the active WuRxs is also quite limited if using the self-mixing envelop detection method, because in such circuits, the square law between the input and output signals still exist.

From the perspective of the channel resources, WuRxs can share the same communication channel with the main radio or use a separate (normally lower frequency channels) for the wake-up activity. The former can reduce the hardware complexity and the cost. For example, it requires less chip area as well as smaller antennas. However, the standby power of the WuRxs at higher frequencies always becomes the bottleneck during implementation. This issue will be discussed in the next section. Separating the wake-up channel from the main communication channel helps to reduce the design difficulty of the WuRx hardware, but it may add up much more system cost as well as the complexity of the data link protocol. Besides that, a low-frequency wake-up scheme (e.g. wake-up a 60-GHz system by a 2.4-GHz WuRx) may produce interference to other wireless systems and cause co-existence problems with other standards.

5.1.2 Solutions for the mmW wake-up receivers

As discussed, in the mmW WPAN application scenario (which is the focus of this thesis), the system requirements on the communication latency and the Rx sensitivity is quite high due to the features such as a high communication speed (i.e. requiring minimum packet overhead) and a high pathloss channel (incapable of providing large power from the Tx to the Rx). Consequently, the conventional passive wake-up schemes become not

[26] The saturation current of silicon diodes is assumed to be 10 pA. Besides, a 25 mV input voltage is used to calculate the output current of the diode. Consequently, the sensitivity of such a diode WuRx is poor.

applicable under current technology levels. For example, in [3], a 2.4-GHz nW-level passive detector is reported. However, its sensitivity is only -28 dBm, which is not feasible, because most of the wireless communications in the WPAN scenario are P2P and no high-power transmitting signal can be obtained like in backscattering systems. As a result, the WuRx has to be highly sensitive in the power constrained situation. Recently, different types of active WuRxs are reported, which are able to achieve about -70 dBm sensitivity by consuming about 50 μW power [4] [5], but none of those WuRx can be used in the mmW range without affecting their power performance, because both of them are based on the conventional envelope detection method and thus suffer from the square law. As a result, the trade-offs among the sensitivity, latency and power become ultimate important in the mmW application scenario. Based on the above discussion and the relevant analyses in Chapter 4, the design challenges and possible solutions of the mmW high-sensitivity, low-latency and low-power WuRxs are summarized in Table 5.1.

Table 5.1 Challenges and possible solutions of the mmW WuRxs

Challenges	**Possible Solutions**
Cost and complexity	-Single chip solution -Shared channel
Latency	-Active wake up -Fast-settling circuitry
Standby power	-Asynchronous duty-cycled wake-up scheme
Connection difficulty [1]	-Dynamic adjustment of the wake-up beacon length or the output power -Efficient MAC protocol (e.g. channel sensing, collision avoidance, localization, etc.) -Machine-learning network

5.2 Technology for mmW low-power applications

There are several drivers for wireless communication systems like cost, power consumption, frequency bands, functionality, and the size of mobile units, and standards and protocols. Compared to the analog and digital technologies, normally more compromises are required for the RF technologies because more competing requirements of a RF system should be achieved simultaneously [6]. This results in very stringent requirements on technologies, especially for the mmW wireless communication systems whose operation frequencies may approach the f_T of a certain technology. In [6], it is reported that for particular mmW applications, the Si-based technologies have become the leading technology in the last decade by advantages in cost and integration levels. This is particularly true for the emerging mmW low-power and low-cost communication. For performance driven applications or high-performance power amplifiers, the market is

still dominated by the heterogeneous III-V compounds like GaAs or InP. As a result, for our applications, the Si-based technologies become the proper choice with respect to the compromise of cost, power and performance. Furthermore, due to the decent performance at the mmW range, better integration level and lower cost (per unit), the deep-submicron CMOS technology is chosen to implement our system instead of the BiCMOS technology. This is also important for the long-term concern, because the technology development of the BiCMOS technology lags about three to four generations behind the CMOS technology due to the lack of product drivers in the market as presented in [6]. In this sense, the advanced CMOS technology is more competitive. The f_T and NF_{min} is simulated as a function of current density of the TSMC 65-nm LVT RF transistor, as shown in Fig. 5.4. It can be seen that with about 0.15 mA/μm bias current density, the f_T is about 120 GHz and NF_{min} is about 3 dB, which is acceptable for our mmW system.

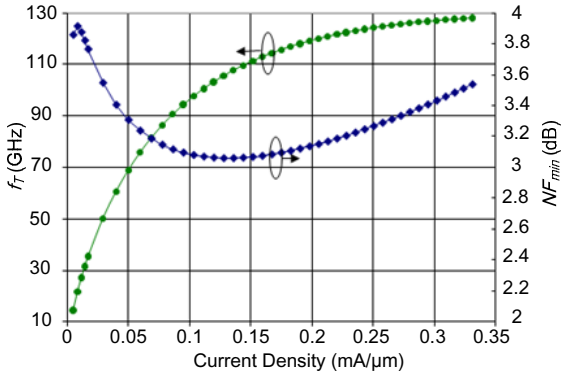
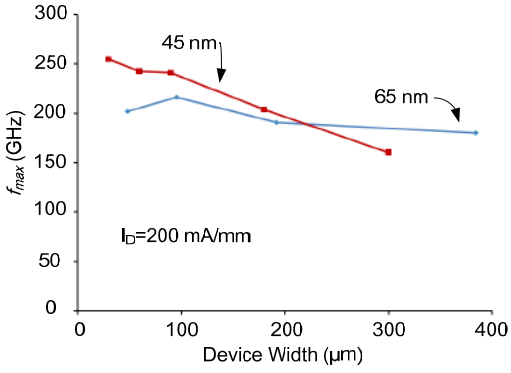
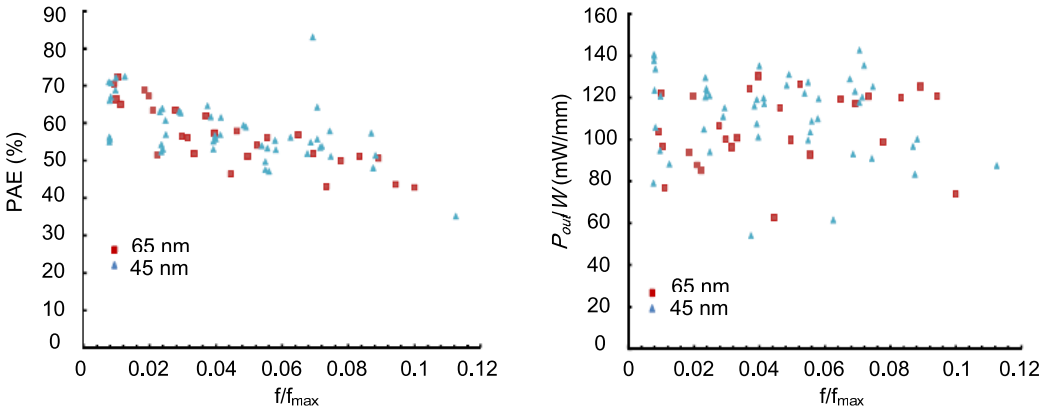


Fig. 5.4 The simulated f_T and NF_{min} curves of the TSMC 65-nm LVT NMOS RF transistor.

In addition, the performance of 45-nm and 65-nm CMOS technologies is compared in [7] with regard to the maximum oscillation frequency (f_{max}), the power-added efficiency (PAE) and the output power per unit width (P_{out}/W), as shown in Fig. 5.5 (a), (b) and (c). Although the results are particularly for the IBM CMOS technology, it also provides insight about comparable technologies from other foundries.



(a) f_{max}



(b) PAE (c) P_{out}/W

Fig. 5.5 The performance comparison of 45-nm and 65-nm technologies.

It can be seen that the RF power performance of the 45-nm CMOS technology only slightly outperforms the 65-nm technology. Since the integration level of the digital block is not yet included within the scope of this work, the 65-nm CMOS technology will be used to implement out mmW system.

5.3 Wake-up receiver circuit design

The design considerations, methodology and relevant discussion of the mmW WuRx are presented in the section. The WuRx architecture is designed and verified by using the ADS™ system simulation tools. The IJLO, mixer, buffer amplifier and the entire WuRx circuits are designed and implemented in the TSMC 65-nm CMOS technology. Measurement results will be shown followed by discussions and recommendations.

5.3.1 Link budget and system specifications

In order to reduce the overall cost as well as the system complexity, the WuRx is designed to re-use the main radio communication channel. Therefore, the wake-up beacon will be sent by the Tx of the main radio, which is a 4-element phased array Tx with a resulting antenna gain of 6 dBi. Assuming a 1-meter hopping distance existing between two communication nodes (as discussed in Section 3.4.2), the link budget parameters of the 60-GHz and 70-GHz systems are summarized in Table 5.2. Taking the 60-GHz link as an example, with a 0-dBm output power of each path of the Tx, 1 meter distance and 6 dBi antenna gain, the maximum power from the Tx to the WuRx is about -56 dBm. The Rx sensitivity (S) can be estimated as the summation of system noise floor (KT multiplied with noise bandwidth), noise figure, SNR and design margin as

$$S = KTB + NF + SNR + Margin \tag{5.2}$$

Equating *S* to -56 dBm, the system margin can be calculated as about 8 dB (with 100-MHz noise bandwidth) when only the wake-up tone is sent, i.e. no demodulation is applied to the wake-up signal.

Table 5.2 Link budget parameters of the 60-GHz and 70-GHz channels

	60 GHz Link (WPAN)	**70 GHz Link (W-band)**
Wake-beacon power from the Tx (single-path)	0 dBm 6 dBm for 4 paths	0 dBm 6 dBm for 4 paths
Antenna gain of the Tx (4-element antenna array)	6 dBi	6 dBi
Path loss over 1 meter (Line-of-Sight)	-68 dB	~-69 dB
WuRx antenna gain	0 dBi	0 dBi
Noise floor (T=300 K)	-174 dBm/Hz	-174 dBm/Hz
Noise bandwidth	100MHz for wake up tone 1 GHz for wake up bits	100MHz for wake up tone 1 GHz for wake up bits
Bandwidth of the IJLO input	<100 MHz	<100 MHz
WuRx noise figure	20 dB	20 dB
Required SNR	~10 dB (wake-up tone)[27]	~10 dB (wake-up tone)
	10 dB (OOK wake-up bits with 1-Gbps data rate)	10 dB (OOK wake-up bits with 1-Gbps data rate)
System margin	~8 dB (wake-up tone)	~7 dB (wake-up tone)
	LNA is required for OOK-modulated wake-up bits or alternatively, we can reduce its data rate	LNA is required for OOK-modulated wake-up bits or alternatively, we can reduce its data rate

It should be noted that for the ID-based wake-up method, the system cannot fulfill the link requirement unless a low-noise amplifier (LNA) is added in front of the WuRx circuitry. The total noise figure of the WuRx can be reduced significantly if the gain of the LNA is sufficiently large. It will increase the peak power consumption of the WuRx as a trade-off. However, with our proposed asynchronous duty-cycled wake-up method,

[27] There is an important trade-off between the noise bandwidth and the detection time of our proposed WuRx. With a pure wake-up tone, the noise bandwidth can be reduced to 100 MHz (or even smaller) because in this situation, the WuRx has better tolerance to high detection time. In an ID-based wake up scheme, however, a certain bandwidth has to be maintained in order to support the high data rate of the wake-up bits. As a result, the fast-detection ability is required for the WuRx. Please refer to Appendix C for more details.

the average power consumption of the WuRx will not be increased too much. Alternatively, we can reduce the data rate of wake up bits. In the link budget calculation, 1-Gbps data rate is used, which is identical to the communication data rate of the main radio. However, a lower data rate (e.g. 100 Mbps) can be adopted particularly for the wake up process and therefore the required SNR of the WuRx can be significantly reduced. Besides, since the data rate of wake up bits is lower, the WuRx has a better tolerance to longer detection time and the noise bandwidth of it can be reduced as well. This will help to relax the link budget further. With a larger link margin, we could either extend the communication range or have a better tolerance to the SNR degradation due to e.g. the frequency selectivity of the phased array front end or the beam pointing errors.

According to Table 5.2, the theoretical minimum detectable signal of the mmW WuRxs, i.e. the sensitivity, is about -60 dBm. In order to make sure our proposed WuRx is sufficiently competitive to other state-of-the-art active WuRxs like [4] and [5], and is compatible to the battery or energy harvesting based WPAN applications, the maximum acceptable average power dissipation is set as 50 μW as the value achieved in [4] and [5] (at much lower frequencies). Referring to (3.55), it can be further interpreted as

$$P_{WuRx,DC} \cdot \sqrt{\frac{n_{event} \cdot k \cdot T_{wu}}{2 \cdot T_{tot}}} \leq 50\ \mu W \tag{5.3}$$

Taking n_{event} as 1000 times/day, k as 100 and T_{tot} as 1 day (86400 seconds), (5.3) can be re-written as

$$T_{wu} < \left(\frac{70\ \mu W}{P_{WuRx,DC}}\right)^2\ s \tag{5.4}$$

Assuming a 10-mW P_{WuRx}, the settling time (T_{wu})[28] of the WuRx should be no longer than about 50 μs. The specifications of the mmW WuRxs are summarized in Table 5.3.

Table 5.3 Specifications for mmW WuRxs

Parameters	**60 GHz Link (WPAN)**	**70 GHz Link (W-band)**
System center frequency	60 GHz	73 GHz
Bandwidth (Locking/tuning range)	>7 GHz (57 to 64 GHz)	>6 GHz (73 to 79 GHz)
Wake-up scheme	Asynchronous duty-cycled wake up scheme	Asynchronous duty-cycled wake up scheme

[28] It should be noted that the settling time and the detection time are different parameters of a WuRx. The former indicates the time used to identify the wake up instruction, and the latter is the time consumed to generate a stable output signal to turn on the Rx of the main radio.

Architecture	IJLO-based self-mixing envelope detector	IJLO-based self-mixing envelope detector
Modulation	None Pure wake-up tone	None Pure wake-up tone
Data rate	1-Gbps for the main radio	1-Gbps for the main radio
Sensitivity	-60 dBm	-60 dBm
Average power constraint	<50 μW	<50 μW
Wake-up time	<50 μs	<50 μs

5.3.2 WuRx architecture and brief system design review

The basic theory on the IJLO-based asynchronous duty-cycled WuRx has been discussed in Section 4.3 where the nature of the injection-locking phenomenon is explained, and the key parameters like the locking range and the settling (locking) time are calculated based on the conceptual model of the IJLO in Section 4.3.1. It is proved that the locking range of an LC oscillator based IJLO can be estimated by

$$\omega_L = \frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}} \tag{4.32}$$

under a weak injection level and the settling time of the IJLO is thus calculated as

$$t_{settling} = \frac{1}{\sqrt{\omega_L^2 - (\omega_0 - \omega_{inj})^2}} \cdot \ln\left(\frac{\omega_L + \sqrt{\omega_L^2 - (\omega_0 - \omega_{inj})^2} - (\omega_0 - \omega_{inj}) \cdot \tan\left(\frac{\arcsin\left(\frac{\omega_0 - \omega_{inj}}{\omega_L}\right)}{2}\right)}{\omega_L - \sqrt{\omega_L^2 - (\omega_0 - \omega_{inj})^2} - (\omega_0 - \omega_{inj}) \cdot \tan\left(\frac{\arcsin\left(\frac{\omega_0 - \omega_{inj}}{\omega_L}\right)}{2}\right)}\right) \tag{4.37}$$

which is a function of the frequency difference between the center and injection frequencies, the loaded quality factor of the resonator and the oscillation and injection current levels. From the above theory, the IJLO-based WuRx architecture is proposed as illustrated in Fig. 5.6.

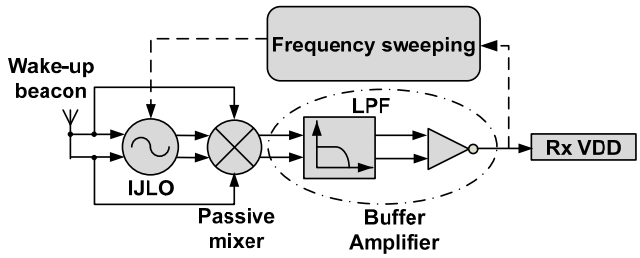
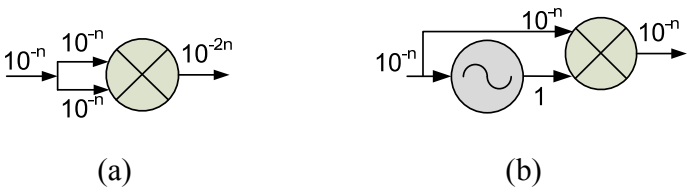


Fig. 5.6 The IJLO-based WuRx.

The wake up beacon is split into two branches after received by the antenna. One is used to trigger the IJLO, where the frequency information will be extracted. The other path is connected directly with the passive mixer as the RF input. When the IJLO gets locked, it will generate a large constant-envelope output voltage with exactly the same frequency as the input RF signal. This output voltage acts as the LO driver of the passive mixer. At the output of the passive mixer, a high DC voltage will appear at the IJLO locking condition and it will be used to switch on the main radio. When there is no RF input or the IJLO is not correctly locked, its natural frequency will be different from the received RF signal, and thus the output of the passive mixer will be a low voltage at DC and therefore the main radio Rx is kept staying at the "off" mode.

With the IJLO, the quadratic behavior or the square law of the self-mixing process is replaced by a linear relationship between the input and output signals. Therefore, the sensitivity of the WuRx is improved significantly, as illustrated in Fig. 5.7 (a) and (b). Assuming a weak input signal level with an amplitude on the order of 10^{-n} (e.g. $3.2\cdot10^{-4}$ V peak voltage referring to a 50-Ohm antenna with a -60 dBm input power), the direct multiplication results in an output signal level on the order of 10^{-2n} (the square of the input), as shown in Fig. 5.7 (a). On the contrary, in Fig. 5.7 (b), by inserting the IJLO into one of the path, the square law is eliminated due to the fact that the output level of the IJLO is always constant envelope and it is independent of the input level under the weak injection assumption.



(a) (b)

Fig. 5.7 Conceptual models of (a) the conventional envelope detector and (b) the IJLO-based linear self-mixing detector.

The frequency-sweeping locking method, as described in Section 4.3.2, is further proposed to extend the effective locking range of the IJLO. The center frequency of the free-running IJLO is kept on being swept with a certain frequency step (by tuning the varactor values, for example) over the entire system bandwidth until the IJLO gets locked to the input RF signal. In this way, the trade-off between the injection level and the locking range, as shown in (4.32) is converted into the trade-off between the total sweeping (locking) range and the sweep time (circuit latency). Consequently, a very high sensitivity level and wide effective locking range (i.e. total sweep range) of the IJLO can be achieved simultaneously.

5.3.3 Design of the injection-locked oscillator

The IJLO, in the first place, should contain a free-running oscillator core, and added to that an extra injection path for the phase correction and lock process. In this section, the design considerations and methodology of the IJLO are presented. The detailed circuit design and implementation of the mmW IJLO will be presented afterwards followed with the measurement results.

Choice of the oscillator core

Different types of oscillators can be used to implement the oscillator core of an IJLO. They can be grouped as the non-resonating oscillators (e.g. ring oscillator) and the resonant oscillators (e.g. cross-coupled LC oscillator, Colpitts or Hartley oscillators, distributed oscillators, etc.).

The ring oscillator is featured with low-power and small chip size properties, but its phase noise performance is relatively poor (at the same current consumption level compared to LC oscillators) due to its inductor-less low-Q nature [8]. However, as it is known, lower Q leads to wider bandwidth, which means it may have a benefit of a larger locking range. The locking range of a conventional ring oscillator can be calculated from the generalized Adler's equation as [9]

$$\frac{d\theta}{dt} = \omega_0 - \omega_{inj} - \omega_0 \cdot g(\theta) \tag{5.5}$$

where ω_0 and ω_{inj} are the natural and injection frequencies, respectively, and $g(\theta)$ is the averaging over the fast varying variables. With a sinusoidal injection signal, $g(\theta)$ can be calculated as [10]

$$g(\theta) = \frac{1}{\sqrt{4\pi^2 + K_0^{\,2}}} \cdot \frac{I_{inj}}{I_{osc}} \cdot \sin(\theta + \zeta) \cdot [K_1 \cdot (e^{\frac{K_0}{2}} + 1) - K_2 \cdot (e^{K_0} + e^{\frac{K_0}{2}})] \tag{5.6}$$

where K_0 equals 2.887, K_1 equals 0.447, K_2 equals -0.1 and

$$\sin(\zeta) = \frac{2\pi}{\sqrt{4\pi^2 + K_0^{\,2}}} \tag{5.7}$$

It can be obtained from (5.5) to (5.7) that, when the ring oscillator gets locked and reaches its steady state, the maximum acceptable frequency difference between ω_0 and ω_{inj} equals the maximum value of $\omega_0 \cdot g(\theta)_{\max}$, i.e.

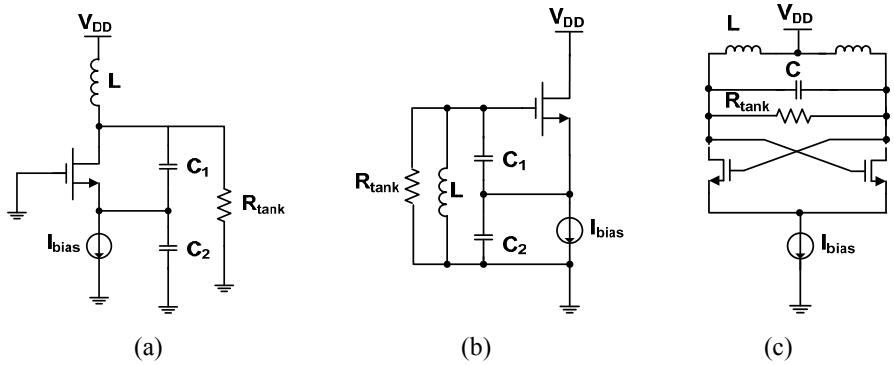
$$\left|\omega_0 - \omega_{inj}\right|_{\max} = \frac{\omega_0}{\sqrt{4\pi^2 + K_0^{\,2}}} \cdot \frac{I_{inj}}{I_{osc}} \cdot [K_1 \cdot (e^{\frac{K_0}{2}} + 1) - K_2 \cdot (e^{K_0} + e^{\frac{K_0}{2}})] \tag{5.8}$$

and thus the locking range ω_L becomes

$$\omega_L = \frac{2\cdot\omega_0}{\sqrt{4\pi^2 + K_0^{\,2}}}\cdot\frac{I_{inj}}{I_{osc}}\cdot[K_1\cdot(e^{\frac{K_0}{2}}+1) - K_2\cdot(e^{K_0}+e^{\frac{K_0}{2}})] = 0.6773\cdot\omega_0\cdot\frac{I_{inj}}{I_{osc}}\Big|_{sinusoidal} \quad (5.9)$$

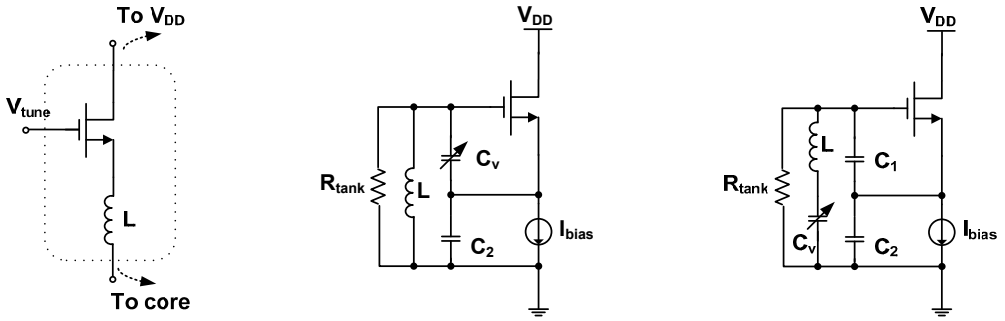
Compared to (4.32), the locking range of a ring oscillator can be theoretically 6 times larger than an LC oscillator when the loaded Q is around 10. Unfortunately, the ring oscillator cannot operate at very high frequencies like the mmW bands at current technology level due to the limited switching speed of MOS gates. However, along with the down-scaling of IC processes as well as the technology improvement, GHz-level ring oscillator may become feasible for low-power mmW applications in the future. Comparing to the ring oscillator, the LC oscillator has an inherent drawback at higher frequencies, i.e. its low tank impedance, which will be limited by $Q_C/(\omega_0 \cdot C)$ (where Q_C is the quality factor of tank capacitor, ω_0 is the resonance frequency and C is the tank capacitance) and can hardly get increased further (if using on-chip passive devices only) when the resonance frequency becomes higher.

As for the resonating oscillators, the Colpitts oscillators (similar principle to that of the Hartley oscillators) and the cross-coupled LC oscillators are widely used in high frequency applications, as shown in Fig. 5.8 (a), (b) and (c). However, the frequency tuning of a Colpitts oscillator is relatively difficult and inexplicit compared to the cross-coupled LC oscillator especially when implemented at very high frequencies like the mmW frequencies. The tuning mechanism can be realized either by using an active inductor, or replacing C_1 or C_2 with a variable capacitor, or by adding a variable capacitor in series with the inductor L (known as the Clapp oscillator), as shown in Fig. 5.9 (a), (b) and (c), respectively.



(a) (b) (c)

Fig. 5.8 Conceptual models of (a) a common-gate Colpitts oscillator, (b) a common-source Colpitts oscillator and (c) a cross-coupled LC oscillator.



(a) (b) (c)

Fig. 5.9 (a) The active inductor tuning, (b) variable capacitor tuning and (c) Clapp oscillator scheme.

In the active inductor tuning scheme, the quality factor of the active inductor path achieves its maximum value only when the tuning transistor is fully on and off, i.e. in the deep-triode region or being completely turned off. A poor quality factor in between of those two states will cause phase noise degradation and very large output amplitude fluctuation, which will impair the frequency sweeping method of our IJLO. Although the non-linear tuning problem can be pre-determined and calibrated in advance, yet, it will introduce design complexity and inaccuracy to the digital control parts. In addition, the current consumption is also an issue when the quality factor is degraded.

The problem of oscillators shown in Fig. 5.9 (b) and (c) is their reduced tuning range. The center frequencies of them can be calculated as

$$f_{0,Colpitts} = \frac{1}{2\pi \cdot \sqrt{L \cdot (\frac{1}{C_v} + \frac{1}{C_2})}} \tag{5.10}$$

$$f_{0,Clapp} = \frac{1}{2\pi} \cdot \sqrt{\frac{1}{L} \cdot (\frac{1}{C_v} + \frac{1}{C_1} + \frac{1}{C_2})} \tag{5.11}$$

Despite the influence of the fixed parasitic capacitance, the center frequency of a cross-coupled LC oscillator is

$$f_{0,cross-coupled} = \frac{1}{2\pi\sqrt{LC_v}} \tag{5.12}$$

Supposing equal values of C_1, C_2 and C_v for calculation simplicity, the center frequency varies by 13.4%, 8.71% and 29.3% in (5.10), (5.11) and (5.12), respectively, when the value of C_v is doubled. The performance comparison among three oscillators is summarized in Table 5.4.

Table 5.4 Performance comparison of three types of oscillators

	Output voltage	**Frequency**	**Tuning Range**	**Phase Noise**	**Chip Area**
Ring oscillator	-	--	+	-	++
Colpitts	+	+	+	++	-
Cross-coupled	+	++	++	+	-

Based on the above discussion, the cross-coupled LC oscillator is chosen as the oscillating core of the IJLO in this work.

Injection signal types and injection positions

Besides the current injection directly to the LC tank as mentioned in Section 4.3.1, multiple other injection approaches are also available. For instance, the injection can be in the voltage mode, current mode or the combined mode, and the injection position can be the LC tank or the biasing path (e.g. the injection signal is fed in through the tail current path) [11], as shown in Fig. 5.10 (a), (b) and (c), respectively.

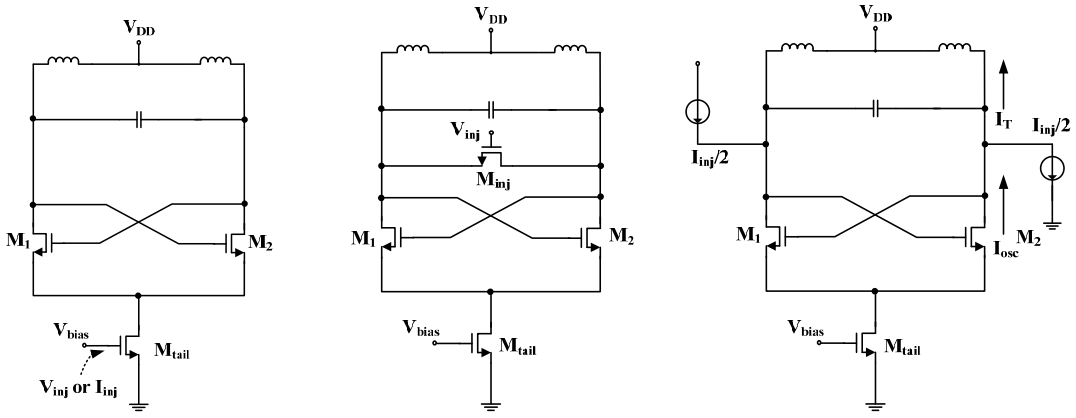


Fig. 5.10 (a) Tail injection in voltage or current mode; (b) direct voltage mode injection and (c) direct current mode injection.

The tail injection method is commonly seen in the frequency divider circuit. It can be implemented by either injecting the voltage signal to the gate of the bottom biasing transistor, i.e. by re-using the biasing transistor as a g_m input stage, or using a current mirror to feed the injection current directly to the source node of the cross-coupled core

transistors. The tail injection method does not load the oscillator and thus does not add extra parasitic capacitance to the LC tank. Therefore, it is beneficial in the sense of maximizing the tuning range of the oscillator. However, this method is very inefficient especially when applied at very high frequencies. For example, in the voltage injection scheme (Fig. 10 (a)), V_{inj} is first converted into a current signal in the tail transistor M_{tail} and then fed into the core transistors M_1 and M_2. In order to deliver a sufficiently large tail current and eliminate the channel length modulation effect of the deep submicron CMOS transistor, the size of M_{tail} has to be chosen with deliberately large value of its length (e.g. W/L=64μm/180nm in our work). As a result, a large amount of injection power leaks away through the miller capacitor or the drain-bulk/source-bulk capacitors, which results in much smaller effective injection power and therefore reduces the locking range. A similar situation occurs when the current signal flows through M_1 and M_2 to the resonator. Due to the high parasitic level, the current gain of M_1 and M_2 is much less than unity, especially at high frequencies. Therefore, the injected signal losses its power again there. Besides the loss of power, it is also not possible to achieve symmetrical injection in the tail injection topology, which will case asymmetrical IJLO outputs and affects the next stage, e.g. by introducing a common-mode signal in the LO driver and thus reducing the conversion gain.

The same problem also happens in the topology shown in Fig. 5.10 (b), which is the most commonly seen direct injection method as reported in [12] and [13]. In order to reduce the R_{on}, the size of M_{inj} has to be relatively large (not as large as M_{tail}, but large compared to M_1 and M_2), and it will introduce considerable parasitic capacitance like C_{gs} and C_{gd} and thus load the LC tank significantly. The shunt inductor peaking method is therefore adopted, for example, in [14], to solve this problem. However, it will increase the chip area and cause coupling problem as well. Consequently, a fully differential direct current-mode injection is chosen in our design as illustrated in Fig. 5.10 (c). In this topology, a differential input buffer is required to provide the current-mode injection signal. In fact, a low-noise amplifier would inherently be an optimum choice, which is able to function as an isolator to the antenna too. The detailed design considerations are presented in the following sections.

I/O buffers and their influence

As discussed, the input buffer of the IJLO is very import. Its main functions can be summarized as: (i) to provide the differential current-mode injection signal with as less loading effect as possible, (ii) to isolate the IJLO from the antenna, (iii) to amplify the injection current and thus to increase the locking range and (iv) to provide the input matching. As a result, the required features of the buffer are as following: (i) it should

have large output impedance and do not load the IJLO; (ii) the buffer circuit should have sufficiently low S_{12} (reverse isolation); (iii) it should have some current gain and (iv) its input impedance should be matched to the designed antenna impedance. Besides these requirements, its noise factor should also be optimized at the operation frequency to avoid additional input-referred noise. In the author's 60-GHz IJLO design [15], a differential current re-use cascode buffer is used to increase the current gain [16], as shown in Fig. 5.11 (a) and (b) where the "load" in the figure means the LC tank of the IJLO. By adding a series inductor L at the gate of the upper transistor, the drain node of MOS transistor M_1 will be coupled to the gate of M_2. As a result, the drain current of M_2 will hold a relation as

$$i_{d2} = g_{m2} \cdot v_{gs2} = \frac{g_{m2} \cdot i_{d1}}{j\omega \cdot C_{gs2}} = \frac{\frac{g_{m2}}{C_{gs2}}}{\omega} \cdot \frac{i_{d1}}{j} = \frac{\omega_{T2}}{\omega} \cdot \frac{i_{d1}}{j} \tag{5.13}$$

where i_{d1} and i_{d2} are the drain currents of M_1 and M_2, respectively, g_{m2} is the transconductance of M_2, v_{gs2} is the gate-source voltage of M_2, C_{gs2} is the gate-source capacitor, ω is the operating angular frequency, and ω_{T2} is the angular cut-off frequency of M_2. In other words, the cascode stage is then converted into a cascaded topology. The current gain is theoretically increased while the DC current is reused by these two stages and stays unchanged. For instance, ω_{T2}/ω can reach 2 in the 65-nm CMOS technology.

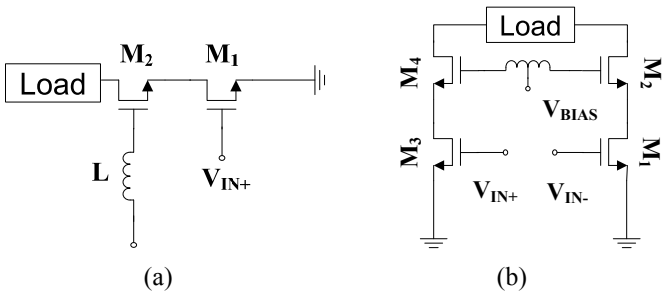


Fig. 5.11 (a) The current-reuse cascode input buffer and (b) its differential version

Though the current gain is improved in this way, the inverse isolation performance becomes impaired as a penalty, i.e. the simulated S_{12} is only about -15 dB in the 60 GHz band, which may cause a reverse leaking of the signal from the IJLO to the antenna in practice. As a result, a simplified differential cascode buffer is adopted in our second design, i.e. the 70-GHz IJLO. Due to the extra phase introduced by the large parasitic capacitor at high frequencies, the cascode structure is less effective in the sense of improving the unilateralization feature of the circuit. Yet, the simulated S_{12} of the buffer is decreased to about -25 dB, which is much better than the cascade buffer. The reverse isolation can be further improved by shunting an extra inductor with the source capacitor

of M_1 as well as the drain capacitor of M_2. The relevant theory and design details will be presented in Chapter 6, the LNA design section.

Although only needed for the measurement purpose, the output buffer should also be carefully designed in order not to load the LC tank seriously. Since the buffer is connected to the LC tank directly, its input capacitance will be merged into the tank. The direct influence is naturally the reduction the resonance frequency. In addition, since it is a fixed capacitance, it will degrade the tuning range of the oscillator as well. Likewise, the resistive input impedance of the buffer will impair the quality factor of the resonator significantly and thus reduce the output power of the oscillator core while increasing the phase noise. These effects can hardly be de-embedded during measurement, and thus should be avoided to the maximum extend. Several widely used buffers are listed in Fig. 5.12. It should be noted that the single-ended versions are shown but in fact, the differential versions are required for the symmetrical loading effect of the IJLO.

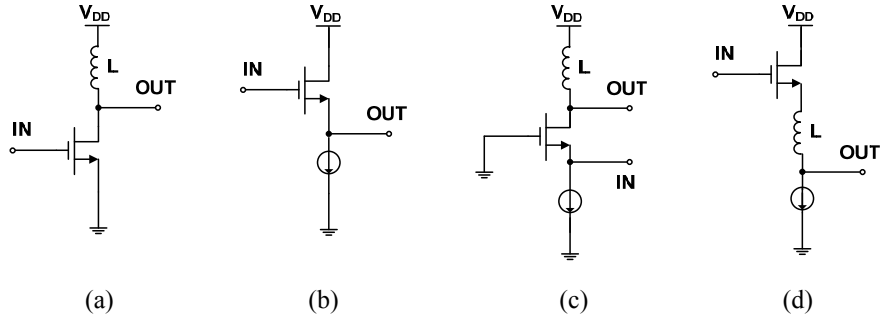


Fig. 5.12 (a) The common-source buffer, (b) the common-drain buffer, (c) the common-gate buffer and (d) the inductor-enhanced common-drain buffer.

These buffers are simulated in combination with our 70-GHz oscillator core using the TSMC 65-nm CMOS models (the LVT RF transistor models and the passive models with built-in parasitics) as well as the parasitic values obtained from the RC-extraction results of the interconnections and metal wires. The simulation results are shown in Table 5.5.

Table 5.5 The comparison of output buffers for the 70-GHz IJLO

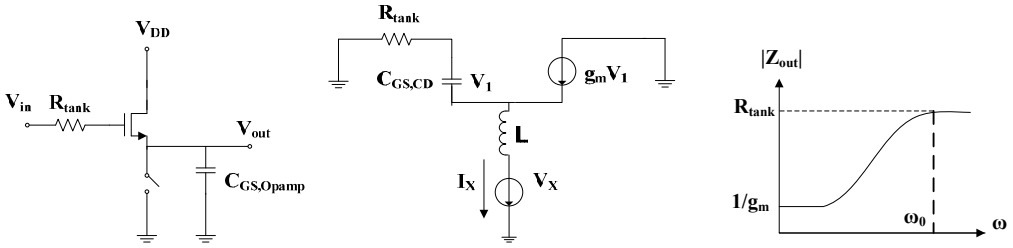
Output Buffer Type	Quality factor	**Capacitive load**	**Power Gain (dB)--unmatched**
Common source	3	12 fF	-3
Common drain	3	8 fF	-5
Common gate	0.2	-	-6
Common drain with inductive peaking	4	5 fF	0

*All the buffers are normalized with a 3-mA bias current from a 1-V supply.

It can be seen that the common-drain buffer with series inductor peaking outperforms other buffers as for the lowest input capacitance and the minimum input resistance, i.e. the highest Q. In this topology, the gate-source capacitance of the common-drain transistor (source follower) is resonated out by the series inductor inserted between the source terminals of the upper transistor and the current source. The capacitive load of the tank is thus minimized while no extra DC consumption is required [17]. Besides, when driven by a relatively large impedance R_S (compared to $1/g_m$), the common-drain buffer shows an inductor-like output behavior [18], i.e. when ω_0 is 0, $Z_{out}=1/g_m$. On the other hand, when ω_0 is identical to the resonance frequency (due to the fact that $C_{GS,CD}$ is resonated out by L), $Z_{out}=R_S=R_{tank}$, as shown in Fig. 5.13. As a result, at the resonant frequency, the output impedance of the buffer is about 78 ohm (the overall tank impedance) and no additional output matching network is required in this design, because S_{22} is inherently smaller than -10 dB at resonance. The equivalent inductance value can be estimated by (5.14), i.e.

$$L_{eq}=\frac{C_{GS,CD}}{gm}\cdot(R_{tank}-\frac{1}{g_m})=\frac{15\cdot10^{-15}}{0.124}\cdot70\approx9\,pH \tag{5.14}$$

In this work, only the flat part of the plot shown in Fig. 5.13 (c) is used for matching, i.e. $|\omega|\approx R_{tank}$ around ω_0.



(a) CD-stage (b) Small-signal model (c) Frequency domain

Fig. 5.13 The inductive behavior of a common drain buffer when driven by a relatively large impedance.

It should be noted that, in principle, the loading effect of the output buffer should be de-embedded from the measurement data, because the buffer is not part of the IJLO and it does not exist in the final WuRx system. However, the input impedance of the buffer is a very close approximation of the real successive stage of the WuRx, i.e. the passive mixer input. As a result, in the IJLO testing section, no de-embedding process of the buffer is done from the raw testing data.

Total capacitance, the loaded quality factor and tuning range

Based on the above design, the loaded quality factor can be calculated as

$$Q_{loaded} = \frac{1}{\frac{1}{Q_L} + \frac{1}{Q_v} + \frac{1}{Q_{inputbuffer}} + \frac{1}{Q_{outputbuffer}} + \frac{1}{Q_{para,cap,core}}} \tag{5.15}$$

where Q_L, Q_v, $Q_{inputbuffer}$, $Q_{outputbuffer}$ and $Q_{para,cap,core}$ are the unloaded quality factors of the inductor of the LC tank, the varactor, the input buffer ('seen' from the IJLO), the output buffer ('seen' from the IJLO) and the parasitic capacitance of the core transistors respectively. The resonance frequency is then calculated as

$$f_{0,cross-coupled} = \frac{1}{2\pi\sqrt{L \cdot (C_v + C_{inputbuffer} + C_{outputbuffer} + C_{para,cap,core})}} \tag{5.16}$$

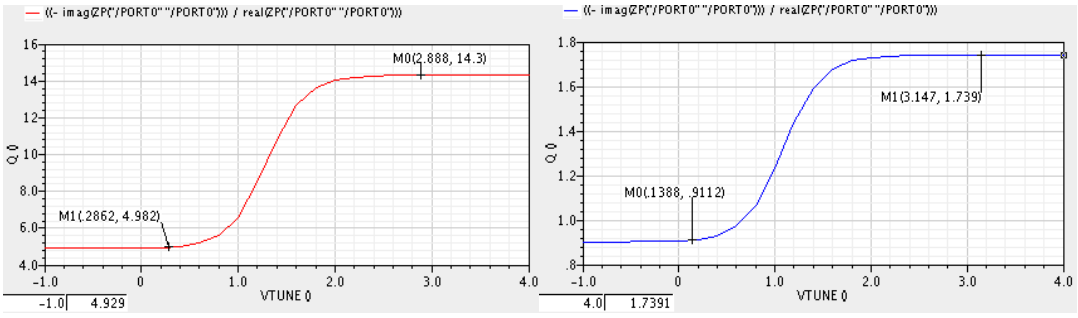
Apparently, due to the loading effect of buffers and the parasitic capacitance of the core, the resonance frequency as well as the overall quality factor of the resonator will be decreased. The capacitance distribution is shown in Table 5.6.

Table 5.6 Simulated capacitance distribution in the 70-GHz IJLO

Varactor capacitance	5fF to 14 fF	$C_{max}/C_{min} \approx 2.8$
Core-induced capacitance	~8 fF	→1.7
Input-buffer-induced capacitance	~5 fF	→1.5
Output-buffer -induced capacitance	~5 fF	→1.4
Others (e.g. inter-connections, inductors)	5 fF *	→1.3
Total Capacitance	28 fF to 37 fF	Overall 1.3

* Post-layout simulation results

According to Table 5.6, the maximum tuning range should be $(1-1.3^{0.5})*100\%$, i.e. 14% for the 70-GHz IJLO. In Fig. 14 (a) and (b), the post-layout simulation results of the open load varactor and overall quality factors are compared.



(a) (b)

Fig. 5.14 (a) The quality factor of the varactor and (b) the overall quality factor with all parasitic capacitances, buffer influence and interconnection imperfections included.

It can be seen that the overall tank impedance R_{TK} becomes

$$R_{TK} = (\omega L \cdot Q_L) // (\frac{Q_v}{\omega C_v}) // (\frac{Q_{inputbuffer}}{\omega C_{inputbuffer}}) // (\frac{Q_{outputbuffer}}{\omega C_{outputbuffer}}) // (\frac{Q_{para,cap,core}}{\omega C_{para,cap,core}}) \quad (5.17)$$

and the peak amplitude of its output voltage $V_{out,peak}$ can then be calculated as

$$V_{out,peak} = \frac{2}{\pi} \cdot I_{bias} \cdot R_{TK} \quad (5.18)$$

Taking a 10-mA bias current and the data from Table 5.5 and Table 5.6 into (5.18), the peak amplitude of the IJLO output voltage (before the output buffer) is about 460 mV. After adding the buffer the final output power becomes about 200-mV referred to the 50-Ohm load, i.e. about -4 dBm. This is consistent to the post-layout simulation results (not including the output transmission line, 50-Ohm on-chip taper and bondpads).

Trade-offs among the phase noise, output power and tuning range

According to Leeson's model [19], the phase noise can be approximately estimated as

$$PN(f_\Delta) = 10 \cdot \log_{10}[\frac{2F \cdot KT}{P_{signal}} \cdot \left(\frac{f_0}{2Q_{loaded} \cdot f_\Delta} \right)^2 \cdot (1 + \frac{f_c}{|f_\Delta|})] \quad (5.19)$$

where f_0 is the natural frequency, f_Δ is the frequency offset, Q_{loaded} is the loaded quality factor, f_c is the flicker noise corner, F is the noise factor and P_{signal} is output signal power. It can be seen that in order to improve the phase noise performance, either the output signal level or the loaded Q should be increased. However, the tuning range and the locking range will be impaired simultaneously, which may lead to higher IJLO and front end power dissipation, as illustrated in Fig. 5.15.

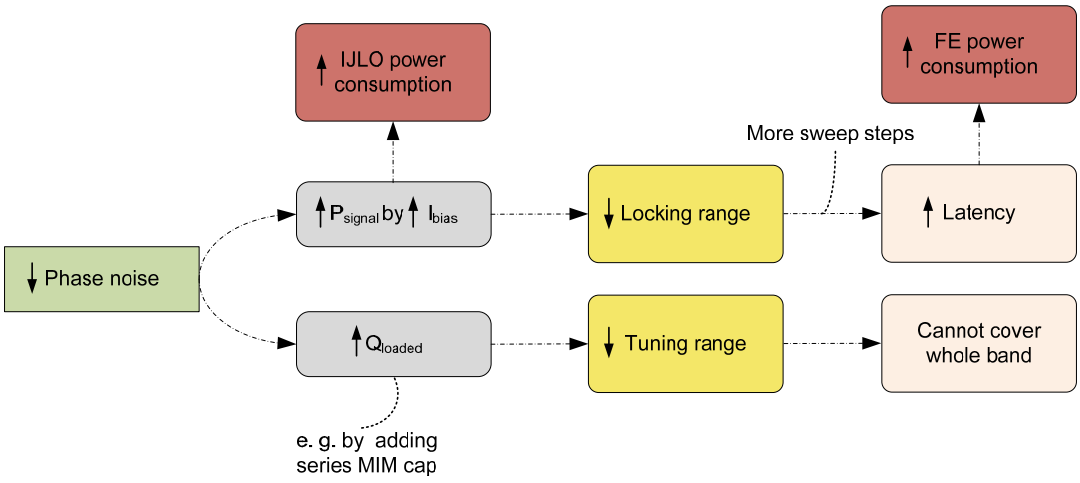


Fig. 5.15 The trade-offs involved in the design of the IJLO.

One thing that should be mentioned is that in the locking condition, the IJLO is inherently capable of reducing the phase noise of the free-running oscillator by adjusting its zero crossings in every period with the injected low-phase-noise signal. In fact, when the frequency difference between ω_0 and ω_{inj} is sufficiently small, the phase noise of the IJLO should be a replica of the injected clean signal [20]. However, this feature cannot be relied on too much in this WuRx design, because the maximum phase noise reduction happens only if ω_{inj} equals ω_0 and any deviation will cause a rapid increase of the phase noise. In our frequency-sweeping locking scheme, it is likely that the IJLO gets locked to the injection signal at the edge of one single-step locking range and in such a situation, the phase noise is close to the phase noise of the free running oscillator. The actual function of injection locking then turns out to be stabilizing the resonance frequency instead of too much noise reduction.

The 70-GHz IJLO core

The complete schematic of the 70-GHz (W-band) IJLO is shown in Fig. 5.16.

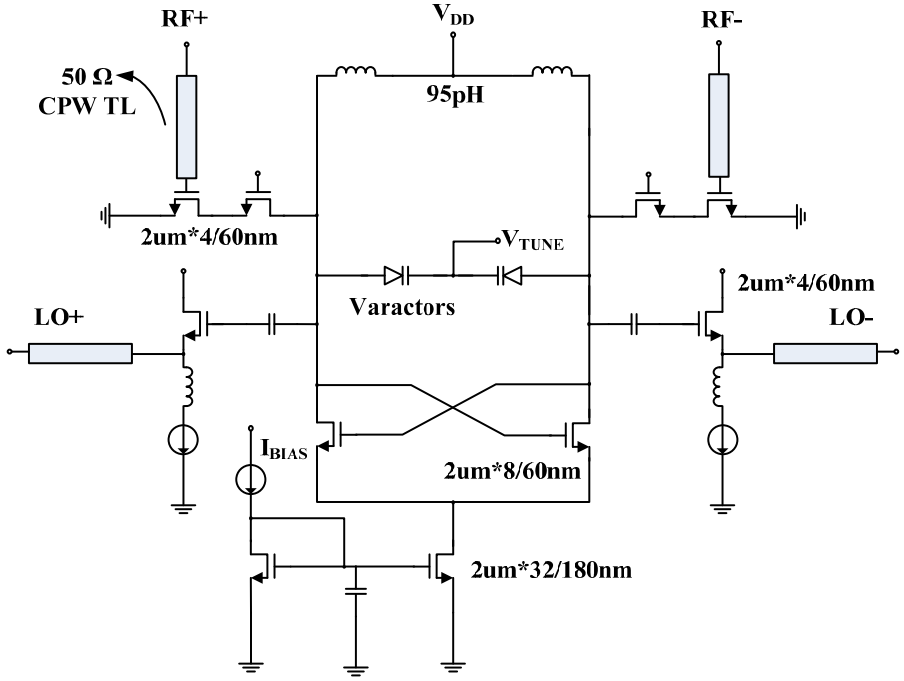


Fig. 5.16 The 73 to 83 GHz IJLO in TSMC 65-nm CMOS technology.

As discussed, the cross-coupled LC oscillator based voltage-controlled oscillator (VCO) is chosen as the oscillation core of the IJLO. The size of these transistors are chosen as 2 μm*8 with minimum length 60 nm. A differential cascode input buffers (2 μm*4) are connected directly to the LC tank so that the injection efficiency is improved and the

input is isolated from the IJLO output. The differential input RF signals are sent through two 50-Ohm characterized slow-wave coplanar waveguide (SW-CPW) transmission lines (TL) [21] from the 50-Ohm signal source to the input buffers. By using the SW-CPW transmission line, the input loss and mismatch are minimized and the chip area occupied by the TL is reduced due to the slow-wave effect. At the output terminals, the differential inductor-enhanced source follower buffers are connected to transfer the IJLO output power through the 50-Ohm TL to the 50-Ohm load. It should be noted that the TL is only used as a low-loss signal path instead of as the matching network. As a result, it is not part of the IJLO circuit and its influence should be eventually de-embedded from the raw measurement data.

One thing to be mentioned is that the phase noise reduction performance may be degraded if frequency locking happens at the edge of each single-step locking range during frequency sweeping. However, it will not cause serious problems in this particular design, because the RF signal in this system is a pure wake-up tone and at the output of the passive mixer, only DC voltage levels as "0" and "1" should be recognized by the comparator, while the side band power is not important as long as the difference in the DC voltages is sufficiently large as the input for a comparator in the baseband. The LO leakage-induced DC offset problem can be further avoided by using the relative detection method or pre-calibration method.

The digital control circuit for frequency sweeping

As discussed in Section 4.3.2, the wake-up tone is recognized in the WuRx by using the frequency sweeping injection locking method, i.e. the Rx keeps on tuning the center frequency of the IJLO with a certain frequency resolution until the IJLO eventually gets locked. The real single-step locking range, i.e. the one calculated by (4.36), is actually small while the total equivalent locking range is large. The worst-case total locking time is modeled by (4.41) and the optimum sweep resolution can be obtained accordingly.

$$t_{settling,tot} = \frac{BW_{tot} \cdot 2\pi}{\omega_{res}} \cdot \frac{1}{\sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2}} \cdot \ln(\frac{\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}} + \sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2} - 2\omega_{res} \cdot \tan(\frac{\arcsin(\frac{2\omega_{res}}{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})})}{2})}{\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}} - \sqrt{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})^2 - (2\omega_{res})^2} - 2\omega_{res} \cdot \tan(\frac{\arcsin(\frac{2\omega_{res}}{(\frac{\omega_0}{Q} \cdot \frac{I_{inj}}{I_{osc}})})}{2})}) \tag{4.41}$$

Taking the post-layout simulation results obtained in Section 5.3.3 of the W-band IJLO, i.e. Q_{loaded}=1.5, I_{osc}=10 mA and BW_{tot}=10 GHz (note: it is 4 GHz larger than the required bandwidth in case any frequency deviation might happen in the measurement data) into (4.41), the total worst-case locking time is plotted in Fig. 5.17.

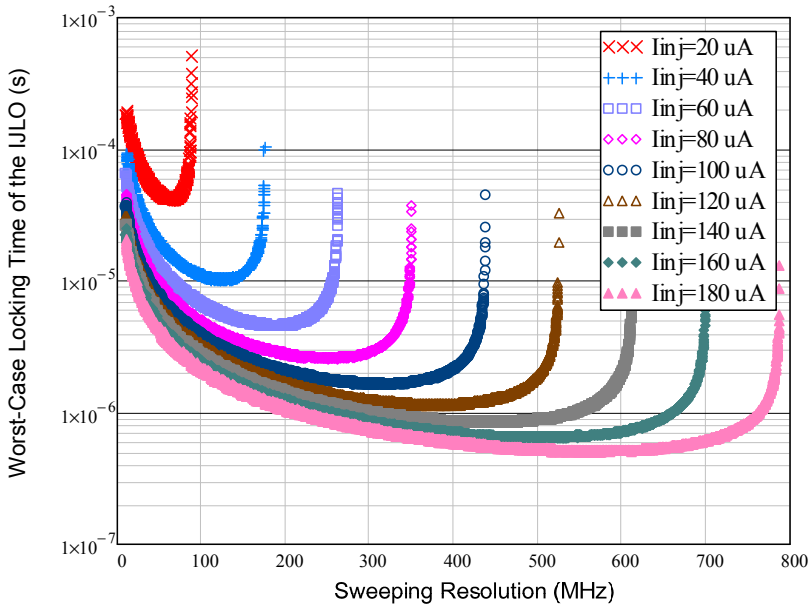


Fig. 5.17 The worst-case locking time as a function of the sweep resolution and the injection current strength. It should be noted that ω_{res} is divided by 2π and transferred into f_{res} as the X-axis.

In the extreme condition, a very weak wake-up tone is received by the WuRx. For example, -60 dBm injection power is obtained in the 50-Ohm antenna of the WuRx, and the antenna gain is assumed as 0 dB. When translated to the current signal and amplified by the input buffer, the resulting current magnitude is around 40 μA (i.e. 20 μA for each side of the differential input buffer). In such a case, the optimum worst-case locking time of the IJLO becomes about 10 μs if the sweep resolution is from 100 to 150 MHz. In the best case, i.e. when the injection frequency lies within the single-step locking range of the IJLO, the total locking time is just about 100 ns and a single sweep step is enough.

In the IJLO, the frequency sweeping can be realized by tuning the bias voltage of the varactors in the resonator. A 100-MHz sweep step can be corresponded to certain voltage tuning steps and pre-calibration is required to remove the non-linear tuning effect. The ultra low-power digital control module can be used to realize such step-like sweep voltages, as shown in Fig. 5.18. With a 7-bit synchronous counter and a 7-bit digital-to-analog convertor (DAC), 128-step sweep voltages are generated as the control voltages. The sweep resolution is reduced to about 78 MHz in this case. Referring to Fig.

5.17, the total locking time is slightly increased. However, smaller sweep step means more locking range overlap between two successive tunings, and it helps to guarantee the locking condition, especially when considering the mismatches and parasitics of the digital circuits. Comparing to 100-MHz required sweep resolution, 78-MHz sweep step would mean about 22% more locking margin.

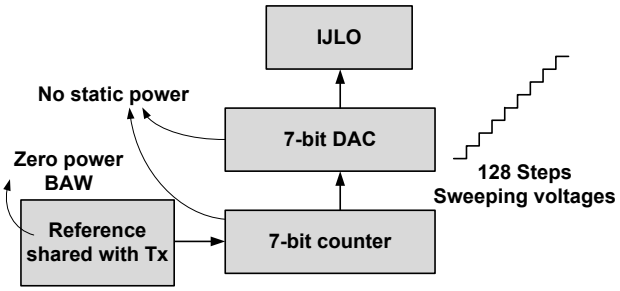


Fig. 5.18 Diagram of the 7-bit low-power sweep voltage generator.

The schematic of the sweep voltage generation module is shown in Fig. 5.19. A synchronous up counter is realized, as shown at the bottom, by serializing seven J-K flip-flops together. In order to make sure their output bits change states simultaneously, NAND blocks are placed between the stages (except for the first and second stages) so that each of the higher-order flip-flops are made ready to toggle (both J and K inputs "high") if the Q outputs of all previous flip-flops are “high” [22]. In all other situations, either J or K or both are “low”, the flip-flop will be in the "latch" mode and it will maintain its present output state at the next clock pulse. The J and K terminals of the first flip-flop are connected to the clock and the supply, respectively, so that it will change its state along with the clock. The second flip-flop is made ready to toggle along with the Q output of the first one, so no AND gate is needed. Similarly, the third flip-flop is made ready to toggle only when both Q outputs of the previous stage are high, which is realized by connecting them with a 2-input AND gate. The rest stages follow the same principle. A 7-bit switched-capacitor DAC is chosen in this design due to the following reasons: (1) it can be realized with ultra-low power, because in principle, there is no current consumption except during the transition moments of the CMOS switches; (2) a 7-bit control voltage resolution does not require high linearity of the DAC and therefore smaller capacitors can be used and no calibration needs to be done; and (3) unlike the real analog-to-digital conversion process, the frequency sweeping voltages do not require stringent tuning accuracy. In the switched-capacitor DAC, the errors mainly come from the mismatches between capacitors[29]. Assuming a 1% mismatch of the MIM capacitors

[29] Since the non-linear relationship between the control voltages and the varactor values of the IJLO is deterministic and can be solved by mapping the required frequencies to corresponding control voltages, we only talk about the stochastic errors in the DAC in this Section.

[23], the resulting frequency tuning error will be about 0.5% (i.e. the frequency tuning value is approximately the square-root of the capacitance tuning value), which will not cause trouble in our application if we reserve about 1% margin in the sweep control voltage. In other words, we leave more locking margin by enlarging the overlap of each single-step locking range by 1% of the sweep resolution (0.4 MHz), and the IJLO is able to get locked in presence of 1% mismatches.

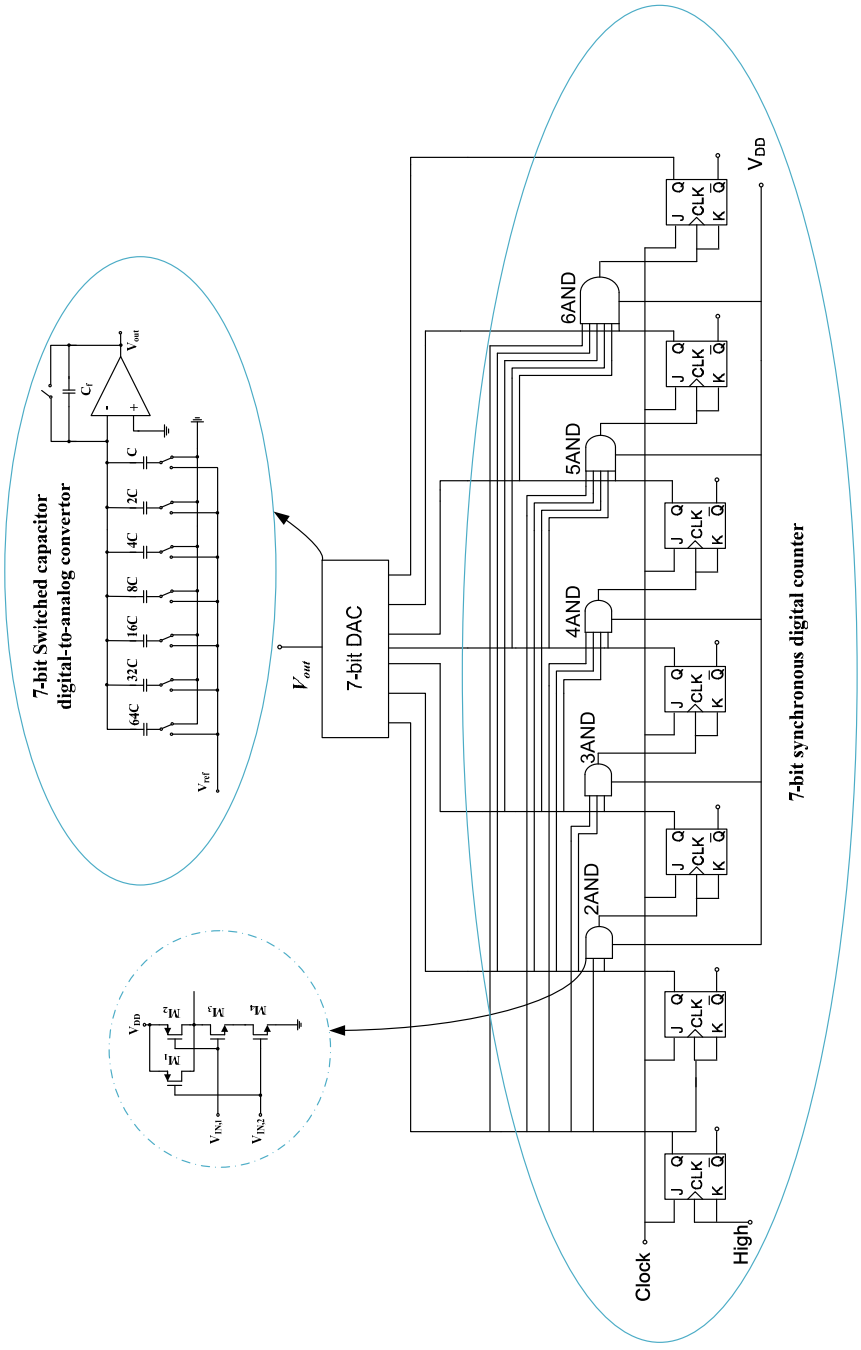


Fig. 5.19 The schematic of the 128-step low-power sweep-voltage generator.

The simulated waveforms of the 128-step control voltages are shown in Fig. 5.20.

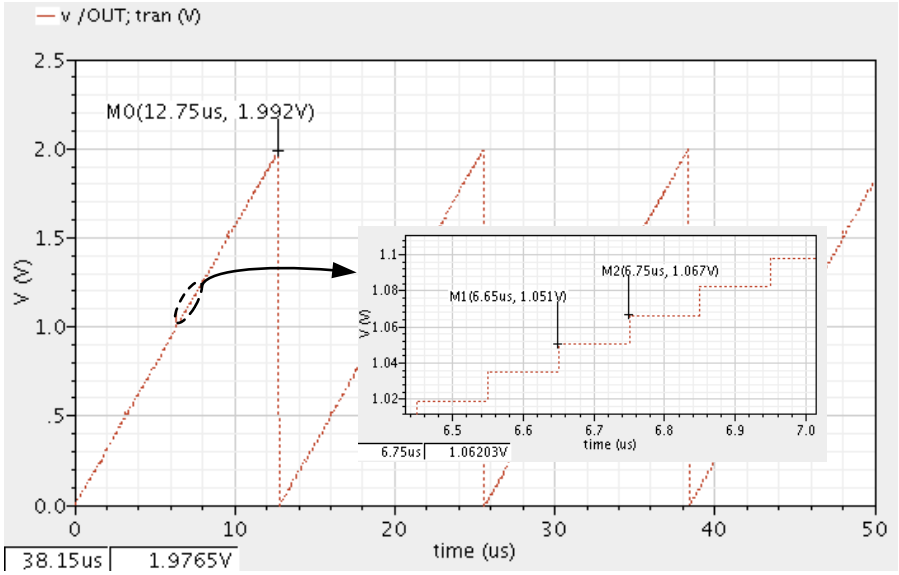


Fig. 5.20 The 7-bit digital control voltages.

5.3.4 Design of the passive mixer and the buffer amplifier

There are several reasons of choosing a passive mixer in our design. Firstly, it features ultra-low power consumption. Secondly, it has good linearity compared to the conventional active Gilbert mixers with output voltage swing limited by the headroom of the switching core and the input g_m-stage. A transformer-coupled method is proposed for the mixer in [24], which is able to solve the headroom problem and use separate supplies for the input and the switching stages, yet, it increases the design complexity, chip area and also the chance of undesired inductive coupling. Thirdly and most importantly, it generates ultra-low flicker noise. In principle, it only generates the flicker noise when both transistors in the pair are partially on or off under the current induced by the DC offset [25], but this problem is largely alleviated in this work due to the low-power IJLO scheme. The basic idea is to deliberately keep the output power of the IJLO as low as -10 dBm while maximizing its voltage swing by choosing small transistors for the mixer core. This method is able to minimize the capacitive loading and maximize the tuning range of the IJLO as well at the expense of large conversion loss of the power, which can be further compensated by the output amplifier buffer. In fact, the power conversion gain is less important in this particular system, because the mixer and the output buffer amplifier operate in the voltage domain, and the power loss in the mixer is not very critical as long as the voltage conversion gain is sufficient and the output voltage is large enough to drive the comparator which typically has a high input impedance at low frequencies. Lastly but not least, due to the impedance transparency property of the passive mixer, the output

low-pass feature is transformed as a dynamic tracking band-pass filter at the input as discussed in Section 4.3.3, and thus the noise bandwidth is reduced significantly. Detailed design considerations and analyses will be presented in the following.

The double-balanced mixer structure is normally required in the conventional Rx to enhance the LO-IF and RF-IF isolations. It is especially important for the superheterodyne receiver. For example, as illustrated in Fig. 5.21, the frequency of LO_1 is 30 GHz and IF_1 (i.e. RF_2) is 30 GHz too. If the LO-IF isolation is poor for the first mixer, it will degrade the input SNR of the second stage dramatically and thus cause demodulation failure. Unlike the active mixer, however, the (open-load) voltage conversion gain of the double-balanced passive mixer is theoretically 6-dB lower (2-times) than the single-balanced passive mixer because the single-balanced passive mixer is able to produce the same output voltage with only half of the input voltage. This is a very different situation than in the case with the active double-balanced mixer where the voltage conversion gain is identical to the single-balanced mixer with the same input level. The fundamental reason is that in an active double-balanced mixer, the output currents of both branches can be summed up while in the passive double-balanced mixer, the outputs are voltages and cannot be added directly [25].

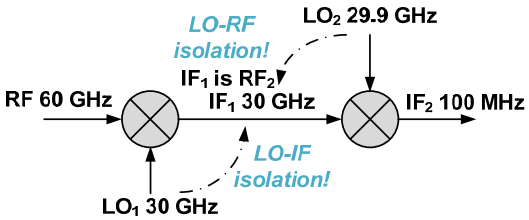
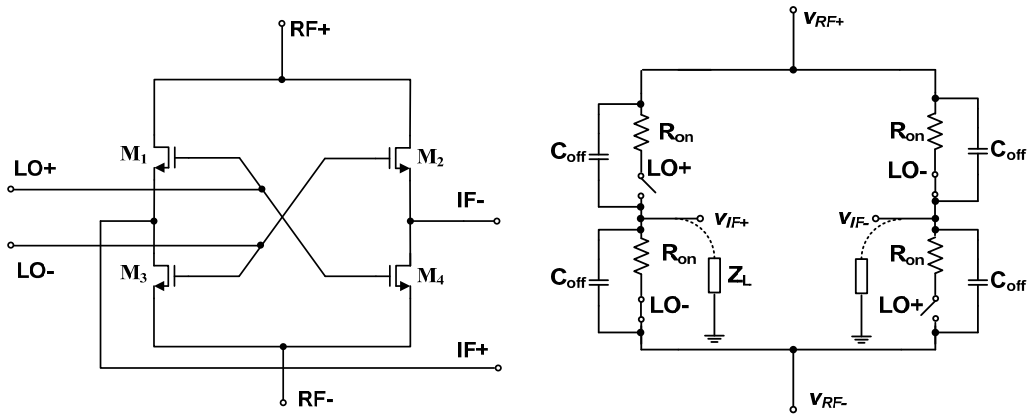


Fig. 5.21 The LO-IF and LO-RF isolations in a superheterodyne system.

Despite a low voltage conversion gain, the double-balanced passive mixer structure is chosen to be used in this WuRx system. The most important reason is that it has to be connected directly to the IJLO output. As discussed in Section 5.3.3, the injection signal should be fully differential and thus the injection paths must be fully symmetrical, otherwise there will be a large common-mode output signal generated. When driving the passive mixer by such an imperfect signal, the output voltage will be reduced and the conversion loss will be increased dramatically. As a result, the fully symmetrical signal paths should always be carefully designed, not only in the system architecture but also in the final layout design. The double-balanced passive mixer is shown in Fig. 5.22 (a) and its conceptual model is illustrated in Fig. 5.22 (b) where the mixer core transistors are modeled as an ideal switch with a small on-resistance R_{on} and a large off-impedance Z_{off}. Z_{off} is approximately identical to $1/j\omega C_{off}$ in which C_{off} is the summation of the gate-drain, gate-source and source-drain parasitic capacitances of the transistor in the cut-off mode.



(a) Double-balanced passive mixer (b) Conceptual model

Fig. 5.22 The schematic and conceptual model of a double-balanced passive mixer.

As discussed in [26], the voltage conversion gain of such a double-balanced passive mixer can be approximated as

$$G_{c,v} = 20\log(\frac{2}{\pi}(\frac{Z_L}{Z_L + R_{on} // Z_{off}}) - \frac{2}{\pi}(\frac{Z_L}{Z_L + Z_{off}})) \tag{5.20}$$

or simplified as

$$G_{c,v} = 20\log(\frac{2}{\pi}(\frac{Z_{off} - R_{on}}{(1+\frac{R_{on}}{Z_L})\cdot(1+\frac{Z_{off}}{Z_L})})) \tag{5.21}$$

where $R_{on}/|Z_L|$ is assumed <<1 in here.

From (5.21), it can be seen that there are three methods that can be used to improve the conversion gain: increasing Z_{off}, decreasing R_{on} and increasing Z_L. Increasing Z_{off} means that smaller transistors should be chosen so that the gate-source and gate-drain capacitance is reduced. Besides, the loading capacitance for the IJLO is reduced too. This is very important because the input capacitance of the mixer affects both the IJLO center frequency and the tuning range, i.e. the minimal fixed capacitance leads to the maximum tuning range. However, small transistors lead to a large R_{on} (i.e. the drain-source on-impedance r_{ds} when the transistor is in the deep-triode region), which in turn degrades the voltage conversion gain. In order to compensate for the degradation of $G_{c,v}$ and keep the small transistor size of the mixer as the optimum load for the IJLO, the output amplifier buffer should be added, which offers a large load impedance, i.e. the gate capacitance of small-size transistors, for the mixer at low frequency. Besides of that, the series parasitic resistance and the gate capacitance of the buffer inherently form an R-C LPF at the output, which can be used to filter out the high-frequency output components

and reduce the noise bandwidth. Based on the above considerations, the transistor size of the mixer core is chosen with a total width of 2 μm*2 and the minimum length of 60 nm for low parasitics. Besides, transistors M_1 to M_4 are biased at their threshold point, i.e. with effective V_{gs} of 0.45 mV so that these switches are made to work in the class B mode. Consequently, a relatively small driving voltage swing is able to turn the switch on and off effectively.

At the output of the mixer, a folded cascode gain-boosted output buffer amplifier is added, as shown in Fig. 5.23. The differential output voltages from the mixer are sent into the differential pair M_1 and M_2, and then further amplified by the folded cascode stages, i.e. the M_3-M_4 pair, and the M_5-M_6 pair. Their outputs are then amplified by M_7 and M_8 respectively as the final output voltages. In this way, the amplifier gain is boosted without requiring additional voltage headroom [18] which is hardly available from the low 1-V supply voltage. Instead of tuning the feedback capacitors C_f, the bandwidth tuning is realized by tuning the gate voltages of the NMOS resistors, i.e. $V_{tune,BW}$ in Fig. 5.23. In order to reserve some margin for the future wake-up bits which may have data rate as high as, e.g. 1 Gbps, the 3-dB bandwidth of the buffer amplifier is deliberately designed larger than 1 GHz in the wideband mode. The overall power dissipation of the buffer is then about 1.4 mA from a 1-V supply. While in the narrow band mode, the 3-dB bandwidth is reduced to about 100 MHz. The noise bandwidth is reduced by 10 dB in this way at the expense of 10-times longer detection time.

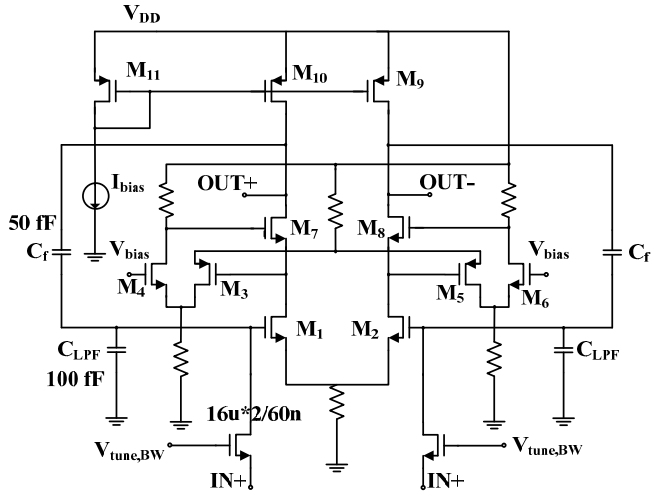
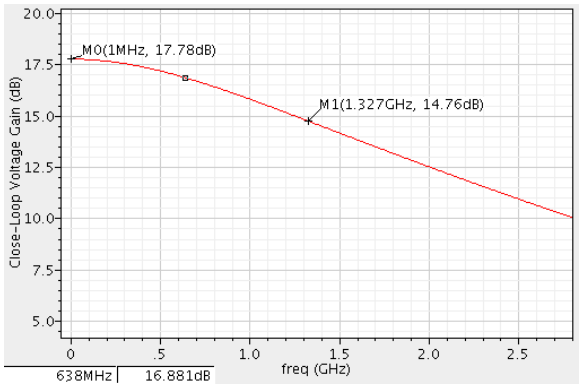


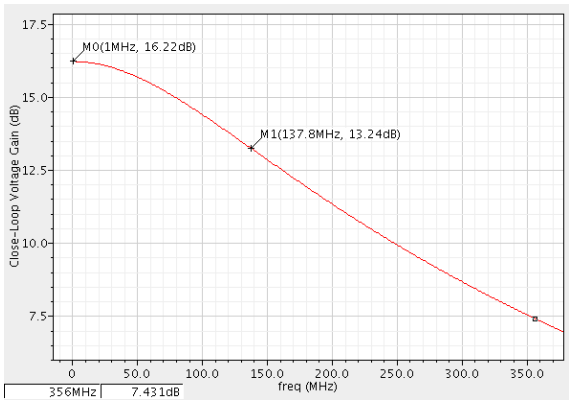
Fig. 5.23 The gain-boosted buffer amplifier.

Simulated AC-responses of the buffer amplifier are shown in Fig. 5.24. In (a), the 3-dB bandwidth of the buffer amplifier is about 1.3 GHz. This can be used for receiving the 1-Gbps wake-up bits and be reused in the main Rx front end. In (b), the bandwidth is reduced to about 100 MHz, which can be used to receive pure wake-up tone. Better

WuRx sensitivity can be further achieved by using more complicated higher-order narrow band output filter.



(a) wideband mode



(b) narrow band mode

Fig. 5.24 The AC-responses of the output buffer.

5.3.5 Design of the wake-up receiver

First, let us have a brief review on the wake-up flow, i.e. how this system works.

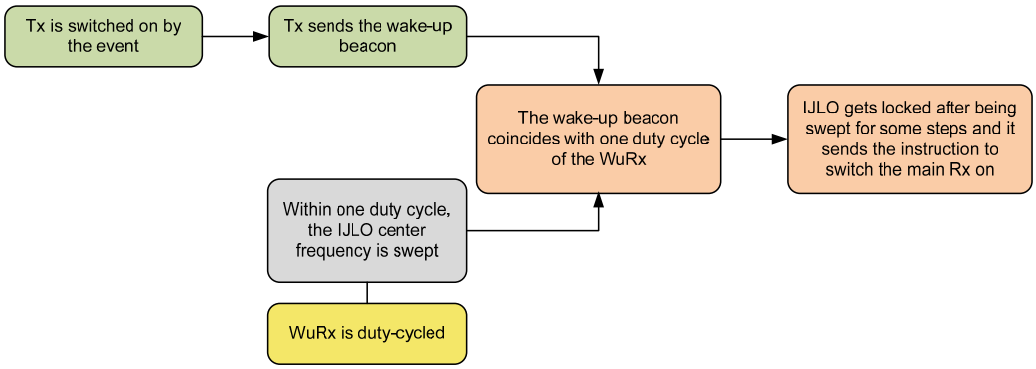


Fig. 5.25 The wake-up process.

It can be seen that in principle, the wake-up beacon only needs to coincide with one duty cycle of the WuRx in order to enable it to get locked. A lower DCF is helpful to minimize the average power consumption of the WuRx, but it will increase the wake-up beacon length and also affect the total average power consumption as a penalty. Recalling (3.46), the minimum average power consumption of the Tx and WuRx in the wake up process can be modeled as

$$P_{ave,DC,min} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot P_{Tx,DC} \cdot n_{event}}{T_{tot}}} + P_{Tx,DC} \cdot \frac{(T_{DC} + 2 \cdot T_{wu}) \cdot n_{event}}{T_{tot}} \tag{3.46}$$

and the optimum DCF of the WuRx is

$$T_{DC} = \sqrt{\frac{P_{WuRx,DC} \cdot T_{wu} \cdot T_{tot}}{P_{Tx,DC} \cdot n_{event}}} \tag{3.45}$$

where $P_{WuRx,DC}$ and $\underline{P_{Tx}}$, DC are the power consumption of the WuRx and Tx, respectively, T_{tot} is the total observing time for a certain event frequency n_{event}, and T_{wu} is the total settling time of the WuRx, i.e. the sweeping (locking) time of the IJLO. From (3.46) and the discussion in Section 5.3.4, the minimal average power consumption decreases when T_{wu} is reduced. In other words, with a higher injection level (which can be seen as a worse WuRx sensitivity), the locking process can be faster and the total average power can be further minimized. Or reversely, if the system sensitivity is already pre-determined, e.g. by the link budget model and the application requirements, there will be no more freedom to accelerate the locking and thus the minimal achievable average power consumption is fixed (assuming the circuits are already optimized, i.e. with the lowest power consumption and an acceptable noise level, as listed in Table 5.2).

Based on the above discussion, the design methodology of an IJLO-based asynchronous duty-cycled WuRx for minimum average power consumption can be summarized as

1) Choose the system parameters: the noise bandwidth, the noise factor of the WuRx from the application requirements, and the system requirements from the link budget model and application requirements.
2) Estimate the system sensitivity from the link budget model with some power margin reserved in case of SNR degradation.
3) Obtain the injection current I_{inj} level from the required sensitivity value and the optimized input stage design.
4) Choose the proper sweeping resolution by taking the actual design results into (4.41). For example, the W-band result of (4.41) is plotted in Fig. 5.17.
5) Design the digital sweep voltage generator accordingly. Some margins should be

reserved, i.e. the frequency overlap of two successive tunings should be sufficiently large, which means about 20%.

6) Optimize the other circuits with regards to power and noise.
7) Take these results into (3.45) and obtain the optimum DCF of the WuRx.
8) Substitute the obtained DCF and all the circuit-dependent values into (3.46) and obtain the minimal achievable average power consumption. It should be noted that n_{event} (in fact, it is n_{event}/T_{tot}, which means n-times of events during T_{tot}) is the only application-determined factor here.

Now, let us continue with the circuit design and optimization, as required in step 6. Combining the IJLO, digital tuning control, passive mixer and output buffer, the complete WuRx is obtained, as shown in Fig. 5.26 (no bias and sweeping circuits).

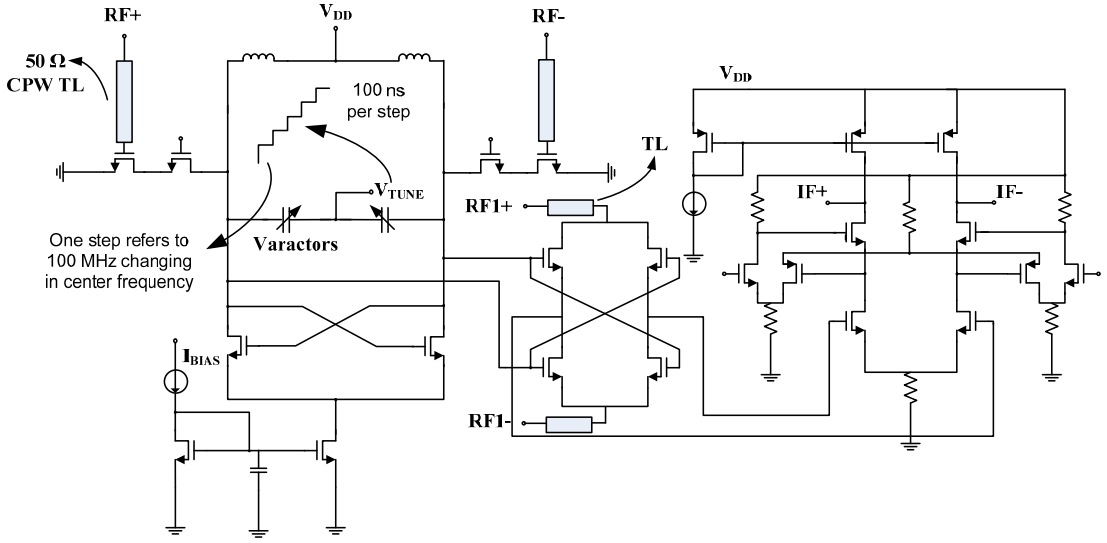


Fig. 5.26 The W-band WuRx circuit.

Since the input impedance of the passive mixer is already optimized to reduce the loading effect of the IJLO, no buffer stage is required. However, compared to the output buffer of the IJLO (Fig. 5.12 (d)), the input parasitic capacitance of the passive mixer will be larger. The reason is that there is no way to add the series inductor to resonate the capacitance out (the inductor will increase R_{on} at the RF frequency). As a result, the center frequency of the free-running IJLO and its tuning range will be decreased a bit. It should be noted that in Fig. 5.26, the terminal RF+ should be connected to RF1+, and the terminal RF- should be connected to RF1-, according to the system architecture design shown in Fig. 5.6. However, they are separated deliberately (also in the layout) for measurement purposes. If the IJLO input frequency is identical to the mixer's RF frequency, when the IJLO gets locked, the output of the mixer will be a pure DC voltage, but the DC power cannot accurately be measured by the spectrum analyzer. Consequently, a slight

frequency difference (e.g. 10 to 100 MHz) is deliberately left to the RF and RF1 ports so that at the output port, a small IF product will appear instead of the DC voltage, and the power and phase noise behavior of such IF signal can be well measured by the instrument.

The working principle of the WuRx can be explained qualitatively as follows. When the OOK data is 1, the output signal of the mixer v_{out} can be calculated from

$$v_{out,RF} = v_{RF}(\omega_0) \cdot S_{LO}(C, \omega_0, ...) \tag{5.22}$$

where $v_{RF}(\omega_0)$ is the input RF signal, S_{LO} is the switching function of the MOS switches and C is the DC offset component. It can be seen that in our system, the frequencies of the LO driver and RF signal are identical and thus the desired output signal is a pure DC voltage. If we remove all the undesired components like ω_0 from (5.22), the output DC will be determined by the product of $v_{RF}(\omega_0) \cdot S_{LO}(\omega_0)$, which can be expressed as

$$v_{out,RF,DC} \propto \left|v_{RF}(\omega_0)\right| \cdot \left|S_{LO}(\omega_0)\right| \tag{5.23}$$

where $|\cdot|$ means the amplitude of the signal. When no wake-up tone is present at the RF port of the passive mixer, the RF input becomes a DC voltage which is identical to the bias voltage of the mixer, i.e. the gate-source voltage V_{GS} (there is no capacitor in between of the mixer and buffer amplifier), so the output signal becomes

$$v_{out,noRF} = V_{GS} \cdot S_{LO}(C, \omega_0, ...) \tag{5.24}$$

After filtering the high-frequency component, (5.24) results in a DC level as

$$v_{out,noRF,DC} \propto V_{GS} \cdot \left|S_{LO}(C)\right| \tag{5.25}$$

It can be seen that no matter there is RF signal or not, a DC voltage will always appear at the final output of the WuRx, but their voltage will differ in amplitude. Consequently, a relative detection method can be adopted in the output comparator, i.e. only detecting the difference instead of the absolute values. Alternatively, a pre-calibration can be done to the output DC voltage and remove the influence of the DC offset.

When looking at (5.23) and (5.25), however, a very important phase problem will occur, which may impair the conversion gain of the passive mixer dramatically. According to the nature of the injection locking, the output voltage of the IJLO should be a replica of the injected RF signal in frequency, but there is no guarantee for the phase performance. As the result, the output signal of IJLO will have some phase shift (φ) with the real input RF signal, and (5.23) will become

$$v_{out,RF} = v_{RF}(\omega_0) \cdot S_{LO}[C, (\omega_0 + \frac{\varphi}{t}), ...] \tag{5.26}$$

After removing the unused components, the final output of the mixer turns to be

$$v_{out,RF,DC,\varphi} \propto \left|v_{RF}(\omega_0)\right| \cdot \cos(\varphi) \tag{5.27}$$

In the worst case, the output DC voltage may drop to 0 (i.e. when φ is $\pi/2$). This situation must be carefully avoided, i.e. the output signal of the IJLO and the input RF signal should be kept in-phase or just with a very small phase difference. There are several potential solutions to solve this problem. For instance, we can insert a phase shifter at the input or output of the IJLO (pre-phase-compensation, or dummy transistors which are replicas of the input buffer for the phase adjustment) or have the additional quadrature (Q) channel and recover the amplitude information afterwards by using the square-root circuit. These methods are compared in Table 5.7.

Table 5.7 Three methods for phase compensation

	Pros	**Cons**
Phase compensation at the output of the IJLO (dummy)	-Good isolation between the IJLO and mixer	-Loading effect -Double power dissipation of the input buffers -The dummy may not produce the exactly identical phase shift because (1) the steady-state phase difference between injection and oscillation signals of the IJLO depends on the values of injection power and injection frequency and (2) the signal levels are different between the input RF and output of IJLO and the capacitance values are different
Phase pre-calibration (phase shifter)	-Easiest -Simple phase shifter	-Loss and reduced single-step locking range -Matching difficulty -No guarantee for the perfect in-phase condition
Q channel	-Easy and straightforward	-Double power consumption -Limited by the accuracy of the Q-phase -Much more serious loading effect, which will cause lower output power and worse phase noise of the free-running IJLO due to the reduction of inductor value of the LC tank -Output is still not maximized

Based on the above comparison, the phase pre-compensation method is chosen in the first generation WuRx system for its simplicity. The basic idea is to insert a low-power phase shifter in the input RF path of the IJLO and pre-compensate the potential extra phase shift. As a result, the additional phase sweeping must be done during the wake-up process. From (5.27), we know that the worst-case phase shift is 90°, i.e. LO and RF are completely out of phase, which means the maximum required phase tuning range of the phase shifter is 90° (lagging or leading). For example, if we set four states for the phase

shifter, which are separated with 22.5° distance and we sweep four states in the very beginning of the wake-up process. There must be one state that leads to the maximum output signal compared to others, and then we will select that state for the rest. It should be noted that the best state does not necessarily make sure the LO and RF are fully in-phase, but it is the state that is most close to the in-phase condition. As a result, the worst-case output signal becomes

$$v_{out,RF,DC,\varphi} \propto \left|v_{RF}(\omega_0)\right| \cdot \cos(22.5^\circ) \tag{5.28}$$

i.e. about 90% of the maximum achievable value. Setting more states are helpful to make the output voltage approach the maximum value but it will surly increase the total settling time (the total latency) of the system. Since the total phase tuning range is 90°, the simple all-pass RC phase shift can be used here, which is zero-biased, as shown in Fig. 5.27.

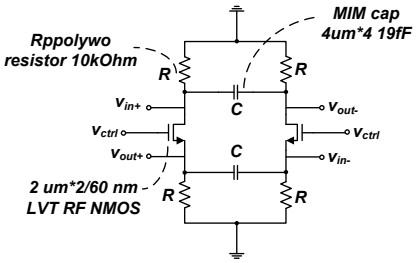


Fig. 5.27 The RC zero-biased all-pass phase shifter.

The magnitude and phase of its transfer function can be calculated as

$$mag(H(\omega)) = \sqrt{\left(\frac{1-\omega^2R^2C^2}{1+\omega^2R^2C^2}\right)^2 + 4\cdot\left(\frac{\omega RC}{1+\omega^2R^2C^2}\right)^2} = 1 \tag{5.29}$$

$$\Phi(H(\omega)) = \arctan\left(\frac{-2\omega RC}{1-\omega^2R^2C^2}\right) \tag{5.30}$$

The simulated phase tuning behavior is shown in Fig. 5.28.

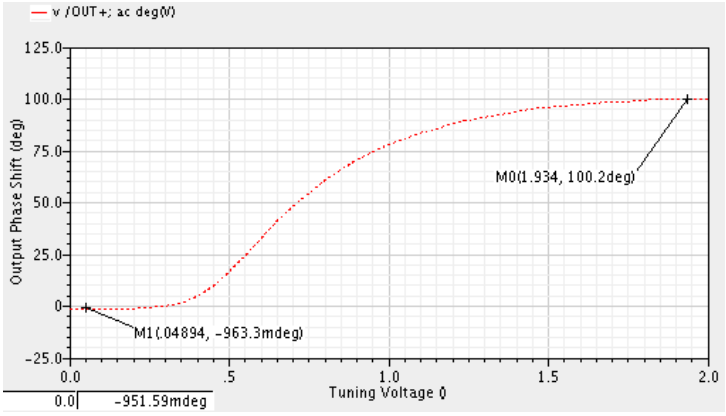


Fig. 5.28 The phase shift range of the RC phase shifter.

As a result, the final WuRx system is shown in Fig. 5.29.

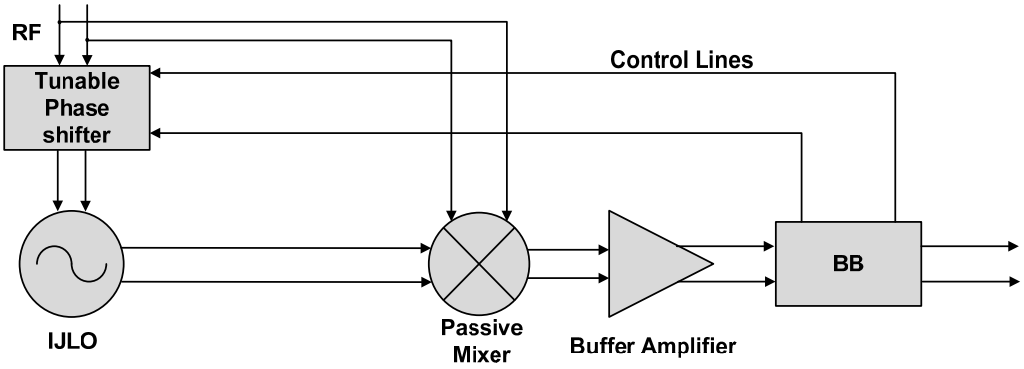


Fig. 5.29 The updated system diagram.

In order to test the locking speed, the input wake-up beacon is modulated by a random bits sequence and then sent to the WuRx, which is already in the locking condition. The result is shown in Fig. 5.30. It should be noted that in this simulation, the frequency-sweeping locking process is already done and the IJLO should keep on staying in the stable locking state as long as the wake-up beacon keeps appearing. If there is no wake-up beacon, the output DC voltage will become low.

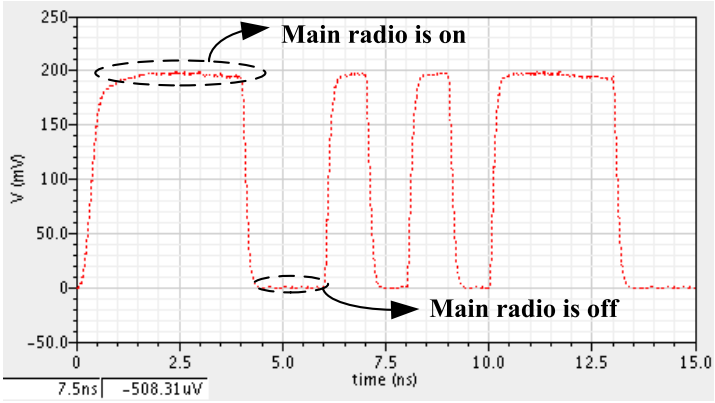


Fig. 5.30 The output signals of the WuRx (open load): random RF amplitudes are used to test the speed of the wake up process. The output signal of the IJLO and the RF signal are tuned in phase in this simulation.

The influence of the phase tuning is illustrated in Fig. 5.31. It can be seen that when the tuning voltage is 1 V, the output DC is very close to the maximum value compared to Fig. 5.30.

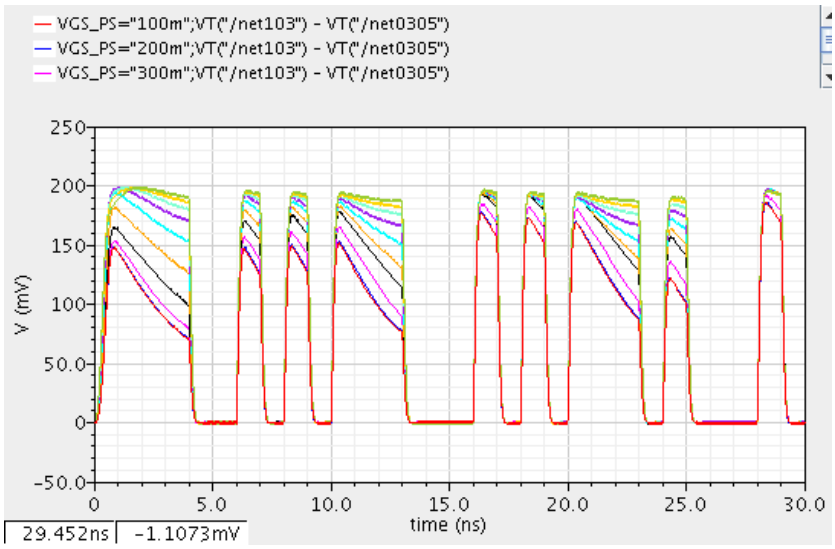


Fig 5.31 The influence of the phase shift.

5.4 Implementation and experimental results

The measurement results are shown in this section. Some relevant topics like the measurement setups and (part of) layout techniques are presented as well. The original measurement data and figures are shown followed with some discussions and necessary data process (e.g. losses testing and de-embedding).

IJLO standalone results

The chip photo of the IJLO is shown in Fig. 5.32. At the top and the bottom are the differential RF GSGSG bondpads, which are characterized to 50-Ohm impedance up to 67 GHz. On the left side are the minimal-size eye-pass bias pads, which have a bandwidth of about 200 MHz. The center-to-center distance between two pads in the eye-pass is 75 μm. This is helpful to make sure the layout is compact and shorter I/O TLs are used. The effective core area (i.e. without the bondpads and I/O TLs) is about 0.055 mm^2. In the empty area, two 50-pF MIM capacitors are placed to locally de-couple the supply to ground. Besides of that, the meshed planes of V_{DD} (metal 6 layer) and ground (metal 1 layer) are adopted in order to reduce the resistance while not violating the metal density requirements.

All the sensitive RF paths except the TL area are protected by the meshed ground layer, as shown in Fig. 5.33, which helps to improve the quality factor of metal wires (like the on-chip inductor mesh) as well as preventing the substrate coupling due to leakage.

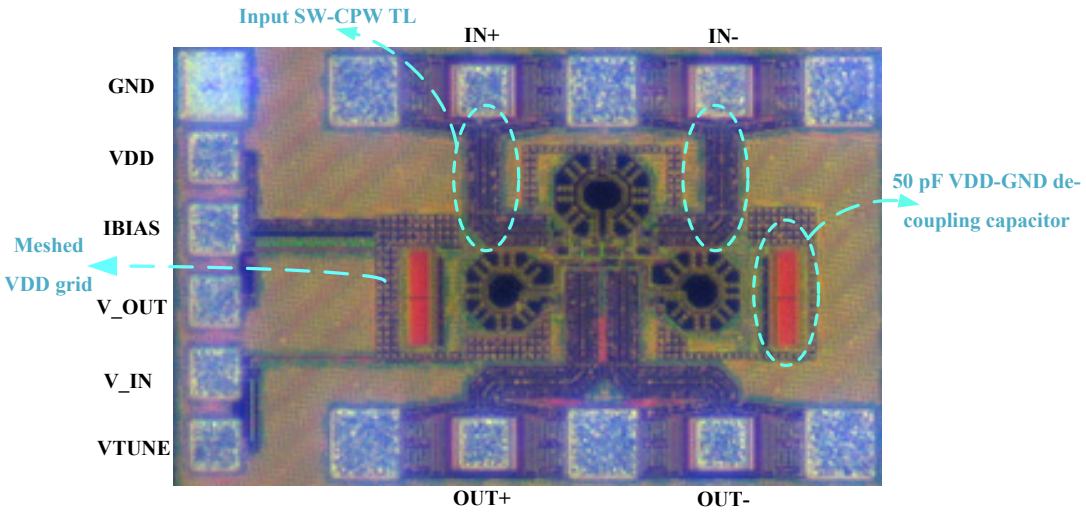


Fig. 5.32 Chip photo of the IJLO.

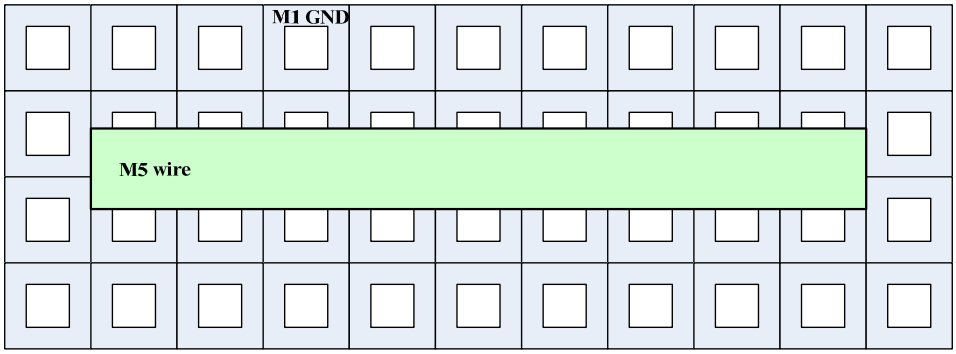


Fig. 5.33 The meshed ground plane.

The measurement setup is shown in Fig. 5.34. Due to the lack of a high-frequency signal generator or network analyzer, the required 70 to 80 GHz signals are generated by up-mixing the low-frequency output signals of two separate generators, as illustrated at the top of Fig. 5.34. The maximum signal frequency obtained from Agilent E8267D is 31.8 GHz. If we double this frequency and use it as the IF frequency, and then use Agilent E8257D to up-mix the IF, we can obtain a weak output signal with the highest frequency of 81.8 GHz.

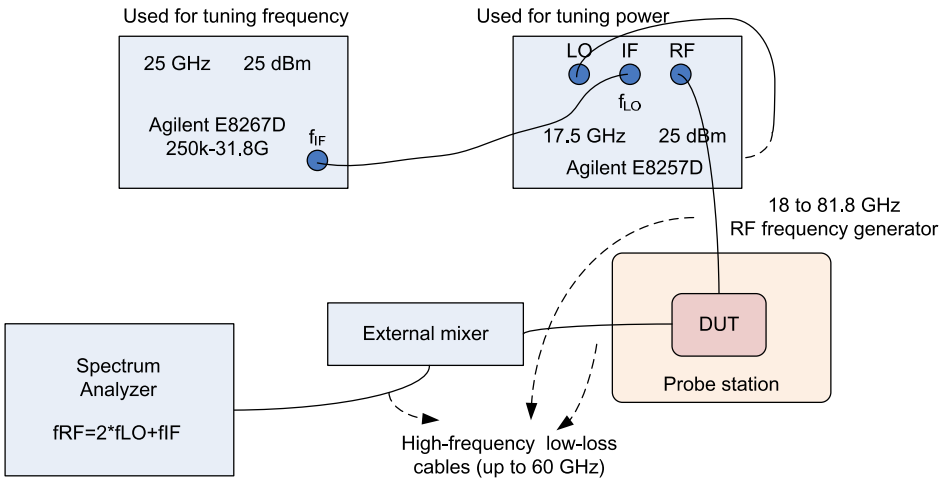
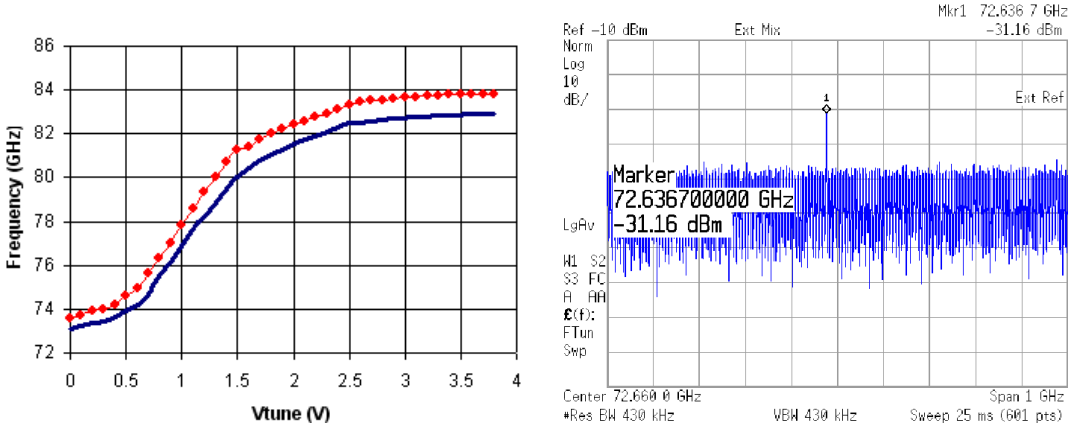


Fig. 5.34 The 70 to 80 GHz signal generation mechanism.

Several samples are tested and the tuning range measurement results are shown in Fig. 5.35 (a). The dotted line is the average curve from samples nearby the center and the solid line is from ones nearby the edge of the wafer. The maximum achievable tuning range of both is around 10 GHz, i.e. from 73 to 83 GHz. The linear range, i.e. from 74 to 82 GHz (V_{tune} is from 0.5 to 2 V) will be chosen for the future design so that the calibration of the control signals becomes easier and reliability problems are avoided.



(a) (b)

Fig. 5.35 (a) the tuning range of the IJLO and (b) output spectrum (not de-embedded)

The spectrum of the IJLO output signal is shown in Fig. 5.35 (b). The output power from the output pad is about -10 dBm. There is more than 15 dB of the total losses from bondpads, cables, off-chip capacitors, baluns and connectors at 73 GHz, because all these devices are able to support the 60 GHz measurement. It is found that when the frequency gets higher than 70 GHz, the losses increase dramatically (almost in a quadratic manner). The waveguide connectors are helpful to reduce the loss, but for the other devices, there

is no proper low-cost solution. The only way is to measure the total loss and de-embed it. This will surely introduce inaccuracy with respect to power measurement, because the matching condition is different as illustrated in Fig. 5.36 (a) and (b).

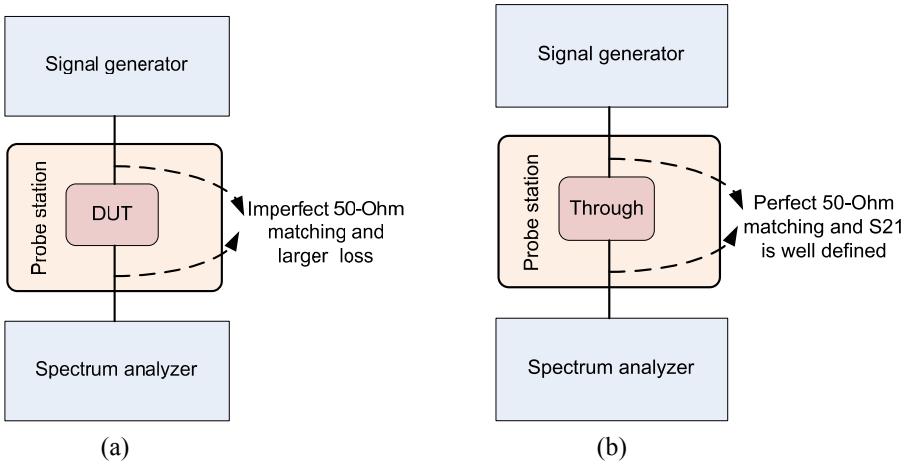


Fig. 5.36 The comparison between (a) losses in the real DUT measurement peripherals and (b) actually measured using a "through" structure.

The actual output mismatch is 6 dB, which is caused by the 90-fF bondpad capacitance [21] which is further transformed via the transmission line to the output of the IJLO. The S-parameters cannot be measured at 73 to 83 GHz by using the network analyzer, which only support measurements up to 67 GHz, but it can be estimated using the post-layout simulation results as follows.

The output impedance of the IJLO 'seen' by the pads becomes

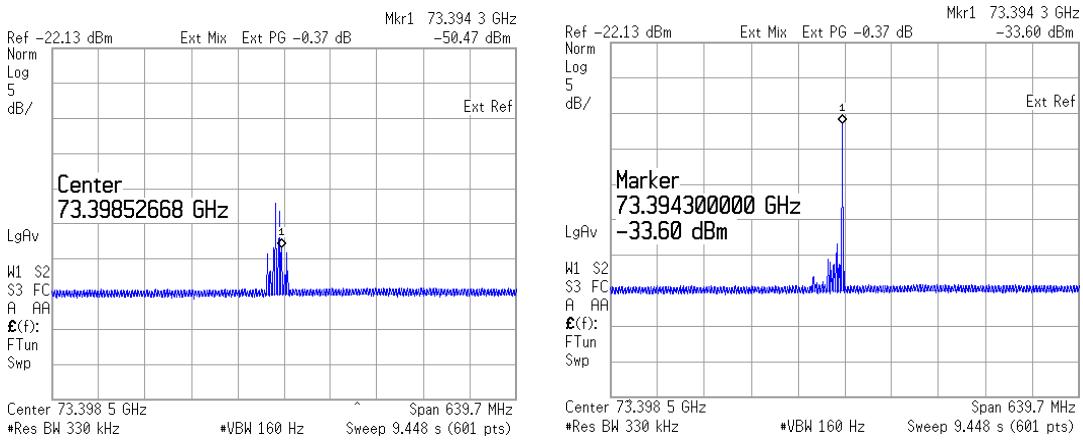
$$Z_{out,transformed} = 50 \cdot \frac{(71 + j46) + j \cdot 50 \cdot \tan(\beta l)}{50 + j \cdot (71 + j46) \cdot \tan(\beta l)} \tag{5.31}$$

where β equals $2\pi f/c$ and l is 250 μm. Adding the 90-fF parasitic capacitance from the pads. i.e. -49j Ω at 72 GHz, the final output impedance seen by the 50 Ω cable becomes about 5-26j Ω, which results in a S_{22} of about -1 dB. Compared to simulated S_{22}, about 6-dB power gets lost due to the mismatch.

Besides the measured cable loss and mismatch, the bandwidth of the coaxial connector, cables and baluns will cause extra mismatch, but they cannot be measured due to the equipment limitation and thus cannot be de-embedded. As a result, -10 dBm output power from the IJLO is a pessimistic estimation, and compared to the post-layout simulation results, the real number might be 6 dB higher. In this work, this conservative measurement output power level is reported, but in the near future, a THz signal generator and spectrum analyzer will be available in the lab and more accurate

measurements will be done to verify that.

The injection locking behavior is also tested. While the IJLO is injection locked, the output spectrum will be constrained around the injected frequency point and the phase noise of the IJLO is much lower, as shown in Fig. 5.37 (a) and (b). When the injection level is -60 dBm of input RF power, the single-step locking range is 5 MHz. It is able to achieve 55 MHz at -40 dBm injection level and 200 MHz at -35 dBm injection level (for the same reason, the mismatches and extra loss caused by measurement equipments are not de-embedded). It should be mentioned that the measured phase noise is about -80 dBc/Hz with 1 MHz frequency offset and -92 dBc/Hz after locking. Since high frequency phase noise measurement of the free-running IJLO using the spectrum analyzer is inaccurate, the data will be measured with a different approach and the details about this measurement will be specified in the WuRx measurement section.



(a) (b)

Fig. 5.37 The measurement results of the standalone IJLO: (a) free-running and (b) locked at the edge of one single-step locking range.

A comparison is shown in Table 5.8 of this IJLO and other injection-locked circuits, which are mostly injection-locked dividers at much lower frequencies, where the gain of transistors is higher, the parasitics are lower, and thus it is easier to achieve low-power operation and low phase noise feature.

Table 5.8 The performance comparison of IJLOs

	Technology	f_{RF} (GHz)	LR (GHz)	Sensitivity (dBm)	P_{DC} (mW)	VDD (V)	Effective Area (mm^2)	*FOM (dB)
[27] RFIC 09	0.12 µm SiGe	5.1	0.15 (3%)	-40	4.8	1.2	0.04	272
[28] CICC 09	0.18 µm CMOS	4.7	3 (64%)	-30	0.9	1.8	0.011	282
[29] ISSCC 09	0.13 µm CMOS	63	7.4 (11.7%)	-27	1.6	0.8	0.0165	291
[30] ESSCIRC 09	0.18 µm CMOS	1.92	-	-	55	-	0.031	-

[31] ISSCC 08	90 nm CMOS	85.2-96.2	11 (12.9%)	-15	3.5	1.2	0.3366	279.5
This work	65 nm CMOS	73 to 83	10 (13.7%)	-60	9	1	0.055	319

FOM= f_{RF}*LR/(Sensitivity* P_{DC})

It can be seen that this IJLO is able to achieve the maximum sensitivity of about -60 dBm, almost the largest (total) locking range of 73 to 83 GHz, and thus has the best FOM.

Passive mixer standalone testing results

The chip photo of the passive mixer is shown in Fig. 5.38.

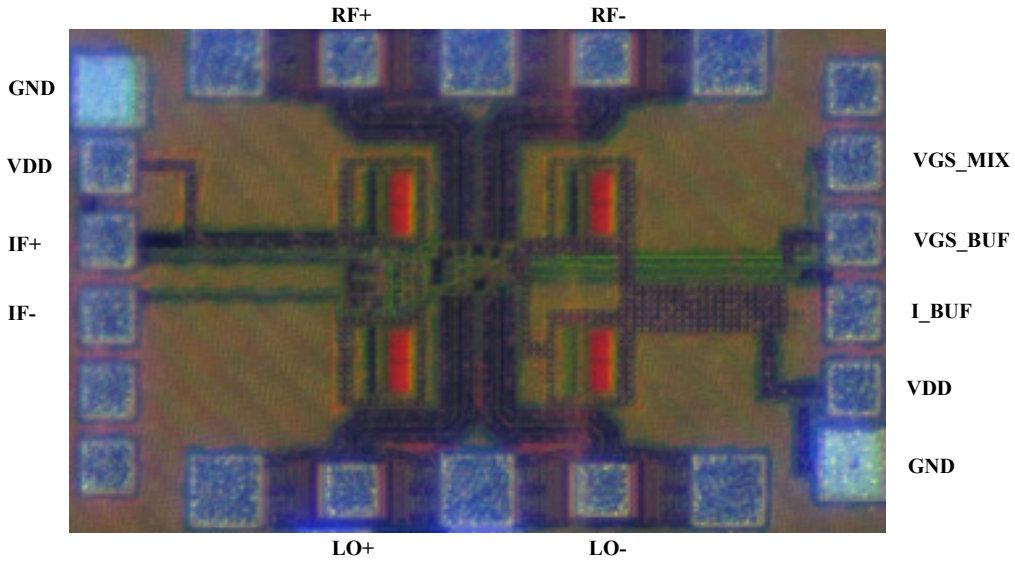
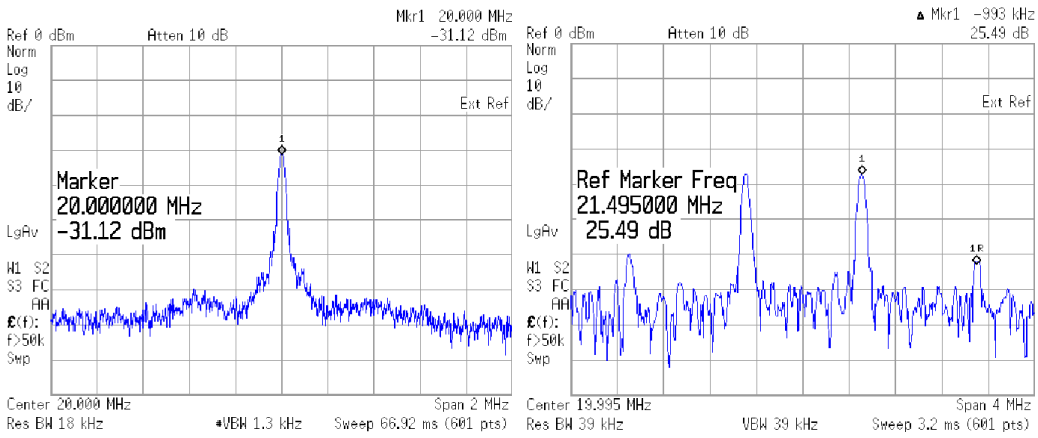


Fig. 5.38 The chip photo of the passive mixer.

The output spectrum of the standalone passive mixer is shown in Fig. 5.39. Despite a 2 dB loss and 6 dB mismatch caused by the measurement setups, the output power into a 50-Ohm load is about -23 dBm, and the voltage conversion gain from the 50-Ohm source to the 50-Ohm load is -3 dB at 70 GHz (including the buffer amplifier). The DC power consumption of the mixer (the buffer amplifier) is 1.4 mA from a 1-V supply. It should be noted that only the narrow-band mode can be tested, because the output terminal of the mixer is connected to the eye-pass pads which only have a bandwidth of 200 MHz. In principle, it has to be connected with the GSGSG RF pads just like the RF and LO input ports. The only reason of doing this is the lack of chip area for this work in the MPW wafer. The two-tone test is shown in Fig. 5.39 (b) and the measured IIP3 is -14.75 dBm, which is mainly limited by the buffer amplifier. Since the linearity requirement of the wake-up process is not high and the input level of the WuRx is normally low, e.g. from -60 to -30 dBm, this IIP3 level is quite acceptable.



(a) (b)

Fig. 5.39 (a) The output spectrum of the mixer and (b) the two-tone test spectrum.

W-band WuRx testing results

The chip photo of the WuRx is shown in Fig. 5.40. It can be seen that the chip area is bondpad limited and the effective area is actually about 0.072 mm^2. Similarly, the SW-CPW TLs are used at the input and output terminals to connect the circuit core to the GSGSG RF bondpad. In order to measure the output spectrum of the WuRx accurately, the inputs of the IJLO and the mixer are separated, whereas they should be connected as one input (as explained in Section 5.3.5). As shown in Fig. 5.40, "RF+" and "RF-" are for the IJLO injection while "RF1+" and "RF1-" are for the RF inputs of the passive mixer. In this way they can have a small frequency difference, e.g. 20 to 200 MHz (limited by the bandwidth of the output pads), and the down-conversion result can be read clearly from the spectrum analyzer.

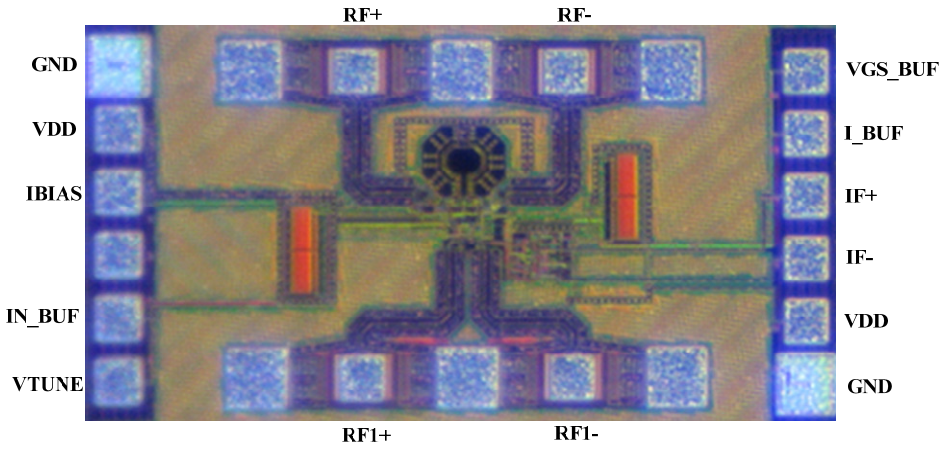
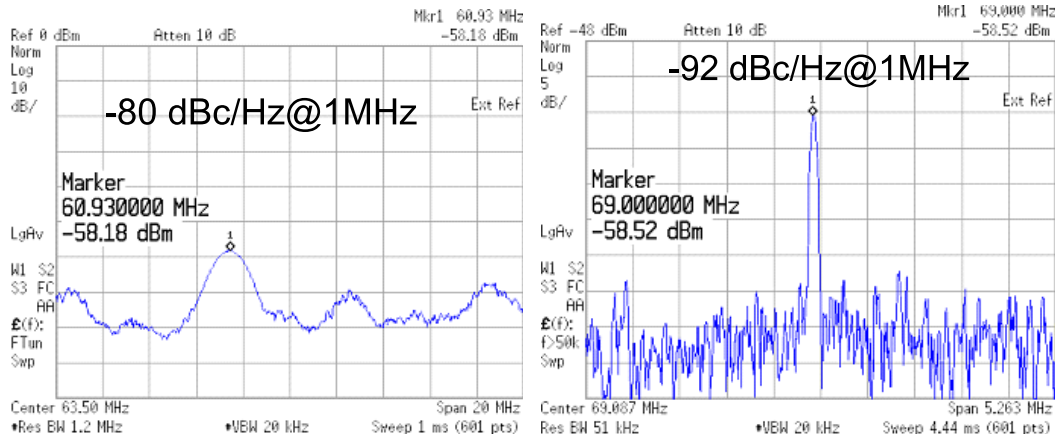


Fig. 5.40 The chip photo of the W-Band WuRx.

The output spectra of the WuRx with and without locking are shown in Fig. 5.41 (a) and (b) respectively. It can be seen that the phase noise of the IJLO is dramatically improved after locking.



(a) (b)

Fig. 5.41 The output spectra of the WuRx (a) free-running and (b) locked.

As discussed in the IJLO measurement section, the phase noise measurement of the 73-GHz free-running IJLO is inaccurate by using the spectrum analyzer. This problem can be solved by measuring the low-frequency output of the WuRx instead, because when mixing the RF signal down to the IF (here we say IF because a small frequency difference is left for the IJLO input and mixer input), the phase noise of the IJLO is also down converted. Although the mixer itself and output buffer amplifier will surely add some thermal noise, the result can still be a good approximation of the real IJLO phase noise. According to the definition, the phase noise can be read directly from the output spectrum by using the following equation

$$PN(1MHz) = 10 \cdot \log\left(\frac{P_{signal} - P_{noise,1MHz}}{BW_{res}}\right) \tag{5.32}$$

where P_{signal} is the output IF signal power, P_{noise},1MHz is the noise power level at 1 MHz distance and BW_{res} is the resolution bandwidth of the spectrum analyzer.

The measured tuning range is from 70.86 to 79.29 GHz. As expected, both center frequency and tuning range decrease a bit due to the capacitive loading effect from the passive mixer. The total power consumption of the WuRx is about 10.2 mW (9 mW for the IJLO and 1.2 mW for the mixer) from a 1-V supply. Additional 4-mW power is consumed for the input buffer. Due to the lack of a high-frequency oscilloscope, the time-domain performance, e.g. the settling behavior cannot be observed in experiment and only the simulated results are provided here. The simulated single-step settling time

(with OOK modulated data with 10-ps rise time) of the WuRx is about 140 ps. At the worst case, when the IJLO freely runs at 70.86 GHz and the injected signal is at 79.29 GHz, the simulated total initial settling time is 8.5 μs. The performance is summarized in Table 5.9.

Table 5.9 Measurement results of the WuRx

	Measured Data
Technology	65-nm CMOS
Frequency (GHz)	70.86 to 79.29
Power consumption (mW)	10.2
Bandwidth (GHz)	8.43
IIP3 (dB)	-14.75
Single-step settling time (ps)	140 (simulated)
Worst-case initial settling time (μs)	8.5 (simulated)
Effective die size (mm^2)	0.072
Energy per bit (pJ/bit)	10.2 (with 1-Gbps wake-up bits)

5.5 Conclusions and follow-up

The design challenges and bottlenecks of the mmW low-power WuRx have been discussed in this chapter. A new mmW WuRx architecture is proposed and analyzed in terms of implementation technology, cost, latency and power consumption. Detailed circuit design and implementation of the proposed W-band WuRx have been presented, followed by the on-wafer measurement results of the circuits implemented in the TSMC 65-nm CMOS technology. These results are in good agreement with our theories, analytical models and simulation results. The measured W-band WuRx has a 8.4-GHz total tuning range (also known as total effective locking range according to the theory), i.e. from 70.9 to 79.3 GHz, which is able to cover the lower end of the W-band from 71 to 76 GHz. The total DC power consumption is about 10 mW from a 1-V supply. The simulated maximum circuit latency is about 8.5 μs.

Recalling the WuRx system design methodology as described at the beginning of Section 5.3.5, let us now finish step 7. Taking the measurement results into (3.45), the optimum DCF of the WuRx can be calculated as 0.86%, which results in an average power consumption of about 10 μW for the WuRx and Tx pair in the wake-up process from (3.46) when the event frequency is 1000 times/day. The completed WuRx architecture together with the digital controls are illustrated in Fig. 5. 42.

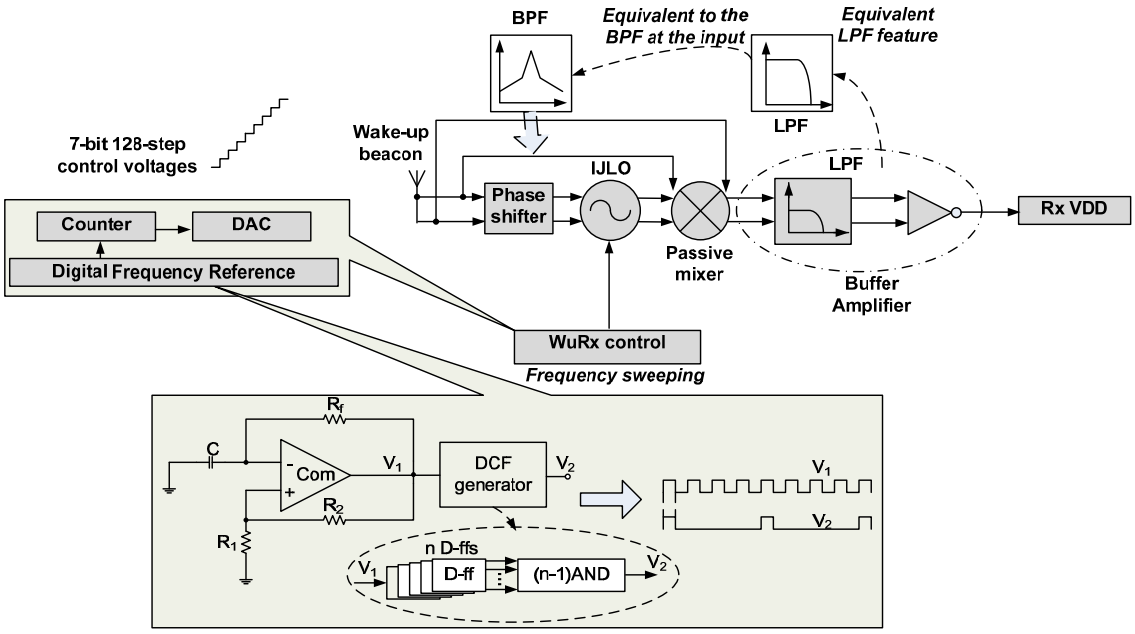


Fig. 5.42 The complete W-band low-power WuRx system.

References

[1] Ilker Demirkol *et al*., "Wake-up Receivers for Wireless Sensor Networks: Benefits and Challenges," in *IEEE Wireless Communications*, vol. 16, no. 4, pp. 88-96, 2009.

[2] Lin Gu *et al*., "Radio-Triggered Wake-Up for Wireless Sensor Networks," in *Real-Time Systems*, vol. 29, no. 2-3, pp. 157-182, 2005.

[3] Joongyung Lim et al., "High Sensitive RF-DC Rectifier and Ultra Low Power DC Sensing Circuit for Waking Up Wireless System," in *Asia Pacific Microwave Conference*, 2009.

[4] Xiongchuan Huang *et al*., "A 1.4GHz/925MHz 51μW Wake-Up Receiver with Offset and Noise Suppression," in *IEEE Solid-State Circuits Conference (ISSCC)*, 2010.

[5] Nathan M. Pletcher *et al*., "A 52 μW Wake-Up Receiver With -72 dBm Sensitivity Using an Uncertain-IF Architecture," in IEEE Journal of Solid-State Circuits, vol. 44, no.1, January 2009.

[6] The International Technology Roadmap for Semiconductors, 2009.

[7] Usha Gogineni *et al*., "RF power potential of 45 nm CMOS technology," in *IEEE 10th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF2010)*, 2010.

[8] Ring oscillator, www.wiki.org, 2011.

[9] R. Adler, "A Study of Locking Phenomena in Oscillators," in *proceedings of the IEEE*, vol. 61, no. 10, October, 1973.

[10] Prakteek Bhansali, *Gen-Adler: The Generalized Adler's Equation for Injection Locking Analysis in Oscillators*, presentation in University of Minnesota.

[11] Kwan Wai Li *et al*., "Low Power Injection Locked Oscillators for MICS Standard," in *IEEE Biomedical Circuits and Systems Conference*, 2009.

[12] Y.-H. Chuang *et al*., "A Wide Locking Range and Low Voltage CMOS Direct Injection-Locked Frequency Divider," in *Microwave and Wireless Component Letters*, vol. 16, no. 5, 2006.

[13] Yu-Hang Wong *et al*., "A 50-to-62GHz Wide-Locking-Range CMOS Injection-Locked Frequency Divider with Transformer Feedback," in *IEEE Radio Frequency Integrated Circuits Conference*, 2008.

[14] Jun-Chau Chien, "Design of Wide-Tuning-Range Millimeter-Wave CMOS VCO With a Standing-Wave Architecture", in *IEEE Journal of Solid State Circuits*, Vol. 42, pp 1942-1952, Sept 2007.

[15] Xia Li *et al*., "A lower-power, high-sensivity injection-locked oscillator for 60 GHz WPAN applications," in *PrimeAsia 2009*, pp. 161-164, November 19-21, 2009.

[16] Chaki Inui, Ivan Chee Hong Lai, Minoru Fujishima, "60 GHz CMOS Current-Reuse Cascade Amplifier," in Proc. Asia-Pacific Microwave Conference 2007, pp. 793-796.

[17] Sakian Dezfuli, P. *et al*., "A 57-63 GHz quadrature VCO in CMOS 65 nm Technology," in *European Microwave Conference*, 2009.

[18] Behzad Razavi, *Design of Analog CMOS Integrated Circuits*, 2005.

[19] Tomas Lee, Oscillator Phase Noise: A Tutorial.

[20] Behzad Razavi, "A Study of Injection Locking and Pulling in Oscillators," in *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 39, No. 9, pp. 1415-1424, 2004.

[21] Paul van Zeijl et al., "Modeling and Measurement on Minimum-Width Transmission-Lines from 10-67

GHz in 65 nm CMOS," in *IEEE Custom Integrated Circuits Conference*, 2010.

[22] Digital Counter, www.wikipedia.com, 2011.

[23] Un-Ku Moon *et al*., "A Switched-Capacitor DAC with Analog Mismatch Correction," in *IEEE International Systems and Circuits Conference*, 2000.

[24] Yanyu Jin et al., "A Low-Power Low-Voltage Millimeter-Wave Transconductance Mixer," in ProRISC 2010.

[25] Sining Zhou *et al*., "A CMOS Passive Mixer With Low Flick Noise for Low-Power Direct-Conversion Receiver," in *IEEE Journal of Solid-State Circuits*, vol. 40, no. 5, pp. 1084-1093, May 2005.

[26] Krenar Komon *et al*., "Fundamental Performance Limits and Scaling of a CMOS Passive Double-Balanced Mixer," in IEEE Circuits and Systems and TAISA Conference, 2008.

[27] Gathman, T.D. *et al*., "Injection locked oscillator arrays for spectrum analysis," in *IEEE Radio Frequency Integrated Circuits*, 2009.

[28] Yung-Chung Lo *et al*, "A 1.8V, sub-mW, over 100% locking range, divide-by-3 and 7 complementary-injection-locked 4 GHz frequency divider," *in IEEE CICC*, San Jose, 2009.

[29] Sujiang Rong *et al*., "0.9mW 7GHz and 1.6mW 60GHz Frequency Dividers with Locking-Range Enhancement in 0.13μm CMOS," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2009.

[30] Yuka Kobayashi *et al*., "A low-phase-noise injection-locked differential ring-VCO with half-integral subharmonic locking in 0.18 11m CMOS," in *ESSCIRC*, Athens, 2009.

[31] K.H. Tsai et al., "3.5mW W-Band Frequency Divider with Wide Locking Range in 90nm CMOS Technology," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2008.

CHAPTER 6 PHASED ARRAY RECEIVER DESIGN

As described in Section 3.2.1, lower E_{bit} value for wireless communication is rendered possible at higher frequency bands by utilizing extra gain of directional antennas which can be realized by the multi-element antenna array beamforming technology[30]. The reason behind it is that the RF power is confined in a narrow beam and the equivalent isotropic radiation power (EIRP) is hence increased. Since this part of gain augmentation is obtained without consuming extra DC power, the energy efficiency of the transceiver is improved. In 4.1.1, a multi-path front-end architecture is proposed as the main transceiver of the wireless wire system based on the above conclusion. In this chapter, the feasibility and effectiveness of the wireless wire directional communication system are discussed. The relevant circuit designs and simulation results of a 60-GHz phased array receiver are given as an example.

6.1 Discussions on beamforming

In general, beamforming usually refers to a signal processing technology that is used to direct the reception or transmission signal power of a receiver or a transmitter in a chosen direction. In a RF front-end, beamforming can be realized by using a multi-element antenna array. Despite diversified implementation details, the general principles of beamforming antenna array can be explained in the following.

As illustrated in Fig. 6.1, both the Tx and the Rx consist of n antenna elements that are separated with a distance d between two adjacent antenna elements. The Rx is located in the direction θ_{in} "seen" by the Tx. When the RF signal is sent from the Tx to the Rx, the signal from path i will arrive slightly earlier than the one from path i+1 by

$$\Delta\tau = \frac{d \cdot \sin\theta_{in}}{c} \tag{6.1}$$

where c is the speed of light.

[30] In fact, the multi-input-multi-output (MIMO) technology can be used to provide extra antenna gain through beamforming or to improve antenna diversity and channel throughput. The methodologies are different either in front end configurations or baseband algorithms [1]. In this work, the antenna array is adopted for the proposed mmW (60-GHz and W-band) wireless wire front-end for which the bandwidth is currently not an issue to support 1 to 2 Gbps peak data rate as required by the applications. As a result, it mainly targets at adding antenna gain.

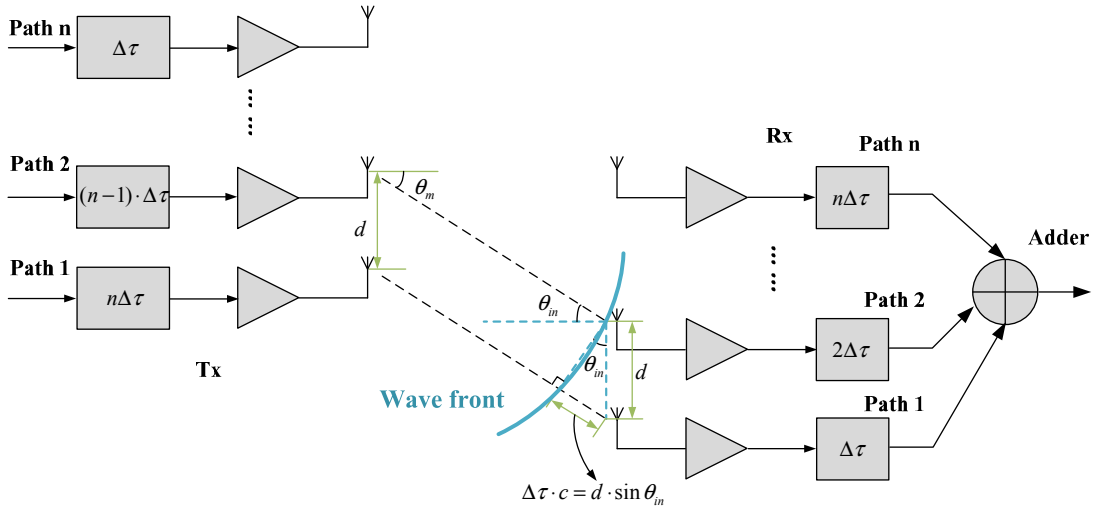


Fig. 6.1 Basic architecture of a beamforming transceiver array.

If we deliberately produce an extra time delay $\Delta\tau$ of the signal in path i compared to path i+1, the wave fronts of these RF signals sent by each antenna of the Tx will coincide with each other in the direction θ_{in}. Consequently, a peak antenna gain will be obtained in the direction θ_{in} and the resultant output power will become n-times larger compared to a single-path Tx if the total LO output power is fixed. This power can be understood as the EIRP, i.e. output power of the Tx plus antenna gain. Likewise, on the Rx side, if we compensate the time delays produced in the Tx (e.g. $i{\cdot}\Delta\tau$) by adding reverse delays (e.g. n-$i{\cdot}\Delta\tau$) and then add the received signals from each antenna element coherently, n-times larger Rx antenna gain will be obtained in the direction θ_{in}, and in other directions, the power from each path will be partially cancelled with each other. The additional antenna gain of the array over the power gain of a single-path front end is defined as the array factor (AF) and it is given by

$$AF(\Delta\tau) = \left[\frac{\sin\left(\frac{n\cdot\left(\omega\cdot\Delta\tau - \frac{\omega\cdot d}{c}\cdot\sin\theta_{in}\right)}{2}\right)}{\sin\left(\frac{\omega\cdot\Delta\tau - \frac{\omega\cdot d}{c}\cdot\sin\theta_{in}}{2}\right)}\right] \tag{6.2}$$

where ω is the angular frequency of the RF signal [2]. It can be seen that the AF achieves its maximum value only when (6.1) is satisfied. If the time delays in the Rx paths cannot perfectly compensate the delays induced by the angle of incidence, the resulting AF will decrease. It will affect the link budget parameters, and therefore needs to be carefully evaluated. The influence and potential solutions are discussed in the following sections.

One of the most important features of a multi-antenna beamforming array is the SNR improvement at the output of the adder of Rx front end. As discussed, in the adder of the

Rx, signals from each path can be coherently added, which results in an extra gain of $10 \cdot \log(n)$ dB, i.e. the gain of the antenna array. On the other hand, the noise from each path is uncorrelated, i.e. generated by independent sources, and thus it will experience no extra antenna gain. The output SNR of the Rx is therefore improved by a factor of n in an n-element beamforming front end.

6.1.1 Benefits and potential problems of beamforming arrays

Based on the above discussion, the benefits of using the multi-element antenna array beamforming are summarized in Fig. 6.2. It should be noted that the influences of mismatches or beam-pointing error are not yet included.

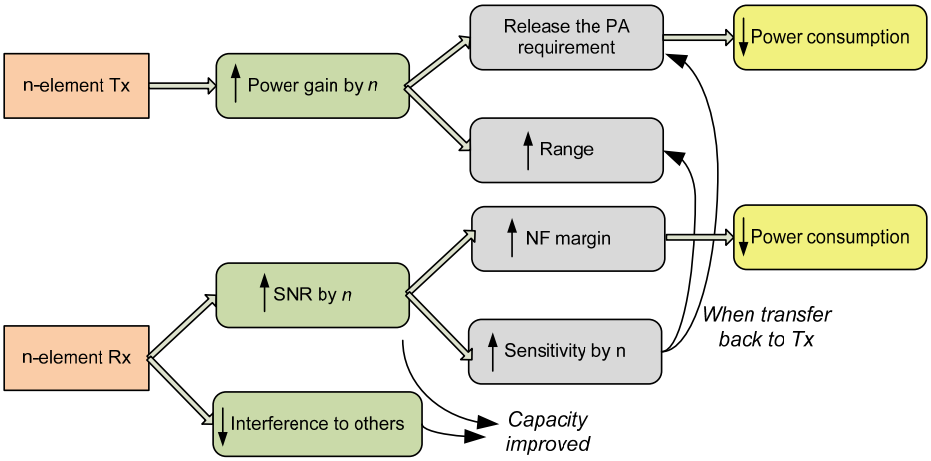


Fig. 6.2 Benefits of beamforming array.

Assuming the output power of the LO is fixed in an n-path Tx compared to a signal-path Tx, an n-element antenna array will introduce n-times higher antenna gain. This extra antenna gain can be used in two ways: first, assuming the total EIRP of an n-path Tx is identical to the output power of a single-path Tx, the required output power of each path will be decreased by n. As a result, the output power requirement of each power amplifier (PA) gets relaxed, which will eventually reduce its DC power consumption if the power-added efficiency (PAE) of the PA is unaffected[31]. In the other situation, assuming each path of the n-path Tx stays unchanged, i.e. the same PA and driver stages, the EIRP of the Tx array will becomes n-times larger compared to the single-path Tx. In this way, the extra antenna gain can be used to combat the pathloss and increase the

[31] With n-times smaller input and output powers, the PA transistors can be re-dimensioned (e.g. smaller transistors, lower currents) resulting in the same operating point. In practice, PAE might be affected either by the limitation of device scaling (very small input and output powers may require the minimum transistor size), or by the limitation of passive power combining (very large power may result in very high impedance transform ratio and extra loss) [3].

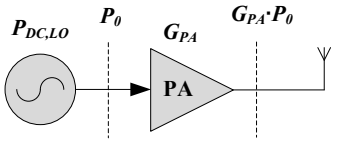
communication range. Since our proposed wireless wire system targets at low power performance, the first usage of the antenna gain is chosen in our work. In an *n*-path Rx, *n*-times better SNR can be used either to relieve the NF requirement under the same link model or to improve the Rx sensitivity by a factor of *n* with the same Rx circuits. The former is helpful to reduce the Rx power consumption while the latter can be transformed back to the Tx and thus relieves the PAs or extends the communication range. In the above arguments, however, two important issues that may potentially impair the benefits of antenna array need to be further discussed:

A. Tx power efficiency

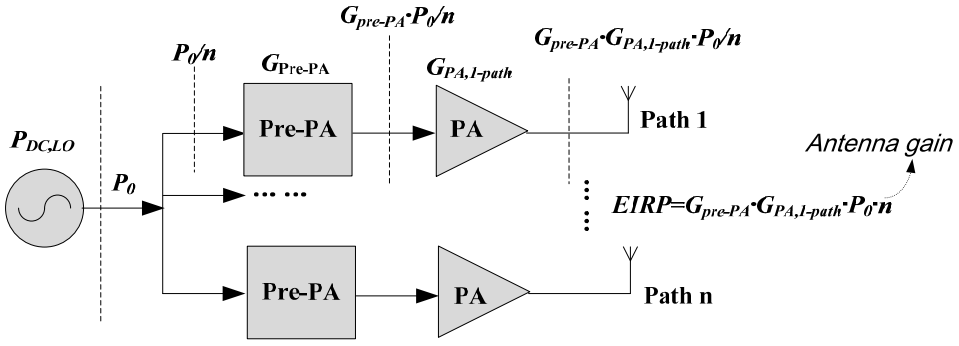
As a metric of the overall power performance, the efficiency of a Tx can be defined as

$$\eta_{Tx} = \frac{P_{out}}{P_{DC,Tx}} \tag{6.3}$$

where P_{out} is the total output power, and in a *n*-path Tx, it refers to the total EIRP. The power budgets of an *n*-path Tx and a single-path Tx are illustrated in Fig. 6.3 (a) and (b), respectively, where $P_{DC,LO}$ is the power consumption of the LO, P_0 is the output power of the LO, and G_{PA} and $G_{PA,1\text{-}path}$ are the power gain of the PA in the single-path Tx and a single path of the *n*-path Tx, respectively.



(a)



(b)

Fig. 6.3 The power budget illustrations of (a) a single-path Tx and (b) an *n*-path Tx. The pre-PA stage refers to the overall effect of other blocks like the delay element or PA drivers that are exclusively required in the *n*-path Tx array.

It can be seen that in order to keep the same output power level, the following equation must be satisfied, i.e.

$$G_{pre-PA} \cdot G_{PA,1-path} = \frac{G_{PA}}{n} \tag{6.4}$$

In each path of the Tx array, the output power is reduced by n^2 compared to the single-path Tx. For calculation simplicity, we assume the power gain of the pre-PA stages to be equal to unity, i.e. they are used to provide the required time delays and thus their power consumptions are only related the amount of time delay and are not affected by the input or output power levels[32]. The efficiencies of the single-path and the *n*-path Txs can be expressed as

$$\eta_{Tx,single-path} = \frac{G_{PA} \cdot P_0}{P_{DC,LO} + P_{DC,PA}} \tag{6.5}$$

and

$$\eta_{Tx,n-path} = \frac{G_{PA} \cdot P_0}{P_{DC,LO} + n \cdot P_{DC,pre-PA} + \frac{1}{n} \cdot P_{DC,PA}} \tag{6.6}$$

It can be seen that if we can generate the required time delay for beamforming with zero power consumption, an infinite number of antennas will be the most power efficient option for the *n*-path Tx. However, it is not feasible in reality. Defining the power efficiency improvement factor (PEIF) as

$$PEIF = \frac{P_{DC,LO} + P_{DC,PA}}{P_{DC,LO} + n \cdot P_{DC,pre-PA} + \frac{1}{n} \cdot P_{DC,PA}} \tag{6.7}$$

it can be seen that PEIF will be larger than 1 only if

$$\frac{P_{DC,pre-PA}}{P_{DC,PA}} < \frac{n-1}{n^2} \tag{6.8}$$

Otherwise the benefit of power reduction by using an *n*-path Tx array will be tampered due to the extra power consumed by the time delay (or phase shift) blocks.

B. Critical issues for SNR improvement

On the Rx side, the SNR improvement is influenced by the system and circuits accuracy significantly. The maximum antenna gain is obtained only if the beam-pointing angle of the Rx, θ_m, is identical to the angle of incidence, i.e.

---

[32] Class A active devices with 0 dB insertion loss are assumed and thus the power dissipation level does not depend on signal powers. This is a moderate assumption particularly for mmW systems.

$$\theta_m = \sin^{-1}(\frac{c \cdot \Delta\tau}{d}) \tag{6.9}$$

and any beam-pointing error will lead to Rx antenna gain degradation and thus impair the SNR improvement feature. The potential causes of errors are illustrated in Fig. 6.4.

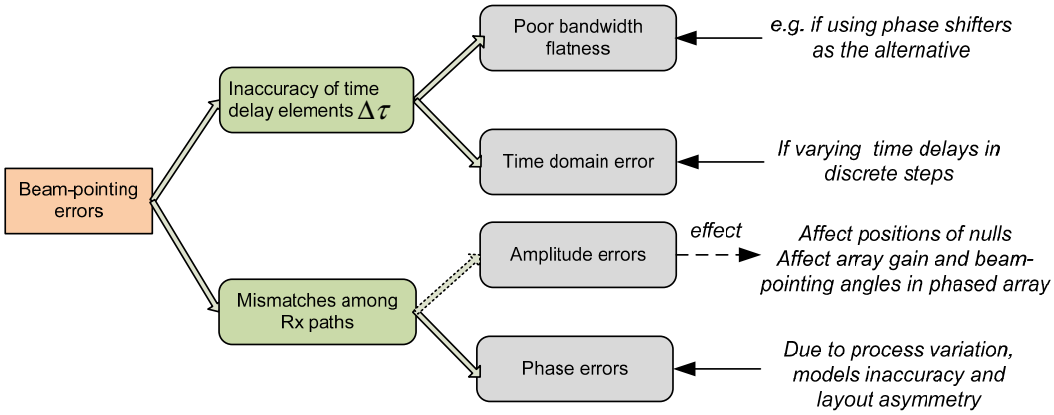


Fig. 6.4 Potential causes of beam-pointing errors.

These causes of beam-pointing errors will be discussed in the following sections.

6.1.2 True time delay array and phased array

True time delay elements are characterized by constant group delay over frequency, and thus their performance is independent of the signal bandwidth. On the contrary, alternative elements like phase shifters can only generate constant phase shifts over the targeted frequency range and therefore are virtually narrowband. However, variable true time delay elements are very difficult to be implemented in silicon, because they are normally lossy, bulky and thus costly, since they typically employ transmission line structures such as the trombone line and the varactor-loaded TL delay elements [2]. These problems can be solved by using (partially) active phase shifters, because they can be implemented by transistors and lumped devices, and therefore are much more compact and exhibit low loss[33]. Although the spatial filtering function requests high linearity of phase shifters to accommodate strong interference rejection, active phase shifters are still very promising candidates for beamforming array. The most important question becomes "can we use these narrow band phase shifters in the mmW beamforming array" and "is the beam-pointing error acceptable"? Let us now take the 60 GHz frequency band as an example for the phased array main receiver design.

In a phased array, two parameters should be emphasized:

- The first one is the distance *d* between adjacent antennas. Since the equivalent

[33] Gain can be obtained of course at the expense of extra power consumption.

aperture size D of a linear antenna array can be approximated as $(n\text{-}1)\cdot d$, a smaller d will reduce the antenna array size and thus reduce the antenna directivity while a larger d may cause multiple main lobes. In a narrow band system, the distance is normally fixed as the half wave length ($\lambda/2$) so that the antenna beamwidth are kept approximately constant over the entire band and it is given by

$$beamwidth = \frac{\lambda}{D} \approx \frac{\lambda}{(n-1)\cdot d} = \frac{2}{n-1} \tag{6.10}$$

However, it is not the case in a wideband system. For example, the maximum wavelength (i.e. λ_{max} at 57 GHz) of the 60 GHz band is 12.28% larger than the minimum wavelength (i.e. λ_{min} at 64 GHz). If the distance between adjacent antennas is fixed as $\lambda_{max}/2$, the beamwidth at 64 GHz will be 12.28% narrower than the one at 57 GHz.

- The second one is the angle of incidence. Every single angle of incidence corresponds to a unit time delay if d is pre-fixed.

In the first step, let us fix the carrier frequency at 57 GHz and the antenna distance of half wavelength at 57 GHz. The relationship between angles of incidence and unit phase shift ($\Delta\Phi$) can be found by

$$\Delta\Phi = \omega\cdot\Delta\tau = \omega\cdot\frac{\lambda}{2\cdot c}\cdot\sin(\theta_{in}) = \pi\cdot\sin(\theta_{in}) \tag{6.11}$$

It can be seen that there is a fixed relation between $\Delta\Phi$ and θ_{in} at a single frequency point. The array factor of a 4-element antenna array can be plotted in Fig. 6.5.

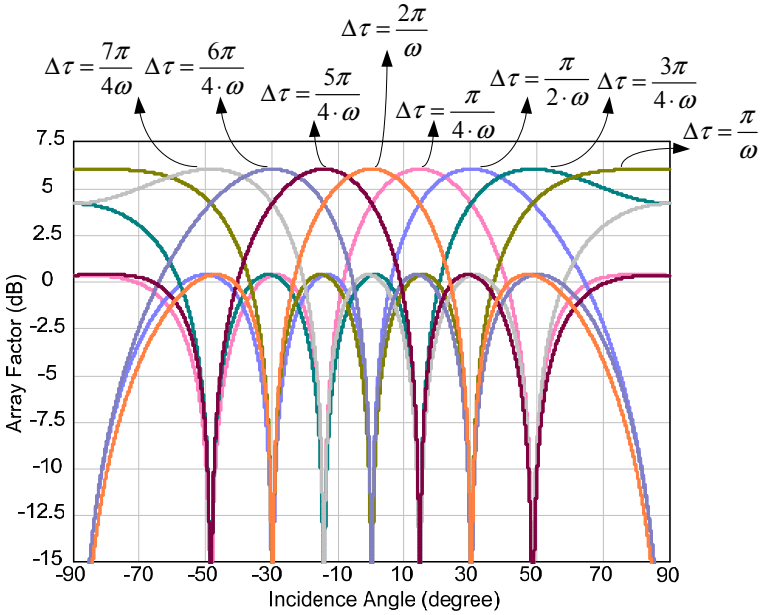


Fig. 6.5 Array factor of a 4-element antenna array at 57 GHz. Similar figure can be obtained at all frequency points in the 60 GHz band.

Now let us fix the unit phase shift ΔΦ as the one at 57 GHz, and study the influence of frequency to the array factor. The results can be illustrated in Fig. 6.6.

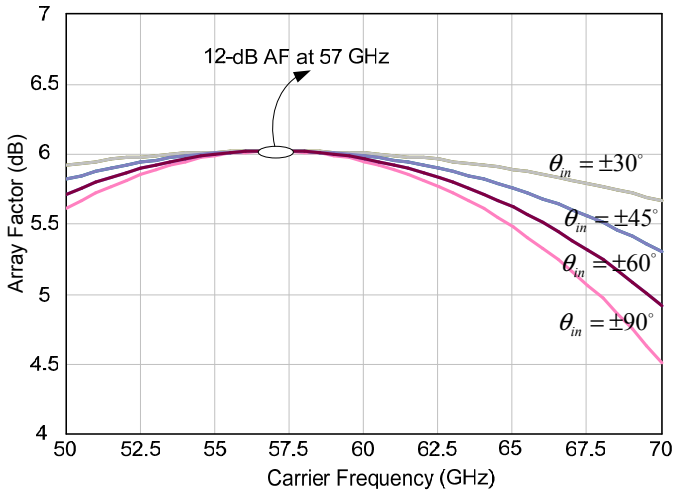
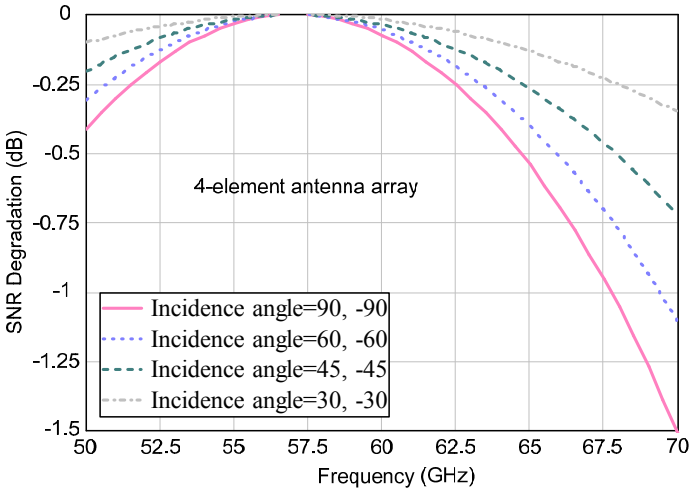
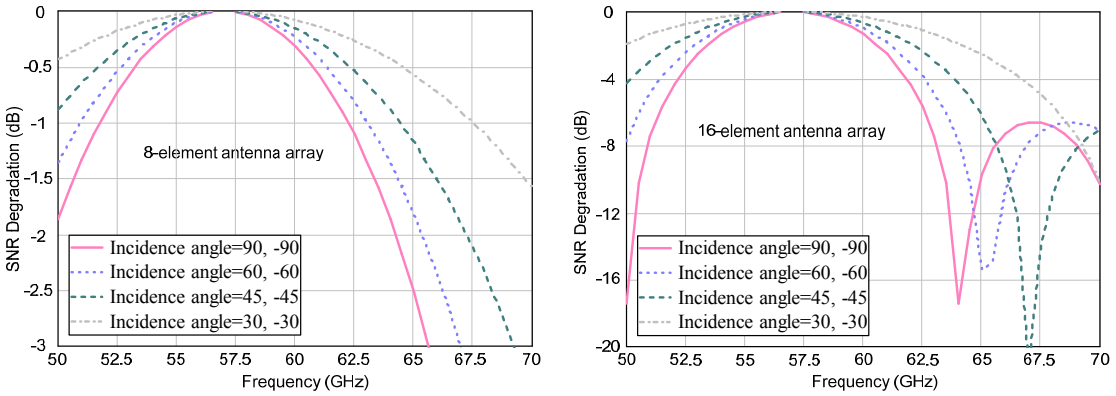


Fig. 6.6 Frequency behavior of AF of a 4-path phased array with unit phase shift target at 57 GHz.

It can be seen that when using phase shifters as an alternative of true time delays in a 60-GHz antenna array, the maximum drop of AF is about 0.5 dB, which occurs at ±90° angles of incidence and at the maximum frequency deviation. The direct consequence of degraded AF is less antenna gain and thus the link budget parameters will be affected. SNR degradations of 4-, 8- and 16-element phased array Rx front ends are illustrated in Fig. 6.7 (a), (b) and (c), respectively. It can be seen that in order to guarantee that phase shifters behave as a good approximation of true time delays across the entire 60 GHz band, no more than 8 elements should be used. As a result, a 4-element phased array is chosen in our work as the main receiver.



(a)



(b) (c)

Fig. 6.7 SNR degradation of a (a) 4-element array, (b) 8-element array and (c) 16-element array due to using of phase shifters.

Besides the frequency selectivity nature of phase shifters, system SNR will be impaired by the inter-symbol interference (ISI) as well, which would occur if the data rate is close to or higher than the signal bandwidth. In a phased array Rx, time delay errors will appear at each path and thus data pulses from each path will not overlap with each other in the baseband. If the data rate is high, i.e. the data symbol time is comparable to the time errors induced by phase shifter, the ISI problem will show up, and the situation will get worse if the number of elements is high. However, since our wireless wire system targets at 1 to 2 Gbps peak data rate and the bandwidth efficiency requirement is not stringent, this problem is not serious. If higher data rate is required, for example, in some other applications, like e.g. the high-definition data streaming, the data shaping technology or more complicated modulation scheme should be used to reduce the ISI [4].

6.1.3 Impact of discrete phase shifting

In Section 6.1.2, it is shown that the phased array front end can be used as an alternative for 60 GHz application as long as the element number is less than eight and the peak data rate is much smaller than the signal bandwidth, otherwise there will be a large SNR degradation that may cause communication failure. The actual situation is much more complicated because phase shifters tend to be digitally controlled in a discrete manner in order to reduce the power consumption, control-induced noise and design complexity. As a result, phase quantization error will appear in phased arrays, which will cause further SNR degradation.

According to (6.10), the antenna beamwidth of a 4-element array is about 2/3 rad or 38°. In order to avoid any blind area (the directions of beam that cause rapid SNR degradation) between two successive phase shift settlings, the beam-pointing resolution of a 4-element array should be no larger than approximately 1/3 rad, or 20°. If we neglect the frequency

influence, the relationship between the beam-pointing resolution ($\Delta\theta_m$) and the phase shifting resolution ($\Delta\Phi$) can be expressed as

$$\Delta\theta_m = \sin^{-1}\left(\frac{d\cdot\Delta\tau}{c}\right) = \sin^{-1}\left(\frac{d\cdot\Delta\Phi}{c\cdot\omega}\right) \approx \sin^{-1}\left(\frac{\Delta\Phi}{\pi}\right) \tag{6.12}$$

As a result, a 20° beam-pointing resolution results in a phase shifting resolution of π rad or 61°. This tuning step can be achieved with a 3-bit or finer digital phase control, because the phase shifting resolution is

$$\Delta\Phi = \frac{2\pi}{2^k} \tag{6.13}$$

where k is the number of binary control bits. Considering the frequency influence (bandwidth limitation) and other potential causes of SNR degradation, a 4-bit phase control is chosen for our 4-element phased array. Consequently, the phase shifting resolution is improved to 22.5° that results in a beam-pointing resolution of 7° when angles of incidence are less than 50 °, as shown in Fig. 6.8. In Fig. 6.8, differently colored lines are used to show the ratio of the actual output SNR (SNR_{ps}) of the 4-element phased array over peak SNR (SNR_{peak}) under different phase settings. It can be seen that one particular phase shift setting is related to one angle of incidence. When the actual angle of incidence has some deviation, SNR_{ps} will drop. The boundary angles (i.e. the ones at the crossing point of two subsequent phase setting) for the worst-case SNR are illustrated in Fig. 6.8 as well. The maximum SNR reduction caused by discrete phase shifting is only 5%. SNR reduction will get worse when increasing the element number or decreasing digits of control bits. In conclusion, the frequency selective nature of a phased array is the most dominating factor of SNR degradation compared to quantization errors. If a narrow beam is compulsory in certain applications, then the signal bandwidth should be reduced accordingly, otherwise the SNR degradation may deteriorate the benefits of beamforming.

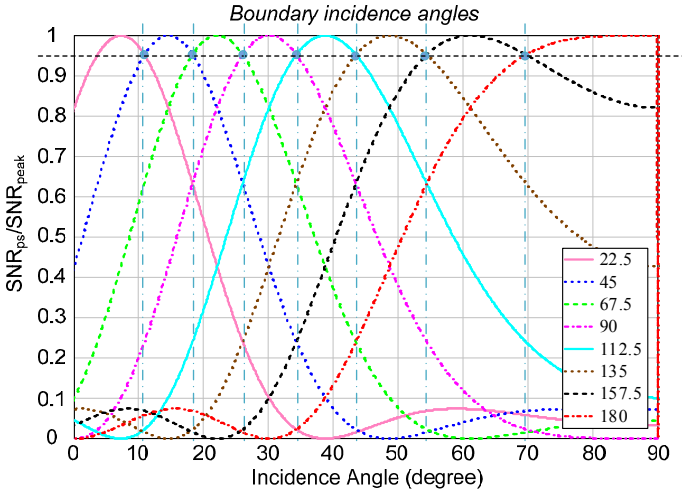


Fig. 6.8 SNR degradation due to discrete phase shifting in a 4-element phased array.

6.1.4 Position of phase shifting

Performance comparisons of several phased array architectures and differences in the position of the phase shifters are summarized in Table 6.1. Generic features are summarized based on discussions in [1] and [2] as well as the state-of-the-art 60-GHz phased arrays designs in [5-9].

Table 6.1 Comparison of different phase shifting positions

Position	Main advantages	Main Drawbacks	Area	Power
RF path [5] [6]	-Minimum RF blocks -Minimum area -Interference cancellation at RF and less dynamic range requirement of analog circuits	-Phase shifting inaccuracy -Phase shifter losses (affect NF and array gain) -Bandwidth limitation	P: Mod A: Low *	P: Low A: Mod
LO path [7]	-Minimum influence to FE gain and noise factor -Alleviated power gain requirement -Minimum bandwidth requirement -Easier design of low-loss IF power combiner	-Multiple high-power LO and/or lossy, bulky phase distribution blocks -Need accurate phase generation, e.g. a PLL or a tuned ring oscillator, etc. -Sensitive to mismatches and process variations for coupling oscillator arrays	P: Mod A: Low	P: High A: Mod
IF path [8]	-Simple and accurate circuits -Minimum loss of RF gain	-Larger size of passives -Multiple RF blocks -Stringent bandwidth requirement of RF blocks -Linearity requirement of RF blocks	Mod	Mod
Baseband [1]	-High versatility (e.g. MIMO, smart antenna) - High Accuracy - High robustness	-Maximum numbers of RF and analog blocks -Maximum linearity requirement of RF and analog blocks	High	High
Mixed analog/digital [9]	-Front-end blocks reduction compared to fully digital beamforming	-Linearity requirement -High power compared to RF beamforming scheme	High	High

* P=passive and A=active.

6.2 Main receiver architecture of the wireless wire system

Based on the above discussion, the 4-element phased array Rx with RF phase shifting is adopted in our main Rx, as illustrated in Fig. 6.9. In addition to the considerations about power consumption, cost, and bandwidth, another very important reason is that in the main Rx, the same IJLO-based self-mixing demodulation block (as in the WuRx) is to be re-used. In order to avoid LO pulling effect that affects injection locking performance, there should be no other oscillating circuits except the IJLO in the RF chain. The reasons we re-use the IJLO and the passive mixer combination are the following: (1) it costs low power compared to a conventional PLL-based LO; (2) it has high sensitivity for demodulation of an OOK signal; (3) it is very compact and wideband; (4) chip area (cost) is reduced by re-using blocks; (5) it generates very low flicker noise and (6) since frequency information is already obtained in the frequency sweeping phase of the WuRx, there is no need for the main Rx to sweep the entire band. Instead, it can be set directly into the correct frequency by the WuRx, and only needs to sweep a small region around that frequency in case of center frequencies deviation between those two IJLOs caused by frequency drifting.

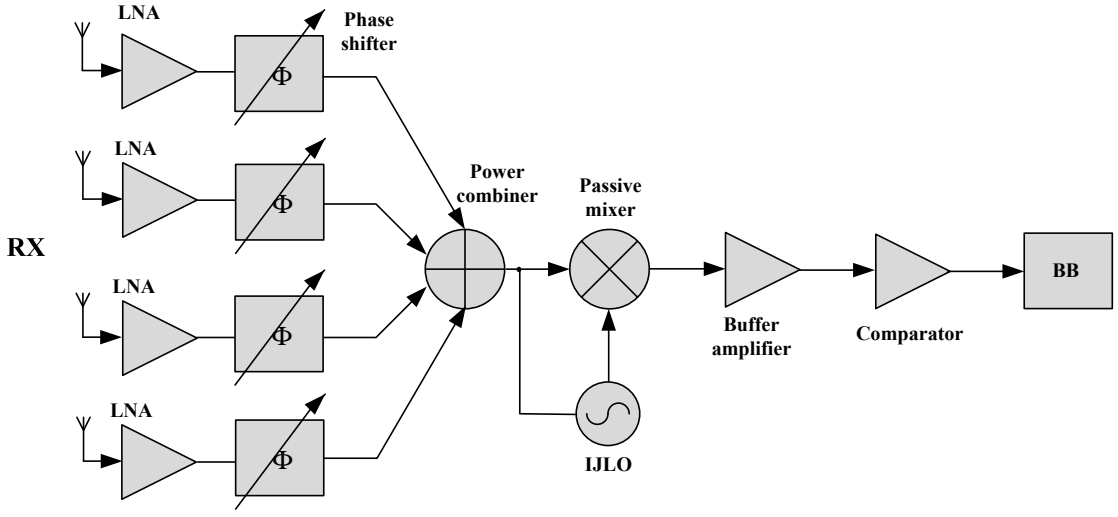


Fig. 6.9 A 4-element phased array self-demodulating Rx.

In fact, viable-gain amplifiers (VGAs) would be required in each path for side-lobe position control. This can also be beneficial to compensate the phase shifter induced loss in the RF chain. These blocks will be added to the 4-element phased array Rx in the next step. The proposed 4-path self-mixing Rx is simulated in ADS. The OOK-modulated 60 GHz input signal and the demodulation data (before the comparator) are shown in Fig. 6.10. The peak data rate of the OOK signal is 1 Gbps. It can be seen that about 10 dB

voltage gain is obtained. It should be noted that in this simulation, the IJLO is not yet included, and therefore the input and output signal still suffer from "square law". As discussed in 5.3.2, by using the IJLO-based self-mixing demodulation scheme, the front-end gain can be improved significantly.

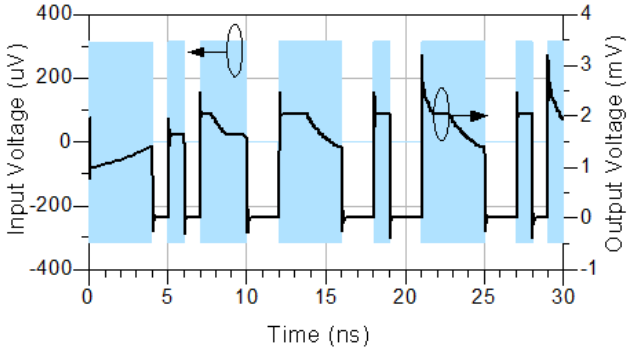


Fig. 6.10 Input and output signal amplitudes simulated in ADS.

Besides the influence of bandwidth, antenna number and angle of incidence, the SNR of a phased array is also affected by mismatches and phase shifting errors in each path. In fact, mismatches are inevitable because of process variation, model inaccuracy, layout asymmetry and non-ideal biasing. For instance, assuming that the phase shifting error in each path follows Gaussian distribution of $N(0, \delta_{path}^2)$, where δ_{path} is the variance of the phase shifting error, the variance of the beam-pointing error can be estimated by [2]

$$\delta_{beam}^{\ 2} = \frac{12 \cdot \delta_{path}^{\ 2}}{\pi^2 \cdot \cos\theta_m^{\ 2} \cdot (n-1) \cdot n^2 \cdot (n+1)} \tag{4.68}$$

where the worst case beam-pointing angle θ_m of 85° is used in the following analyses[34]. A rough estimation (not from the stochastic point of view) can be made as follows. With 5° maximum phase shifting error in each path, the resulting beam-pointing error is 9° in a 4-element phased array Rx. Compared to a 38.2° half-power bandwidth, 9° beam-pointing error would not cause obvious loss of antenna gain and thus not affect system SNR improvement. Compared to phase mismatch, gain mismatch has a more direct influence on the array factor and the positions of side-lobes. In addition, it is particularly harmful when a vector-sum type phase shifter is used in the array, because the amplitude mismatch in the I/Q phase generation will be translated to the phase error and deteriorate the beam-pointing error. In [1], it is calculated that in order to achieve less that 1.7 dB RMS gain error and 11.25° RMS phase error in a 4-bit vector-sum type phase shifter, the required I/Q signal generation accuracy should be better than 2.8 dB for the amplitude and 16° for the phase. Considering mismatches in the adder, these

[34] Due to the large beamwidth and poor beam pointing resolution when angles of incidence are close to 90°, the worst case scenario is estimated by 85°angle of incidence and 5° phase pointing error will not degrade antenna gain significantly.

specifications will be more stringent.

6.3 Design of 60-GHz low-noise amplifier

A low-noise amplifier (LNA) is normally the first stage of an Rx front end, i.e. it is located directly after the antenna. The main functions of the LNA are: (i) to provide RF gain; (ii) to offer noise matching and to ensure low noise figure (*NF*); (iii) to provide matching to antenna impedance for maximum input power; and (iv) to offer band-pass filtering. Therefore, it is a key block and has a significant influence on the overall receiver performance.

Along with the progress in wireless technology, wireless communication begins to enter into the mmW region for higher throughput. However, this trend introduces extra design difficulties in the LNA design too. Firstly, the existing LNA design theory is significantly extended and more complicated due to the non-neglectable parasitics that affect the gain and noise performance significantly. For example, in [10], very complicated signal-flow functions are derived by applying Mason's rule on the small-signal model of the proposed LNA. However, when corresponding these equations to the LNA circuit, the relations become very complicated and it is difficult to guide the LNA design. The extra complexity mainly comes from the Miller capacitor and a simple transformer feedback block. Along with the down-scaling process of implementation technologies and the increase of operation frequencies, this influence will become more significant due to higher levels of non-negligible parasitics. Secondly, at mmW frequencies, a mature and generic design methodology of an LNA is required. A certain design methodology is particularly derived for a specific type of LNA without offering a general model or design method. One example is given in [11] where an algorithmic design method for a two-stage 60-GHz LNA is well described, but it is hardly to be applied to a different LNA configuration.

In this section, an LNA design methodology is proposed by using a black-box approach. As an assumption, this methodology is derived in a power-constrained scenario. Parameter optimizations among gain, *NF* and linearity are discussed, followed by design examples of several 60-GHz low-power LNAs.

6.3.1 LNA design methodology

Since the LNA is a complicated component, especially at mmW frequencies and with constrained power dissipation, the design procedures should be sufficiently orthogonal and simple in order to provide an insight for the overall optimization. A minimum number of iterations should be kept. As a result, a structured design method is proposed.

A. System-level pre-optimization in power consumption

The system level pre-optimization of a transceiver should be discussed before designing the LNA. The following parameters are important to evaluate the performance of an LNA: gain (G), NF, input-referred third-order intercept point ($IIP3$), S-parameters, DC power consumption (P_{DC}), bandwidth (BW) and center frequency (f_0). Parameters like f_0 and BW are normally pre-determined and there is not so much freedom for further optimization at the system level. S-parameters basically define the small-signal gain, the input-output isolation and the matching conditions of the LNA, and they can be optimized at circuit level. Therefore, only G, NF, IIP_3 and P_{DC} are considered in the following model.

Instead of investigating a specific circuit, structure independent transforms (SITs) can be used for system-level parameters determination [12]. Based on different SITs models, a benchmark called power linearity factor is defined as

$$\kappa = \frac{P_{DC}}{G \cdot IIP3} \tag{6.14}$$

From (6.14), it can be seen that with certain technology, κ is a fixed value while G, P_{DC} and $IIP3$ can be treated orthogonally and a trade-off between G and $IIP3$ exists.

$$P_{DC} = \kappa \cdot G \cdot IIP3 \tag{6.15}$$

In order to maximize the overall design flexibility of a LNA, a two-stage circuit structure is assumed, which in principle is able to double the number of tunable circuit-level parameters. The overall noise figure (NF_{tot}) of such a cascaded LNA can be calculated as

$$NF_{tot} = NF_1 + \frac{NF_2 - 1}{G_1} \tag{6.16}$$

where NF_1, NF_2 are the noise figure of the first and second stages respectively and G_1 is the gain of the first stage. It can be seen that, in order to minimize NF_{tot}, NF_1 should be minimized and G_1 should be large enough. However, the boundary condition is not stated. From (6.15), denoting the total power of the two-stage LNA as P_{tot}, P_{tot} can then be modeled as

$$P_{tot} = \kappa_1 \cdot G_1 \cdot IIP3_1 + \kappa_2 \cdot G_2 \cdot IIP3_2 \tag{6.17}$$

where $IIP3_1$ and $IIP3_2$ are the $IIP3$ of the first and second stages, respectively, and G_2 is the gain of the second stage. The total gain of the LNA G_{tot} becomes

$$G_{tot} = G_1 \cdot G_2 \tag{6.18}$$

Substituting (6.18) into (6.17) and letting $\frac{\partial P_{tot}}{\partial G_1} = 0$, it can be obtained that

$$G_{1,opt} = \sqrt{\frac{G_{tot} \cdot \kappa_2 \cdot IIP3_2}{\kappa_1 \cdot IIP3_1}} \tag{6.19}$$

If G_{tot} is pre-determined by the system level calculation and therefore is taken as a design constant, increasing $IIP3_2$ and decreasing $IIP3_1$ are effective methods to enhance G_1, which will reduce the overall noise figure according to (6.16). However, since $IIP3_{tot}$ is calculated as

$$IIP3_{tot} = \frac{1}{\frac{1}{IIP3_1} + \frac{G_1}{IIP3_2}} \tag{6.20}$$

which is dominated by the lowest *IIP3* in the cascaded stages, the effect of decreasing of $IIP3_1$ should always be critically checked by (6.20).

Substituting (6.19), (6.20) into (6.17), $P_{tot,opt}$ can be written as

$$P_{tot,opt} = 2 \cdot \sqrt{\kappa_1 \cdot \kappa_2 \cdot G_{tot} \cdot IIP3_1 \cdot IIP3_1} \tag{6.21}$$

In the power-constrained applications, $P_{tot,opt}$ is limited by the total LNA power budget. As a result, the product of $IIP3_1$ and $IIP3_2$ can be determined in advance. Parameters of a two-stage power-constraint LNA can be chosen in the following way: (i) G_{tot} and P_{DC} are obtained from system level specifications; (ii) the product of *IIP3* parameters of two stages of the LNA can be determined by (6.21); and (iii) with this constraint, the optimum gains of the two stages can be obtained from (6.19) and (6.18).

B. The black-box method on circuit-level optimization

The optimization method described above mainly focuses on the system-level parameters. In this section, a black-box method is used to provide the generic solutions. As discussed in Section A, the two stages of the LNA can be optimized orthogonally with regard to different aspects. The first stage targets at optimizing NF and the gain (with the assumption of certain P_{DC}) and the second stage focuses on optimizing the distortion and input-output isolation. The overall LNA bandwidth is optimized as the last step.

The optimization procedures are as follows: (i) Tuning the current density of the first stage to the optimum current for lowest noise figure. For example, the optimum bias current density of TSMC 65-nm LVT NMOS transistor is shown in Fig. 5.4. When biased on the optimum point, e.g. 0.15 mA/μm, the transistor shows a lowest NF_{min}. (ii) Maximizing the gain of the first stage at the expense of lower $IIP3_1$ according to (6.19). Methods like g_m-boosting or current re-using can be used. In the meantime, the obtained $IIP3_1$ values should also be checked by (6.20). (iii) Optimizing $IIP3_2$ (according to (6.19)) and input-output isolation (S_{12}) in the second stage by neutralizing the circuit to eliminate

the miller effect. Since the well-known cascode structure is less effective in the mmW bands, enhanced cascode or other neutralization methods should be used (details will be discussed in the next section). (iv) The bandwidth of each stage can be traded with the DC power because it is related to the quality factor of the load LC tank at a certain center frequency. The overall bandwidth of the two-stage LNA can be extended by making the peaks of the responses of these two stages non-overlapping (twin-peaks). (v) Slight iteration (tuning $IIP3_1$) might be required if the $IIP3_2$ specification obtained from (6.19), (6.20) and (6.21) is non-realistic or cannot be easily realized. The black-box model is illustrated in Fig. 6.11.

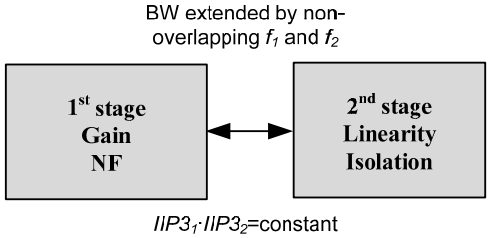


Fig. 6.11 Black-box optimization model.

6.3.2 Design examples of the 60-GHz low-power LNA

In this section, several 60-GHz low-power LNAs are designed in TSMC 65-nm CMOS technology as examples. As a prerequisite, the total power limit is set as 10 mW.

The first stage of the LNA should be featured with high gain and low noise simultaneously. It is proved that the common-source structure can achieve better noise figure and higher gain (at the same bias condition) compared to the common-gate structure as the input stage. Besides, with a source degenerative inductor and an input series inductor, the optimum power matching and noise matching can be achieved simultaneously. As a result, this configuration is chosen. Based on this basic structure, the gain enhancement methods are investigated.

Since the voltage gain of the LNA is linearly proportional to its transconductance, the first solution is boosting the g_m of the core transistors. The basic idea is to 're-use' the transistor nodes to get multiple inputs and obtain extra gain. This can be done by using a transformer or simply by cross coupling the source and gate nodes of the input transistors as shown in Fig. 6.12 (a) and (b). The gate-source voltage of each input transistor is increased to V_{IN+}-V_{IN-}, i.e. it is doubled. Theoretically speaking, the voltage gain is increased by 6 dB [13].

The other direction is to re-use the DC current as shown in Fig. 6.12 (c). In a cascode LNA structure, an inductor is inserted between the source of the upper transistor and the

drain of the lower transistor. This inductor is resonating with the gate capacitor of the upper transistor at the center frequency. In this way, the signal is coupled directly to the gate of the upper transistor and the 'common-gate' stage behaves like a common-source configuration, i.e. provides both voltage and current gains. The DC current is re-used in these two transistors. From (6.15), with certain P_{DC} budget, maximizing the gain will cause worse *IIP3*. A little iteration is required in this step to check that $IIP3_{tot}$ requirement is satisfied and thus $IIP3_1$ is sufficiently large. Under the "maximum allowable minimal $IIP3_1$", $IIP3_2$ can be calculated from (6.21).

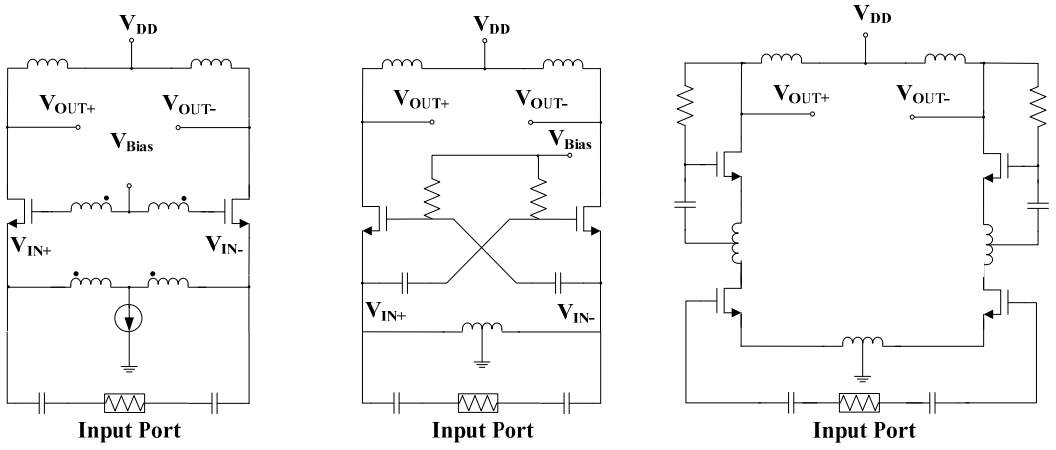


Fig. 6.12 LNA input stage: (a) (b) gm-boosted; (c) current-reused.

In the second stage, the linearity and unilateralization problems are focused on. According to (6.15), less gain is required compared to the first stage. Since the overall noise figure can be optimized due to the large gain of the first stage, the noise figure requirement of the second stage is less stringent, too, which leads to lower DC power consumption.

The most common method to enhance the unilateralization is to use the cascode structure. However, at 60 GHz, the cascode becomes less effective due to the parasitic capacitor between the source of the upper transistor and the drain of the lower transistor, which causes an extra pole to the transfer function. As a result, the gain drops dramatically and the input-output path is not well neutralized.

The first intuitive solution is to use an inductor to resonate out the parasitic capacitor and eliminate the pole. Two different configurations can be used as shown in Fig. 6.13 (a) and (b). In (a), an inductor is shunt-resonating with the parasitic capacitors, i.e. the source-bulk capacitor C_s and drain-bulk capacitor C_d. The DC block capacitor is not used in the differential structure because the DC voltages of two nodes of the inductor are identical. When looking into the LC tank, the impedance is high at the center frequency

and the signal will be fed into the upper transistor other than leak via the capacitance into the ground. The cascode behavior is then restored. In (b), an inductor is inserted between the source of the upper transistor and the drain of the lower transistor. An artificial transmission line is formed, which achieves the optimum power matching of the two internal stages and cancels the effect of the parasitic capacitor path too. One problem of solution (b) is that when the quality factor of the inductor is not high enough (which is true for standard CMOS on-chip inductor), the non-neglectable resistor exists in the signal path, which will lower that gain and add up extra noise.

The second method is very straightforward, i.e. shunting an extra inductor in parallel with the miller capacitor directly, as shown in Fig. 6.13 (c). In this case, the direct path "seen" at the input and output is blocked at the center frequency and no cascode is needed. With a lower count of active devices, the noise can be lower. Locating the extra inductor in the final layout is not easy. Normally a transmission line inductor is preferable here for its flexibility in shaping [13].

Some other neutralization techniques like the net charge cancelling or input-output transformer feedback are discussed in [10]. The net charge cancelling configuration requires a capacitor with precisely the same value as the miller capacitor, or otherwise the cancellation will be non-effective. However, it is very difficult to have a small non-linear capacitor in a standard CMOS technology. Besides, the parasitic capacitor normally has very large variation in its value due to layout, process and other reasons, so this method is not considered in our work. The transformer feedback technique is useful to neutralize the transistor as well as to extend the bandwidth significantly. However, it requires higher power consumption as a trade-off.

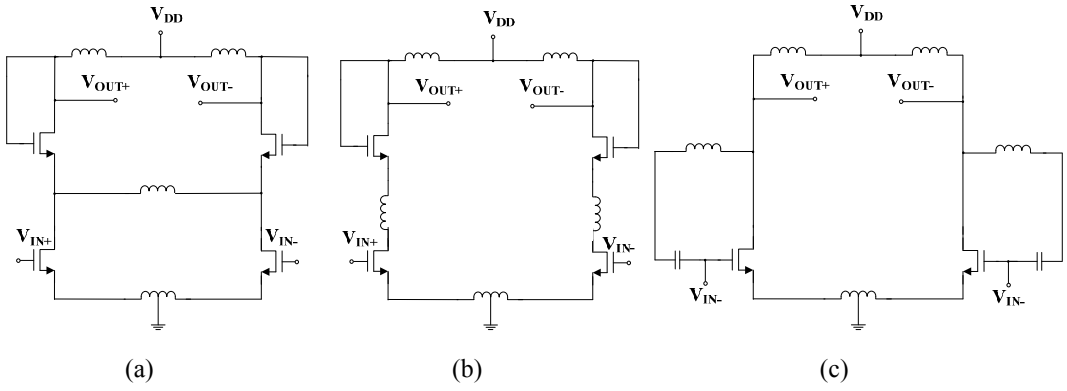


Fig. 6.13 LNA second stage: (a) middle inductor shunt peaking; (b) artificial transmission line based inter-stage matching; (c) shunt inductor based neutralization on miller effect.

The bandwidth of the LNA is tuned as the last step in the design by using the method described in Section B. More than 2 GHz 3-dB bandwidth can be achieved, which is compatible with the IEEE 802.15.3c standard.

Based on the methods discussed above, four 60-GHz LNAs are designed by using TSMC 65-nm CMOS models. Simulation results are shown in Table 6.2. It can be seen that all these designs are able to achieve about 15 dB power gain, 2 GHz *BW*, less than 5 dB NF, and about -4 dBm *IIP3* at a power consumption of 10 mW. It can be seen that, according to (6.15), under a certain level of κ, the total power budget can be further reduced at the expense of worse linearity performance. As a result, in the next step of research, the total power constraint of the LNA will be scaled down to 5 mW and in principle *IIP3* will be decreased by 3 dB.

Table 6.2 LNA Performance Comparison

	Gain (dB)	***BW* (GHz)**	***IIP3* (dBm)**	**P_{DC} (mW)**	**NF (dB)**	**FOM***
Fig. 6.12 (b)+6.13 (a) (simulation)	14	2	-4	9.6	4	0.69
Fig. 6.12 (c)+6.13 (a) (simulation)	19.8	2	-4	10	5	1.76
Fig. 6.12 (b)+6.13 (b) (simulation)	13	2	-4	9.6	4	0.55
Fig. 6.12 (b)+6.13 (c) (simulation)	15.2	2	-4	10	4.7	0.67
ISSCC 2011 [14]	18	2	-5	20.7	6	0.32
VLSI 2009 [15]	17.1	7	-6.9	21.4	8.2	0.09
MTT 2009 [16]	20.4	7	-12	65	8.6	0.02

*FOM=$(Gain \cdot IIP3)/(P_{DC} \cdot (NF-1))$

6.4 Design of 60-GHz phase shifters

In a phased array Rx front end with RF phase shifting, phase shifters are normally deployed as the successive stage of the LNA. The noise performance becomes less critical if the LNA gain is sufficiently high. However, their insertion loss, power consumption and phase shifting accuracy are very important. If its power loss is too high, the benefit of array gain will be deteriorated and the Rx sensitivity may be impaired due to a worse noise figure. Phase shifting inaccuracy leads to errors of the beam-pointing angle and therefore causes SNR reduction, which will eventually affect the Rx sensitivity and increase the Tx power consumption, because more output power is required. In the wireless wire application scenario, the peak power of the main Rx is less important, because it operates in a burst mode (high data rate with short active time) under control of the WuRx. However, we still have to reduce the peak power consumption in order to alleviate the power spike from the supply and maximize the battery life. The basic principle has been described in Section 4.1.3. In the following sections, we will discuss the classification and the design considerations of different types of RF phase shifters. Two design examples of 60-GHz phase shifters are given.

6.4.1 Phase shifters comparison

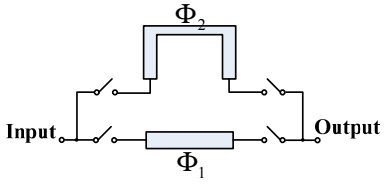
Different types of phase shifters are summarized in Table 6.3 and illustrated in Fig. 6.14. Their performance is discussed especially for mmW applications and for silicon technologies.

Table 6.3 Comparisons of several well-known RF phase shifters [1] and [17-19]

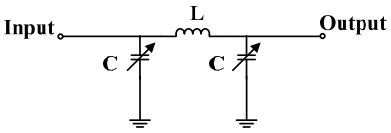
Type (Fig. 6.14)		**Pros**	**Cons**
Passive *	Switched line (a) $\Delta\Phi = \beta \cdot \Delta l$	-Straightforward design -Robust	- Bulky, lossy at mmW range - S_{21} fluctuation -Performance degradation due to poor mmW MOS switches
	Loaded line (b)	-Robust	-Lossy at mmW range - S_{21} fluctuation (in [1], S_{21} varies from -12 to -6 dB among different states) -Limited phase tuning range due to poor MOS varactor (normally smaller than π)
	Reflection (c)	-Difficult to achieve a large phase tuning range	- Bulky, lossy at mmW range - S_{21} fluctuation - Performance is limited by accuracy of quadrature phase generator
	Switched filter (d)	-Compact when implemented by lumped elements	-Lossy at mmW range - S_{21} fluctuation -Poor mmW MOS switches
Active	All-pass (e)	-(Relatively) low loss	-High power and high noise -Low dynamic range -Narrowband -Non-linear control - S_{21} fluctuation -Limited phase tuning range at mmW range
	Reflection (c)	-Low loss when using negative resistance as compensation -Able to achieve large phase tuning range	-Narrowband -(Relatively) high power -Non-linear control - S_{21} fluctuation -Potentially instable -Performance is limited by accuracy of quadrature phase generator
	Vector sum (f) $\Phi = \tan^{-1}(A_Q / A_I)$	-Low loss -Wideband -Compact -Constant S21	-(Relatively) high power -High noise -Low dynamic range

		-Constant power dissipation -Inherently suitable for accurate digital tuning	

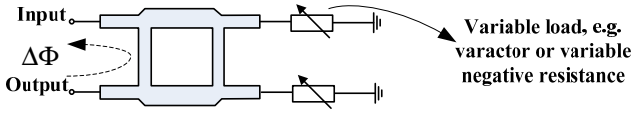
* These passive phase shifters can be implemented in distributed or lumped elements. For simplicity, only one of the versions is illustrated.



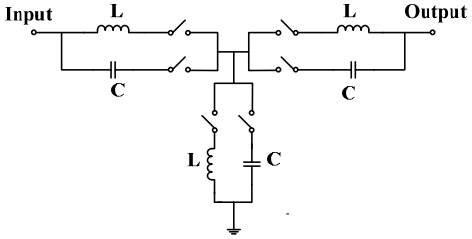
(a) Switched line;



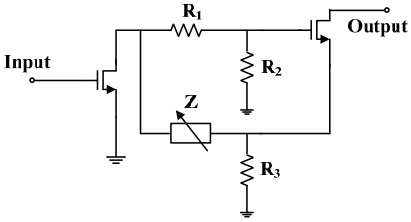
(b) Loaded line;



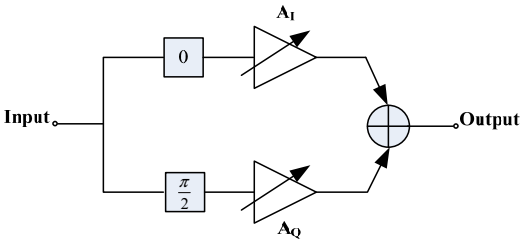
(c) Reflection;



(d) Switched filter;



(e) All pass;



(f) Vector sum.

Fig. 6.14 Typical phase shifters.

Since the asynchronous wake-up scheme is adopted in our wireless wire system, the peak power of the main Rx is less stringent because its active time is almost identical to the payload length and the average power can be very low. Consequently, active phase shifters become more suitable choices due to their low insertion loss, small chip area and wide phase shifting range. In the following sections, a partially-active 60-GHz narrowband reflection type phase shifter and a 60-GHz 4-bit wideband vector-sum type active phase shifter are designed and analyzed using the standard TSMC 65-nm CMOS

models. Their simulation results are shown in the following sections.

6.4.2 Design of a 60-GHz reflection-type phase shifter

Besides with distributed lines like the one illustrated in Fig. 6.14 (c), the reflection type phase shifter (RTPS) can be implemented by lumped elements, as shown in Fig. 6.15. The lumped version achieves more compact layout at the expense of bandwidth and robustness. However, if the load is carefully designed, the circuit can achieve relatively low loss with a wide phase shifting range.

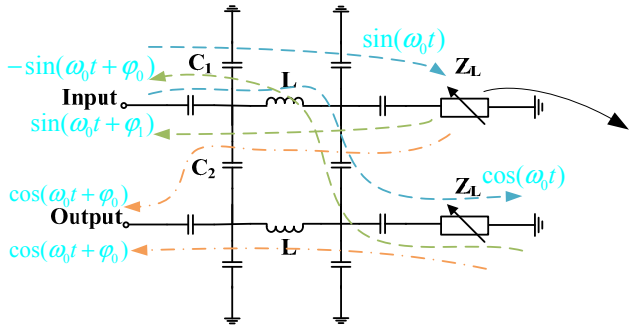


Fig. 6.15 The reflection type phase shifter using lumped elements.

It can be seen from Fig. 6.15 that the input signal is split into two branches with equal amplitudes (ideally) and quadrature phases: one has a 0° phase and another has a 90° phase. When those two signals are reflected back by the reflection loads, they are combined destructively at the input port and constructively at the output port. As a result, the RF signal experiences a phase shift

$$\Phi = -180^{\circ} - 2\cdot\tan^{-1}(\frac{Z_L}{Z_0}) \tag{6.22}$$

where

$$Z_0 \approx \sqrt{\frac{L}{(C_1+C_2)}} \tag{6.23}$$

in this lumped circuit. Since this circuit is fully symmetrical and the loads are equal, the input matching is not a problem and this structure can be relatively easily cascaded for larger phase shifting range. However, when the parasitics and loss of the passive devices are taken into account, its phase shifting range will be reduced and at the input port, the reflection signals are not able to completely cancel each other due to unequal amplitudes. As a result, the input isolation of this circuit structure should be carefully designed. In this work, we do not use the cascade structure in order to avoid high insertion loss. Instead, an active reflection load is chosen to provide wide phase tuning range and low insertion loss features simultaneously.

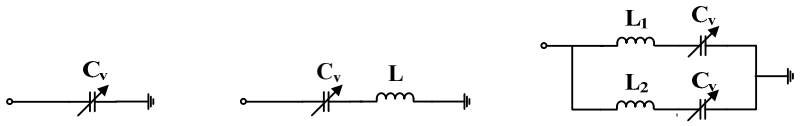


Fig. 6.16 Typical refection load configurations: (a) Capacitive load (CL); (b) resonating load (RL) and (c) dual resonating load (DRL).

The most well-known reflection load configurations are discussed in [19], and they are shown in Fig. 6.16. Due to the limited CMOS varator tuning range and the high on-chip parasitic capacitance level, the maximum phase shifting of the circuit in (a) is normally less than 40° [19], which is calculated by

$$\Delta\Phi_{CL} = 2\left|\tan^{-1}\left(\frac{Z_{L,\max}}{Z_0}\right) - \tan^{-1}\left(\frac{Z_{L,\min}}{Z_0}\right)\right| \tag{6.24}$$

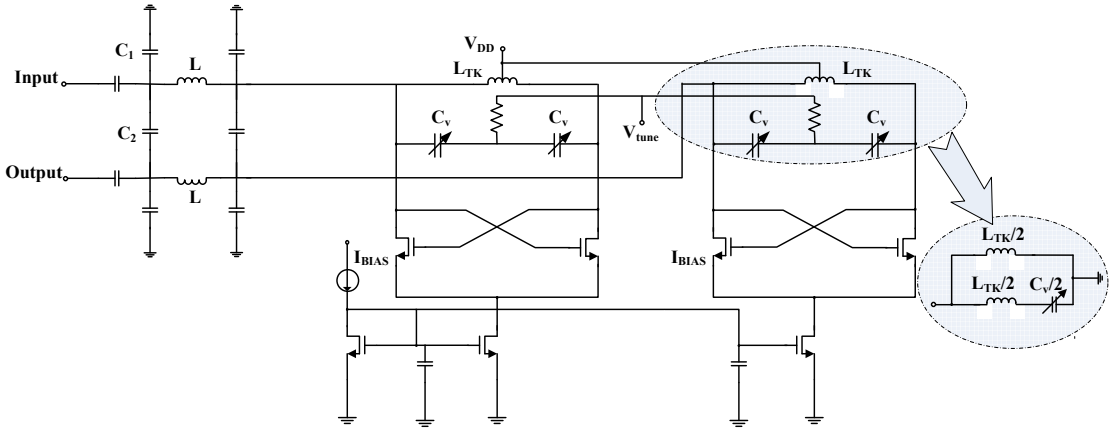
where $Z_{L,max}$ and $Z_{L,min}$ are the maximum and minimum load impedance, respectively. When adding a series resonating inductor L, as shown in (b), the phase shifting range becomes

$$\Delta\Phi_{RL} = 2\left|\tan^{-1}\left(\frac{Z_{C_{\max}\ or\ C_{\min}}}{Z_0}\right)\right| \tag{6.25}$$

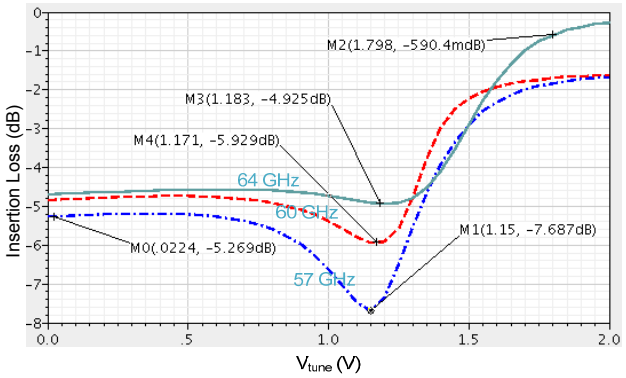
if L resonates with the average value of C_{max} and C_{min} at ω_0. The phase shifting range can be further extended by adding another series LC path, i.e. by adding an extra zero in its transfer function, as shown in Fig. 6.16 (c), which is normally referred as a dual resonated load (DRL). If not considering the parasitic capacitance and limited quality factor, a DRL is able to achieve 360° phase shift. However, its disadvantages are also obvious, i.e. doubled chip area, increased coupling and higher insertion loss.

Provided a differential signal is used (as discussed in Chapter 5), a 180° phase shift is required for the phase shifter. Besides, since the phase shifter is deployed in the RF signal path, its insertion loss must be minimized. As a result, a new reflection load with negative resistance loss compensation mechanism is proposed. Unlike the L-matching impedance transformation reflection loads proposed in [19], our reflection loads are composed of an inductor in parallel with a series LC path of Fig. 6.17 (a). However, when looking at the load itself, it looks like an LC oscillator. The only difference is that this active load does not oscillate itself. The main function of the cross-coupled transistor pairs is to provide negative resistance to compensate part of the loss of the reflection loads. It should be noted that the phase shifting range of this load can be extend by using one side of the varator, but this will require a large DC block capacitor and make the load structure asymmetrical. As a result, a two-varactor scheme is chosen as a trade-off. As shown in the lower right corner of Fig. 6.17, the effective varator capacitance is $C_v/2$ instead of C_v.

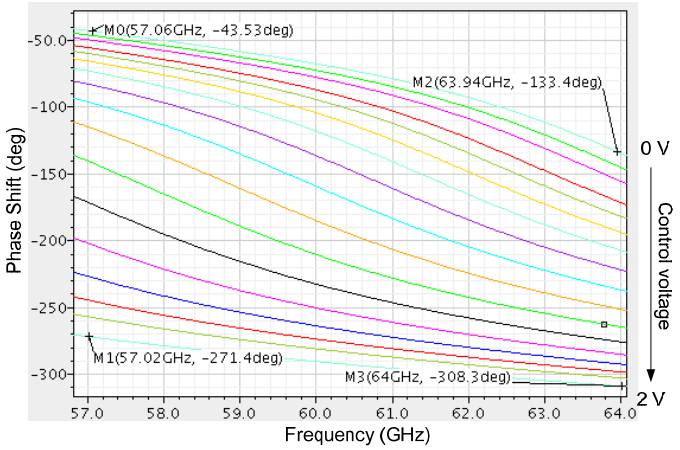
The direct result is that the maximum achievable phase shifting range is reduced form 360° to about 180°. However, this is acceptable in our differential system.



(a)



(b)



(c)

Fig. 6.17 (a) RTPS with negative resistance loss compensation loads, and RTPS performance of (b) Insertion loss at 57 GHz, 60 GHz and 64 GHz and (c) Phase shift for different control voltages (calibration is not done).

It can be seen that this RTPS is able to achieve about 180° phase shift range across the entire 60 GHz band, i.e. from 57 to 64 GHz. However, the insertion loss at 57 GHz is very large due to the low quality factor of MOS varactors. Besides of that, the narrow bandwidth of the quadrature-phase generation network also degrades the insertion loss. The simulated noise figure varies from 5.5 to 11 dB with different control voltages and frequencies. The peak power consumption (by the active loads) is about 6 mW. Compared to [20], the negative resistance of the load is able to compensate part of the insertion loss. Therefore, the insertion loss is at least 2-dB better than [20] from 60 to 64 GHz and comparable to it at lower frequencies. The input-referred -1 dB compression point ($P_{-1dB,in}$) of this RTPS is beyond -10 dBm with different control voltages, as illustrated in Fig. 6.18.

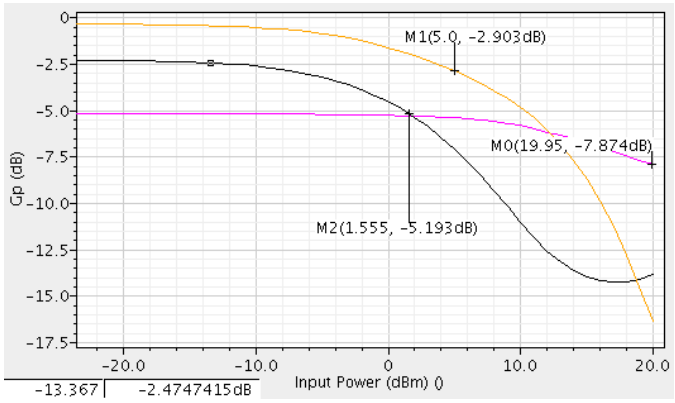


Fig. 6.18 $P_{-1dB,in}$ of the RTPS with different control voltages (at 60 GHz).

6.4.3 Design of a 60-GHz 4-bit vector-sum type phase shifter

In [17], [20] and [21], vector-sum type active phase shifters are presented. By modulating the amplitudes of two quadrature signals, their resulting vectors will have constant amplitude with different phases, as shown in Fig. 6.19 (a). This type of phase shifter can be used for analog continuous phase shifting and discrete digital phase shifting as well by controlling weights of biasing currents in digital bits, as shown in Fig. 6.19 (b).

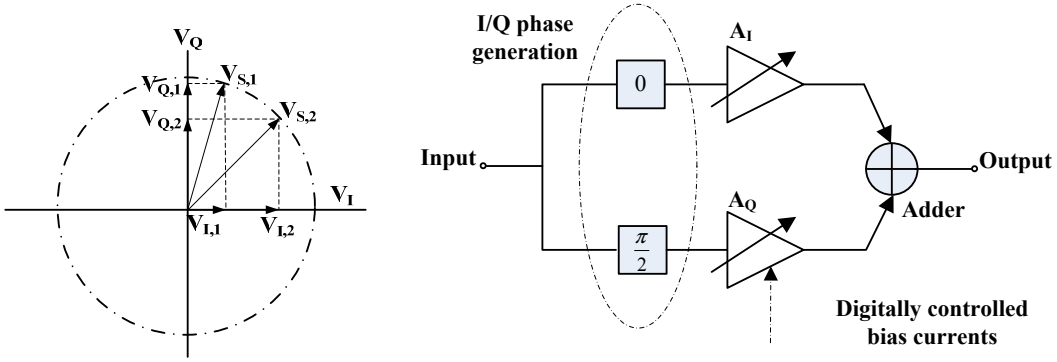
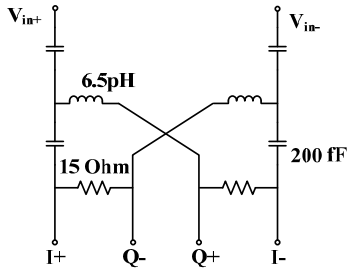
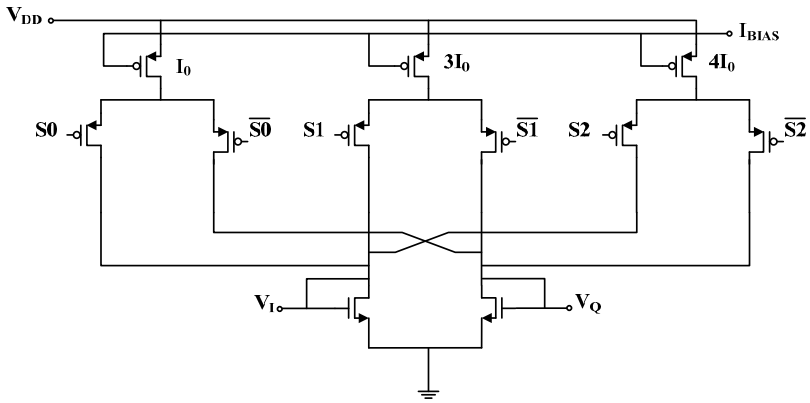


Fig. 6.19 (a) Concept and (b) architecture of a vector-sum type phase shifter.

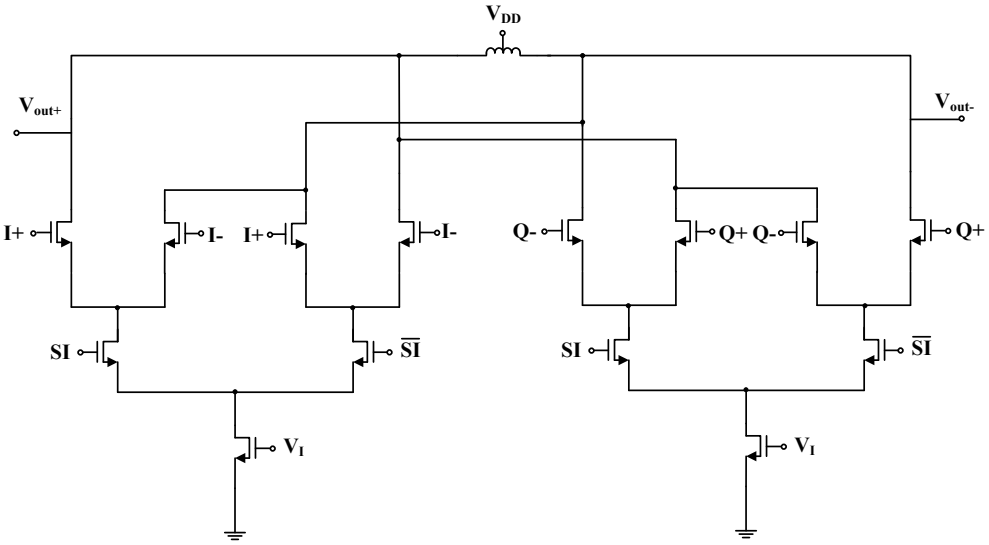
Circuit implementation details are shown in Fig. 6.20 (a), (b) and (c) respectively. An all-pass I/Q generation network is shown in (a). Compared to a traditional poly-phase RC network, this all-pass network has a comparable phase error as a 2-stage poly-phase filter (wide-band feature), but has the highest voltage gain (3-dB higher than a 1-stage poly phase filter) among other passive phase generation networks [17].



(a) An all-pass I/Q Generation circuits



(b) a 3-bit DAC bias control



(c) I/Q VGA

Fig. 6.20 Schematic of a 4-bit 60-GHz vector-sum type phase shifter.

A 3-bit programmable tail current control module is illustrated in (b) and the I/Q signal adder (i.e. VGAs) is shown in Fig. 6.20 (c). With differential input and output feature, a 360° phase shift can be achieved by switching the polarities of input I/Q signals and altering the weights of biasing currents, i.e.

$$\Phi = \tan^{-1}(A_j / A_i) \tag{6.26}$$

and

$$A_{out} = \sqrt{A_i^2 + A_j^2} \tag{6.27}$$

where A_i, A_j and A_{out} are the amplitudes of total in-phase, quadrature-phase and output signals, respectively. Simulation results are shown in Fig. 6.21 to Fig. 6.24.

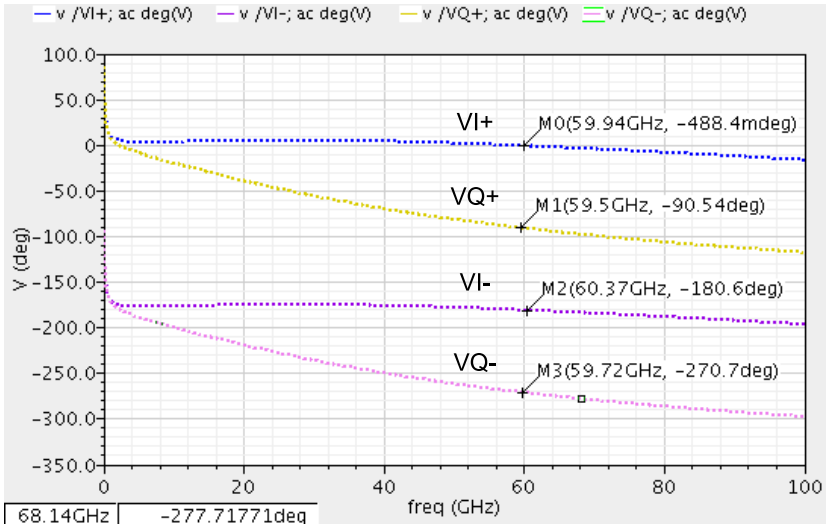


Fig. 6.21 I/Q phases.

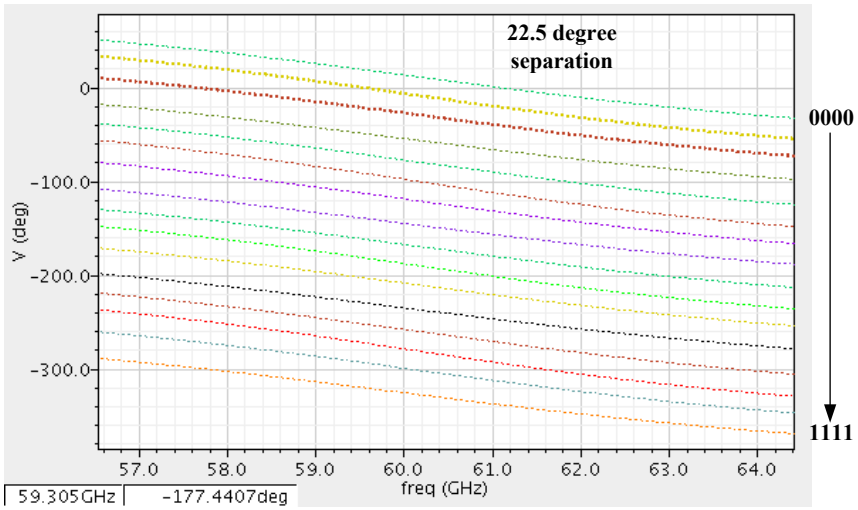


Fig. 6.22 Insertion phases (deg)

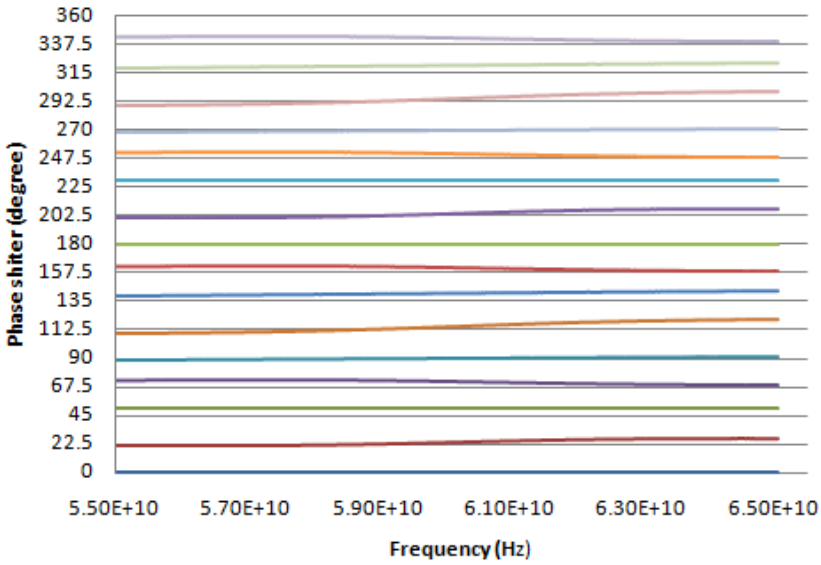


Fig. 6.23 Relative phase shifts of different settings.

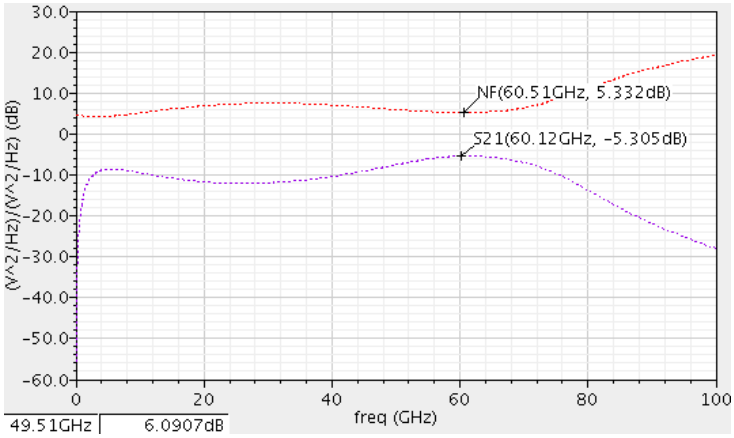


Fig. 6.24 S_{21} and *NF* of the stand-alone I/Q generation network.

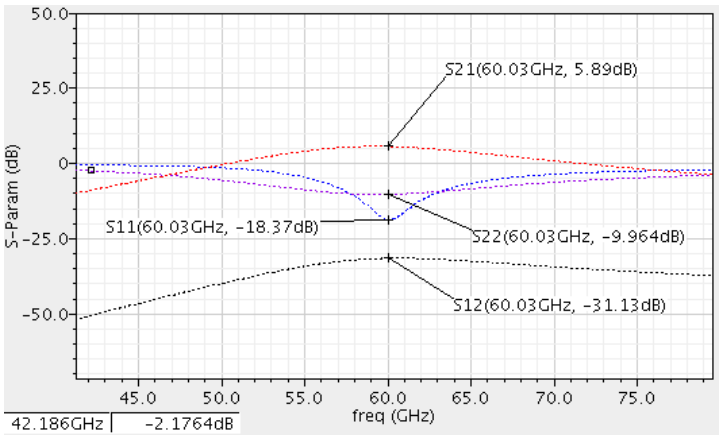


Fig. 6.25 S-parameters of the stand-alone adder.

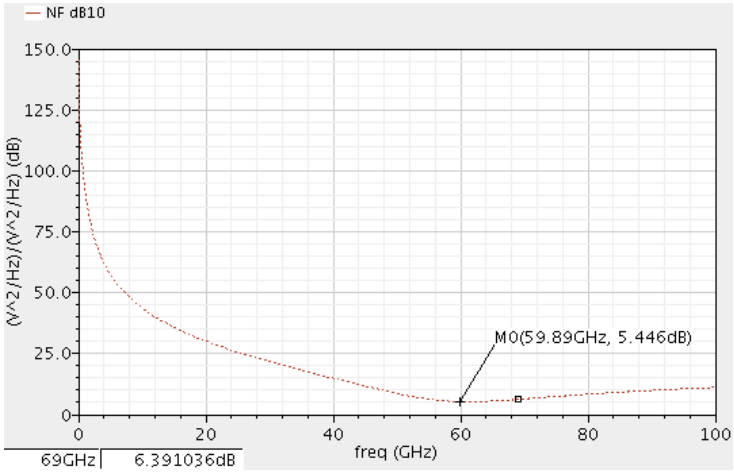


Fig. 6.26 *NF* of the stand-alone adder.

The vector-sum phase shifter achieves around 0 dB insertion loss with 12.5 mA current from a 1-V supply. The simulated NF is about 17 dB and P_{-1dB} is -7 dBm.

Considering phase and amplitude mismatches of the I/Q generation part and gain mismatch of adders, the phase deviations from the standard phase settings of our phase shifter are less than 6° in the worst case, which is about 30% of the 22.5° step. If we assume that the phase deviation in each path follows a Gaussian distribution with its expectation identical to the desired phase shift value and a variance of 6°, the maximum beam-pointing error of a 4-element phased array becomes 0.85°. As discussed in Section 6.1.3, the minimum beam-pointing resolution is about 7°, and the phase error in each path will introduce about 12% beam steering error. Compared to 38° beamwidth, this phase error will not cause obvious SNR degradation. Referring to discussions in Section 6.2, this 4-bit vector-sum type phase shifter fulfills the system accuracy specifications.

6.5 Design example of a 4-elements phased array Receiver

In this section, two versions of the 60-GHz 4-elements phased array Rx are shown as design examples. The first one is based on the reflection-type phase shifters, as illustrated in Fig. 6.27. Since the polarity switching stage is not yet added, this Rx is able to achieve a 180° analog phase shifting.

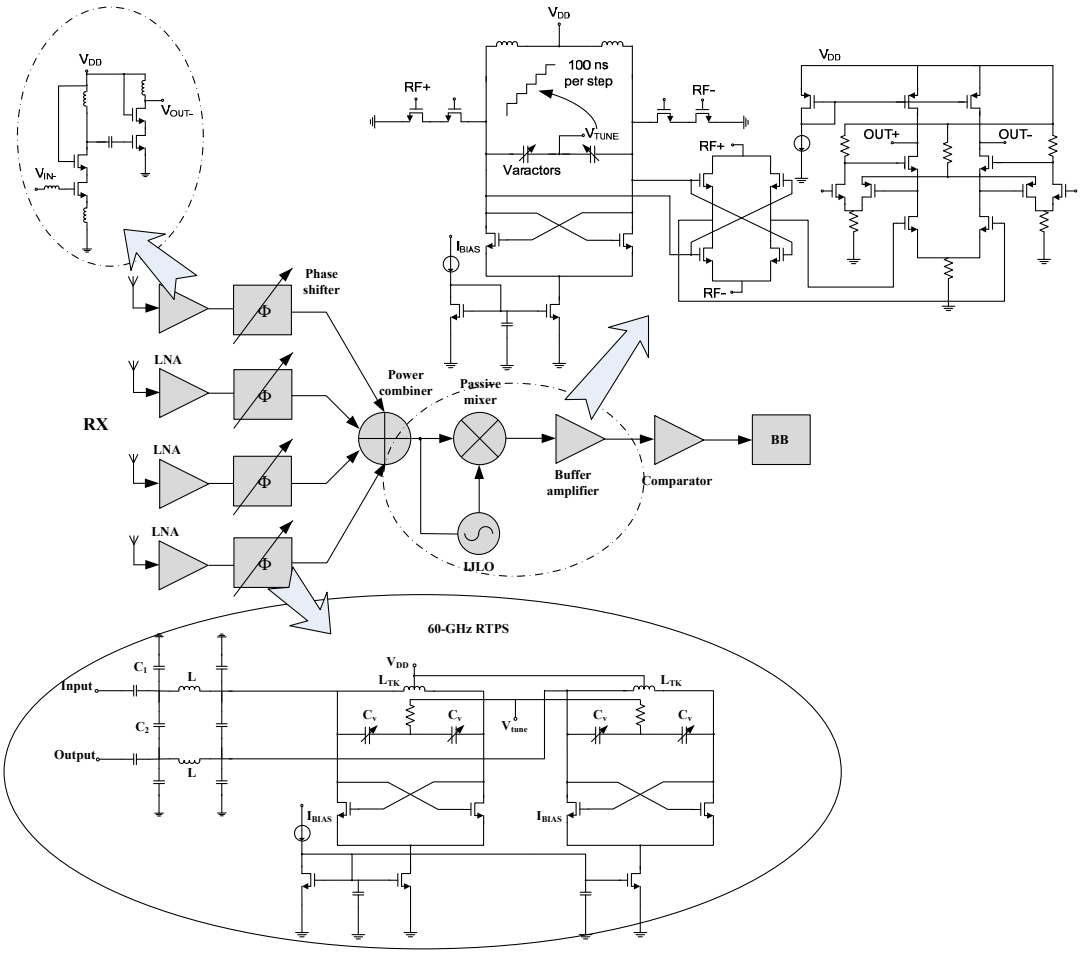


Fig. 6.27 Schematic of a 4-element phased array Rx based on the reflection-type phase shifters.

The effects of tuning the control voltage for different phase settings on the output signals (without any modulation) are illustrated in Fig. 6.28. Amplitude differences are coming from different insertion losses of the RTPS as well as different impedance level between the RTPS and the subsequent stages.

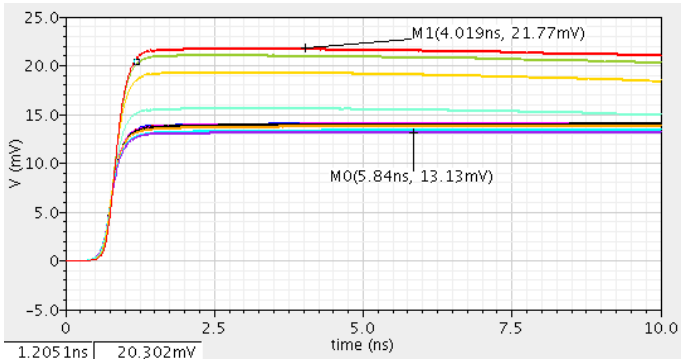


Fig. 6.28 Output signal amplitude variations under different phase settings (input power is -60 dBm and its output terminals are connected to 50-Ohm loads by adding source follower output buffers).

The OOK-modulated -60 dBm 60-GHz input RF signal is sent to the LNA with 1 Gbps peak data rate. When the tuning voltage of the RTPS is 1.5 (middle of tuning range), the Rx is able to achieve -60 dBm to 15 mV RF-DC conversion. The total power consumption is about 70 mW from a 1-V supply. The simulated NF of each Rx path varies from 6 dB to 6.5 dB under different phase settings. The simulated IIP3 is about -10 dBm.

The other version of the Rx based on the 4-bit vector-sum-type phase shifters is shown in Fig. 6.30.

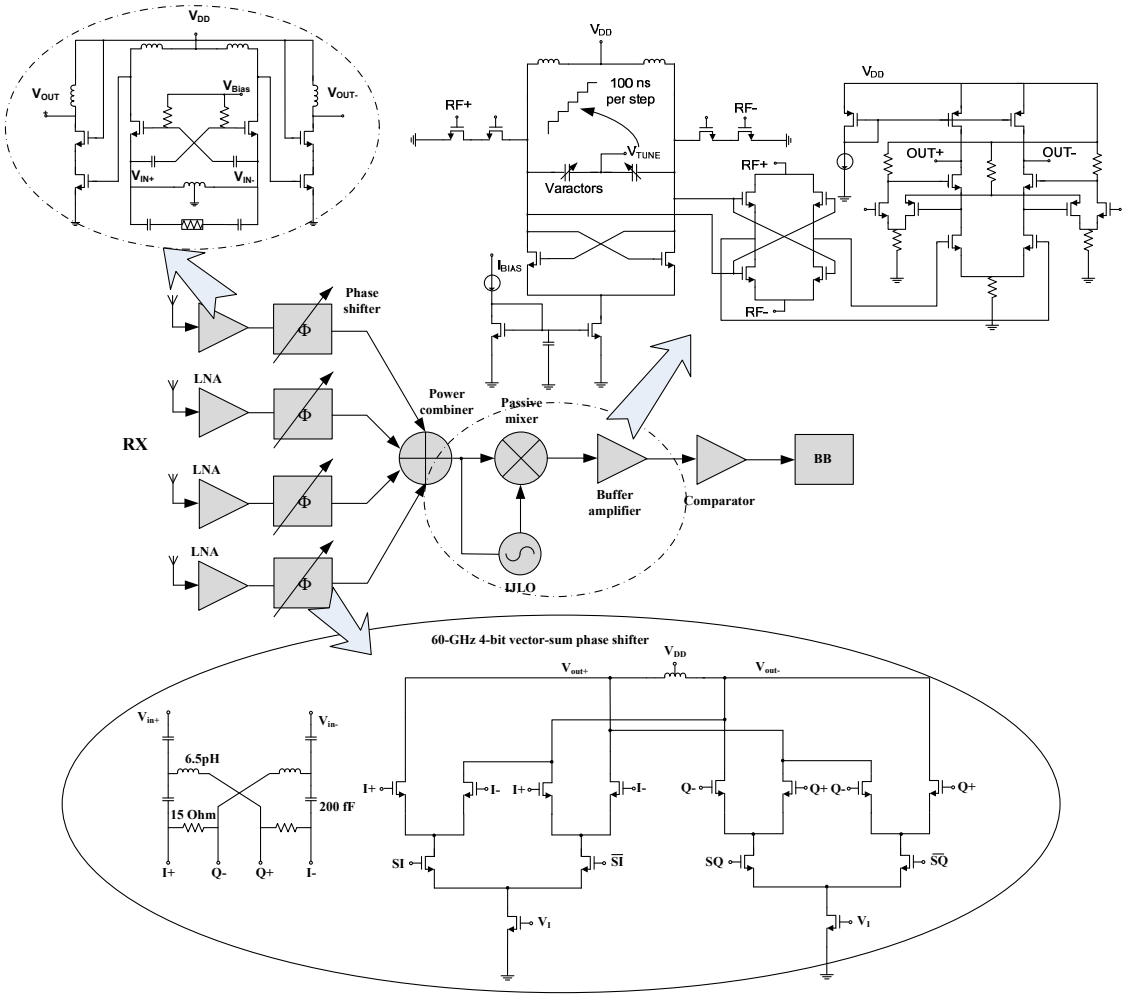


Fig. 6.29 Schematic of a 4-element phased array Rx based on vector-sum-type phase shifters.

This version of Rx is able to achieve -60 dBm to 50 mV RF-DC conversion capability with about 100 mW power consumption from a 1-V supply. The simulated NF of each Rx path is 10 dB and IIP3 is about -10 dBm.

6.6 Conclusions

Multi-element beamforming techniques have been discussed in this chapter. The relevant building blocks of the 60-GHz phased array Rx have been designed and simulated by using the TSMC 65-nm CMOS models. Two versions of the 4-element beamforming Rx have been designed and simulated based on different types of phase shifters and different LNA designs. It is shown that the RTPS-based Rx is able to achieve a -60 dBm to 15 mV power-to-voltage transform ratio at a 1-Gbps data rate. The total power consumption is about 70 mW from a 1-V supply, which means the raw E_{bit} of such a system is about 70 nJ/bit. The vector-sum phase shifter based Rx is able to achieve a -60 dBm to 50 mV power-to-voltage transform ratio at a 1-Gbps data rate at the expense of 100 mW power consumption (further optimization can be down to improve the performance). However, the positive aspects of the second version of Rx are: (i) the LNA of the latter Rx is differential and thus less sensitive to ground resistance or inductance while implemented on chip; (ii) the latter is digitally controlled and has higher accuracy without particular calibration process of control voltage, and (iii) it has better power-to-voltage conversion capability. In a next project, the power consumption of the second version of Rx can be further reduced (by scarifying the linearity). As we have discussed in Chapter 3, in such an mmW communication scenario, the power received by the Rx front end is normally weak, and a very high linearity is not required for the Rx.

References

[1] Yikun Yu, *Design Methods for 60GHz Beamformers in CMOS*, Ph.D. dissertations, Eindhoven University of Technology, Eindhoven, The Netherlands, 2010.

[2] Ali M. Niknejad *et al*., *mm-Wave Silicon Technology 60 GHz and Beyond*, Springer, New York, 2008.

[3] Discussion with Prof. Peter Baltus, 2011..

[4] http://www.scribd.com/doc/50166124/88/Pulse-shaping-to-reduce-ISI.

[5] Yikun Yu *et al*., "A 60GHz Phase Shifter Integrated with LNA and PA in 65nm CMOS for Phased Array Systems," in IEEE *Journal of Solid-State Circuits*, vol. 45, NO. 9, pp. 1697-1709, August 2010.

[6] Maryam Tabesh *et al*., "A 65nm CMOS 4-Element Sub-34mW/Element 60GHz Phased-Array Transceiver," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2011.

[7] Wei L. Chan *et al*., "A 60GHz-Band 2*2 Phased-Array Transmitter in 65nm CMOS," in *IEEE International Solid-State Circuits Conference*, Feb 2010, pp. 42-43.

[8] K. Raczkowski *et al*., "A Wideband Beamformer for a Phased-Array 60GHz Receiver in 40nm Digital CMOS," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2010.

[9] J. Nsenga *et al*., "Mixed Analog/Digital Beamforming for 60 GHz MIMO Frequency Selective Channels," in *proceedings of IEEE International Conference on Communications*, 2010.

[10] David J. Cassan *et al*., "A 1-V Transformer-Feedback Low-Noise Amplifier for 5-GHz Wireless LAN in 0.18-μm CMOS," in *IEEE Journal of Solid-State Circuits*, vol. 38, NO. 3, 427-435, March 2003.

[11] Terry Yao et al., "Algorithmic Design of CMOS LNAs and PAs for 60-GHz Radio," in *IEEE Journal of Solid-State Circuits*, vol. 42, NO. 5, 1044-1057, May 2007.

[12] Peter Baltus, *Minimum Power Design of RF Front Ends*, PhD dissertation, Eindhoven University of Technology, Netherlands, 2004.

[13] Chao-Shiun Wang *et al*., "A 60-GHz Phased Array Receiver Front-End in 0.13-μm CMOS Technology," in *IEEE Transaction on Circuits and Systems—I: Regular papers*, vol. 56, no. 10, 2009.

[14] Kenichi Okada et al., "A 60GHz 16QAM/8PSK/QPSK/BPSK Direct-Conversion Transceiver for IEEE 802.15.3c," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2011.

[15] Yo-Sheng Lin, *et al*., "Low-Power 48-GHz CMOS VCO and 60-GHz CMOS LNA for 60-GHz dual-conversion," in *International Symposium on VLSI Design, Automation and Test*, 2009.

[16] Bo-Jr Huang *et al*., "Design and Analysis for a 60-Ghz Low-Noise Amplifier with RF ESD Protection," in *IEEE Transactions on MTT*, vol. 57, issue 2, Feb. 2009.

[17] Kwang-Jin Koh, *Integrated Microwave and Millimeter Phased Array Designs in Silicon Technologies*, PhD dissertation, University of California, San Diego, US, 2008.

[18] David Viveiros *et al*., "A tunable All-Pass MMIC Active Phase Shifter," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no.8, August, 2002.

[19] Hossein Zarei *et al*., "Reflection-Type Phase Shifters for Multiple-Antenna Transceivers," in IEEE Transactions on Circuit and Systems—I: Regular Papers, vol. 54, no. 8, August 2007.

[20] Ming-Da Tsai and Arun Natarajan, "60GHz Passive and Active RF-path Phase Shifters in Silicon," in IEEE Radio Frequency Integrated Circuits Symposium (RFIC), 2009.

[21] Holger Erkens *et al*., "A Comparison of two RF Vector-Sum Phase Shifters Concepts," in *German Microwave Conference*, March 2009.

CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS

The next generation of personal communication systems requires high speed data transmission to support versatile applications like P2P communication, high speed downloading and synchronization. Furthermore, it requires low power consumption for the communication systems in order to achieve high mobility and autonomous operation as well. In this thesis, potential solutions to achieve such low power and high speed communications have been investigated and discussed, and the relevant systems and circuits have been designed and demonstrated. Conclusions of the thesis are as following:

- We have shown the feasibility of achieving high energy efficiency of communication at high frequency bands. By introducing extra antenna gain of directional antennas, it is possible to (over) compensate the "frequency-caused losses" with the "frequency-induced gain" under certain IC technology level, and therefore the transceiver front ends are able to achieve a better E_{bit}. Moreover, based on analytical models, optimized PHY parameters have been identified to support the targeted applications in this research.
- The lowest possible average power dissipation of the entire Rx front end for low-event-frequency personal communication applications can be achieved using the asynchronous duty-cycled wake-up power management method proposed in this work. This has been verified by comparing with other widely used low-power MAC power management methods.
- Important features like low average power consumption, high communication speed and fast settling behavior of a front end can be achieved simultaneously by adopting the asynchronous IJLO-based duty-cycled WuRx proposed in this work. It has been shown that this approach can be integrated into a low-power and high-speed wireless wire system and it helps to minimize the average power consumption of such a system efficiently.
- The quadratic relation between the input and output signals of an energy detector can be eliminated by inserting an IJLO into one of the branches. As a result, the relation becomes linear and the front end sensitivity can be significantly increased.
- The equivalent noise bandwidth of the passive mixer in an energy-detection WuRx can be reduced, at the expense of longer detection time, by adding a low-pass filter and controlling its bandwidth. This will also improve the sensitivity of the WuRx significantly.
- The trade-off between sensitivity and locking range of the IJLO can be eliminated by using the frequency-sweeping method proposed in this work.
- Combining above approaches, a 70GHz WuRx can achieve 8.4GHz system bandwidth, -60dBm sensitivity and 10 μs worst-case settling time at a power dissipation of 10 mW, without introducing missed alarms (other than through strong

interferers). This was verified through the design, implementation and characterization of a 70-GHz IJLO-based high-sensitivity, low-power and fast-responding WuRx in TSMC 65-nm CMOS technology.

- The main Rx of the wireless wire system can achieve an energy efficiency of less than 100 pJ/bit at 1Gbps data rate and -60dBm sensitivity using the technologies and methods described in this work. This was verified through the designs of two versions of 4-element phased array Rxs with IJLO-based self-mixing demodulation module using TSMC 65-nm CMOS models. The first version of the Rx contains a 60-GHz LNA that achieves about 14 dB gain, 2 GHz bandwidth, -4 dBm *IIP3* and 5 dB *NF* with 10 mW power consumption, and a passive negative-resistance-loaded reflection type phased shifter that achieves -7 to -1dB insertion loss, 5.5 to 11 dB NF with less than 6 mW power consumption. An alternative 4-bit 360° vector-sum type active phase shifter is used in the second version of the Rx as well. This version of the phase shifter is able to achieve 0-dB insertion loss, 17-dB *NF* and -7 dB *IIP3* with 12.5 mW power consumption.
- By using above methods, the entire wireless wire Rx system is able to achieve about 20 μW average power consumption with 10 Gb data load per day and 1000 event/day. This is verified by analytic models.

For future research, following aspects can be addressed or further optimized:

- More comprehensive simulations, verifications and optimizations should be carried on for the phased array Rx. As indicated by the power-linearity factor, lower power consumption can be achieved at the expense of worse front end linearity feature, which is not critical in this particular mmW communication.
- ID-based WuRx should be investigated and implemented to avoid false alarms.
- By adding an LNA in front of the WuRx, better sensitivity can be achieved.
- An advanced MAC protocol may be added to avoid the worst-case beamforming scenario in order to further save power.
- The Realization of the prototype that combines the main Rx and WuRx would be interesting and useful to verify the effectiveness of our wireless wire system.
- A low power Tx front end is required. Among all blocks of the Tx, a low power and high accuracy mmW PLL, which will influence the effectiveness of injection locking based wake up and self-mixing demodulation concepts is critical, and therefore is worth being investigated.
- A configurable multi-element phased array system is recommended. More flexibility of adjusting antenna beamwidth, antenna gain (communication range) and power consumption level can be obtained by adjusting the number of antenna elements (front end paths).
- Integration with an antenna array is required.
- An RF-signal based wireless charging module can be further added. Using

multi-stage charge pumping method mentioned in Chapter 2, nodes can be charged by an asymmetrical power source (e.g. a multi-element access point) during their idle times.

APPENDIX A NOISE MODEL

The direct relationship between the noise and power consumption is hard to be modeled simply and accurately. Besides, the circuit noise has to be optimized in a proper design and the minimum noise is fundamentally decided by the frequency and technology, which has not so much freedom to be changed. As a result, the influence will be shift to the power gain. The noise factor of a matched cascaded stages Rx (LNA plus mixer) can be calculated as

$$F_{tot} = F_1 + \frac{F_2 - 1}{G_1} \tag{A-1}$$

or

$$F_{tot} - F_1 = \frac{F_2 - 1}{G_1} \tag{A-2}$$

where F_1, F_2 and F_{tot} are the noise factors of the first, second and total Rx respectively and G_1 is the gain of the first stage. If the n-times worse noise factor is caused by less gain of the first stage, it will obtained that

$$n \cdot F_{tot} - F_1 = \frac{F_2 - 1}{G_1'} \tag{A-3}$$

where G_1' is the required gain of the first stage to achieve the new noise factor.

If dividing the left and right parts of (A-2) by the left and right parts of (A-3) respectively, it can be obtained that

$$\frac{F_{tot} - F_1}{n \cdot F_{tot} - F_1} = \frac{G_1'}{G_1} \tag{A-4}$$

Since the gain of the first stage, i.e. normally the LNA is proportional to its power consumption (the transconductance is linear to the biasing current), (A-4) can be re-written as

$$\frac{F_{tot} - F_1}{n \cdot F_{tot} - F_1} = \frac{P_{DC,1}'}{P_{DC,1}} \tag{A-5}$$

where $P_{DC,1}$ and $P_{DC,1}'$ are the power consumption of the first stage before and after increasing the noise respectively. Assuming the ratio of the total noise factor and the first-stage noise factor is ε, it can be seen that

$$P_{DC,1}' = \frac{\varepsilon - 1}{n \cdot \varepsilon - 1} \cdot P_{DC,1} \tag{A-6}$$

and the total power consumption of an n-path Rx becomes

$$P_{Rx,n\text{-}path,1} = n \cdot \left(\frac{\varepsilon - 1}{n \cdot \varepsilon - 1} \cdot P_{DC,1} + P_{DC,2} \right) \tag{A-7}$$

where $P_{DC,2}$ is the power consumption of the second stage of the Rx.

The interesting conclusion can be drawn from (A-7) is that when the path number n is very large, the noise tolerance will be infinitely improved and thus

$$\frac{\varepsilon - 1}{n \cdot \varepsilon - 1} \to 0 \tag{A-8}$$

i.e. the LNA is not compulsory anymore and the total power model becomes

$$P_{TRx,n\text{-}path,1} = \frac{1}{n} \cdot P_{LO,DC,1} + \frac{1}{n} \cdot P_{PA,DC,1} + n \cdot P_{DC,2} \tag{A-9}$$

and n_{opt1} becomes

$$n_{opt,1} = \sqrt{\frac{P_{LO,DC,1} + P_{PA,DC,1}}{P_{DC,2}}} \tag{A-10}$$

Re-using the data in Table 4.2 and assuming the first and second stages of the Rx consume equal power, the power budget becomes

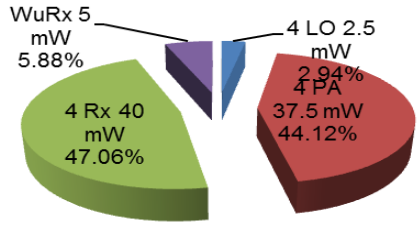


Fig. A-1 Power budget with the worse-noise 4-element Rx.

It can be found that by evenly distributing the power budget to the Rx and Tx, the entire system is able to achieve 85 mW total power consumption power with a 4-element transceiver front end.

APPENDIX B. RF HARVESTING

The surface area of a sphere with radius R can be calculated as

$$S = 4\pi R^2 \tag{B-1}$$

The beam-pointing angle of an n-antenna array is

$$\theta_b = \frac{2}{n-1} \tag{B-2}$$

in which the distance between the antenna element is set as λ/2, i.e. half wavelength.

The EIRP of the n-element Tx is

$$P_{out,tot} = P_{out,1-path} \cdot n^2 \tag{B-3}$$

where $P_{out,1\text{-}path}$ is the output power of one branch of the Tx array.

Assuming an omni-directional antenna is used for the charging Rx, the received power density can be calculated as

$$P_{d,Rx} = \frac{P_{out,tot} / PL}{\frac{\theta_b}{2\pi} \cdot S} = \frac{P_{out,1-path} \cdot n^2 \cdot (\frac{\lambda}{4\pi R})^2}{\frac{2}{\frac{n-1}{2\pi}} \cdot 4\pi R^2} = \frac{P_{out,1-path} \cdot n^2 \cdot (\frac{\lambda}{4\pi R})^2}{\frac{2}{\frac{n-1}{2\pi}} \cdot 4\pi R^2} \tag{B-4}$$

If R is 1 meter, for a 60-GHz wireless charging beam, $P_{out,1\text{-}path}$ is 10 mW and the required power density for charging is 20 μW/cm^2 (0.2 W/m^2), the resulting n of the Tx becomes about 800.

If R is 1 meter, for a 2.4-GHz wireless charging beam, $P_{out,1\text{-}path}$ is 10 mW and the required power density for charging is 20 μW/cm^2 (0.2 W/m^2), the required n of the Tx is about 100.

SUMMARY

With the rapid development of communication technologies, wireless personal-area communication systems gain momentum and become increasingly important. When the market gets gradually saturated and the technology becomes much more mature, new demands on higher throughput push the wireless communication further into the high-frequency and high-data-rate direction. For example, in the IEEE 802.15.3c standard, a 60 GHz physical layer is specified, which occupies the unlicensed 57 to 64 GHz band and supports gigabit links for applications such as wireless downloading and data streaming. Along with the progress, however, both wireless protocols and physical systems and devices start to become very complex. Due to the limited cut-off frequency of the technology and high parasitic and noise levels at high frequency bands, the power consumption of these systems, especially of the RF front-ends, increases significantly. The reason behind this is that RF performance does not scale with technology at the same rate as digital baseband circuits.

Based on the challenges encountered, the wireless-wire system is proposed for the millimeter wave high-data-rate communication. In this system, beamsteering directional communication front-ends are used, which confine the RF power within a narrow beam and increase the level of the equivalent isotropic radiation power by a factor equal to the number of antenna elements. Since extra gain is obtained from the antenna beamsteering, less front-end gain is required, which will reduce the power consumption accordingly. Besides, the narrow beam also reduces the interference level to other nodes. In order to minimize the system average power consumption, an ultra-low power asynchronous duty-cycled wake-up receiver is added to listen to the channel and control the communication modes. The main receiver is switched on by the wake-up receiver only when the communication is identified while in other cases it will always be in sleep mode with virtually no power consumed. Before transmitting the payload, the event-triggered transmitter will send a wake-up beacon to the wake-up receiver. As long as the wake-up beacon is longer than one cycle of the wake-up receiver, it can be captured and identified. Furthermore, by adopting a frequency-sweeping injection locking oscillator, the wake-up receiver is able to achieve good sensitivity, low latency and wide bandwidth simultaneously. In this way, high-data-rate communication can be achieved with ultra-low average power consumption. System power optimization is achieved by optimizing the antenna number, data rate, modulation scheme, transceiver architecture, and transceiver circuitries with regards to particular application scenarios. Cross-layer power optimization is performed as well. In order to verify the most critical elements of this new approach, a W-band injection-locked oscillator and the wake-up receiver have been designed and implemented in standard TSMC 65-nm CMOS technology. It can be

seen from the measurement results that the wake-up receiver is able to achieve about -60 dBm sensitivity, 10 mW peak power consumption and 8.5 μs worst-case latency simultaneously. When applying a duty-cycling scheme, the average power of the wake-up receiver becomes lower than 10 μW if the event frequency is 1000 times/day, which matches battery-based or energy harvesting-based wireless applications. A 4-path phased-array main receiver is simulated working with 1 Gbps data rate and on-off-keying modulation. The average power consumption is 10 μW with 10 Gb communication data per day.

ACKNOWLEDGEMENT

First of all, I would like to say sorry to my parents. I apologize for my naughty childhood and the troubles and headaches for you. I apologize that I am still a maverick and perhaps will never be a perfect daughter as you have expected. I apologize that as the only child of the family, I choose to stay far away from you for so many years. I apologize that I was not there when you felt lonely, dull or ill. I apologize that I cannot celebrate your birthdays for more than ten years. I apologize that I have never said I love you but I definitely do…I owe you so much that I have no way to compensate by any means. Thank you so much for your love and support. I love you more than you know!

My special thanks to my boyfriend Chao. With your unconditional love, accompany, support and help, I do not feel lonely and always have sufficient courage to face all challenges and problems. Life is a long journey and it will never be easy. There might be some other unexpected problems in the future. I do not fear, because I know you are always here with me.

I would like to express my sincere gratitude to my promotor prof. Peter Baltus. Thank you for giving me the chance to pursue my PhD research. Thank you for your trust and encourage all the time. In the most difficult moment, sometimes when I was almost desperate, you were always there, helping me out with your comprehensive knowledge and experience and cheering me up with your enthusiastic attitude. I believe the most precious things I have learnt from you are not only the knowledge itself, but also are your attitude of life, your passion to work and your selflessness to people. These priceless gifts from you will stay with me and influence me all the time.

I would like to thank my second promotor prof. Arthur van Roermund. Thank you for your trust to give me the PhD position as well. Thank you for the effort you have put in reviewing my thesis. I was a bit shocked when I saw your detailed comments and feedbacks, from the technical contents to the misuse of a single punctuation. I really appreciate your efforts very much!

I wish to thank my supervisor dr. Dusan Milosevic. Thank you very much for your efforts to the weekly meetings and to my technical papers and PhD thesis. I appreciate your help through all these years. Many thanks to my supervisor dr. Paul van zeijl. I really appreciate the interesting and inspiring discussions with you during the progress meetings. Thank you very much for providing me the opportunity to access the facilities for tape-outs in Philips research, which definitely helps a lot to my research work.

My appreciation to other members of my promotion committee, prof. Paul van den Bosch,

prof. John Long, prof. Bram Nauta, prof. Jean-Paul Linnartz and dr. Guido Dolmans, for your time and effort. Thanks the dean Prof. Ton Backx for your permission of my PhD promotion. Special thank to prof. John Long, for your help and guidance on my master education in TU Delft.

I wish to express my sincere appreciation to our beautiful secretary, Margot van den Heuvel. Thank you very much for your time spent on dealing with my issues these years. I really enjoy the 'lady talking' with you as well. I warmly thank Piet Klessens. Thanks a lot for your help about those Cadence problems and about the measurement stuff. I appreciate that very much!

I would like to thank all members in MsM group for the cozy and pleasant working environment, for the helps and discussions, and for the friendship. My special thanks to Pooyan Sakian Dezfuli. Thank you so much for your help on the measurements. I really appreciate your time and effort very much! It is hard to imagine how these difficult millimeter wave measurements can be done smoothly without your experience and useful guidance. Thank my colleague Wei Deng as well for your help on the measurement. I appreciate the interesting and useful discussions with you. My warmly thanks also go to my office mate Johan van den Heuvel. I really enjoy the cozy atmosphere of our office. The old office dinners are also wonderful! I wish to thank 'the Dutch gentlemen' Pieter Harpe very much! I believe you have a nearly perfect personality which is valuable and rare. I enjoy the talks and discussions with you very much! Thank all the members in MsM group and CWTe, Reza Mahmoudi, Jan Haagh, Eugenio Cantatore, Georgi Radulov, Hans Hegt, Hammad M. Cheema, Xiaopeng Yu, Yu Lin, Yikun Yu, Hao Gao, Erwin Janssen, Maarten Lont, Mehdi Sarkeshi, Jaap Essing, Chuang Lu and Foad Arfaei Malekzadeh and so on.

I wish to thank the members of Philips research. Thank Bob Theunissen for your support of my tape-outs and especially for your help on solving the troublesome Cadence problems. Thank Dave van Goor for your help on the measurement issues. Thank Henry van der Zanden for your helps and discussions, too. Thank the ex- group leader Neil Bird and my supervisor of the master project Mihai Sanduleanu. I have learned a lot from you! My special appreciation goes to John Mills. Thank you for your friendship for all these years. Thank you for your encouragement when you found I was struggling with tape-out in Philips late in the night. Thank you for your concern!

I wish to thank Mr. Mark Tomesen and all members of ST-Ericsson for offering me a nice job. This may be the greatest gift for graduation I've received.

Last but not least, I wish to thank the mystery and complication of the existence, which always stimulate me to learn more.

BIOGRAPHY

Xia Li was born on July 18th, 1983 in Xi'an China. She received the B.S. degree in electronics and information engineering from Harbin Institute of Technology in China and M.S. degree in microelectronics in Delft University of Technology, the Netherlands. She is currently working towards her Ph.D degree in Eindhoven University of Technology, the Netherlands.

Her research interests include wireless communication system research, cross-layer design and optimization, millimeter-wave RF IC design, low-power wireless system and circuit design, device modeling and design.

LIST OF PUBLICATIONS

Li, X., Baltus, P.G.M., Deng, W., Milosevic, D., Zeijl, P.T.M. van, Roermund, A.H.M. van. "60 GHz ultra-low power radio," in *19th Annual Workshop on Circuits, Systems and Signal Processing*, pp. 206-209 , Netherlands, November, 2008.

Wu, W., Sanduleanu, M., Li, X., Long, J.R., "17 GHz Front-Ends for Low-Power Wireless Sensor Networks," in *IEEE Journal of Solid-State Circuits*, vol. 43, NO. 9, pp. 1909-1919, September 2008.

Li, X., Baltus, P.G.M., Milosevic, D., Deng, W., Zeijl, P.T.M. van, Bird, N., Roermund, A.H.M. van, "Wireless wire - the 60 GHz ultra-low power radio system," in *Proceedings IEEE Radio and Wireless Symposium 2009 (RWS 2009)*, pp. 215-218, San Diego, California, 18-22 January 2009.

Li, X., Baltus, P.G.M., Milosevic, D., Zeijl, P.T.M. van, Roermund, A.H.M. van, "An EFOM for cross-layer optimization of low-power wireless networks," in *Proceedings of 20th Annual Workshop on Circuits Systems and Signal Processing*, November 26-27, 2009.

Li, X., Baltus, P.G.M., Zeijl, P.T.M. van, Milosevic, D., Roermund, A.H.M. van, "A lower-power, high-sensivity injection-locked oscillator for 60 GHz WPAN applications," in *Proceedings of CAS PrimeAsia 2009*, pp. 161-164, Shanghai, China, November 19-21, 2009 (best paper award).

Li, X., Baltus, P.G.M., Zeijl, P.T.M. van, Milosevic, D., Roermund, A.H.M. van, "An EFOM for Cross-Layer Optimization towards Low-Power and High-performance Wireless Networks," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*, Paris, May 2010.

Li, X., Baltus, P.G.M., Zeijl, P.T.M. van, Milosevic, D., Roermund, A.H.M. van, "A 70 GHz 10.2 mW Self-demodulator in 65-nm CMOS Technology," in *Proceedings of IEEE Custom Integrated Circuit Conference (CICC 2010)*, San Jose, US, Sep. 2010.

Li, X., Baltus, P.G.M., Milosevic, D., Zeijl, P.T.M. van, Roermund, A.H.M. van. "A Low-Power 60 GHz LNA," in *21st Annual Workshop on Circuits, Systems and Signal Processing*, November 2010.

Li, X., Baltus, P.G.M., Zeijl, P.T.M. van, Milosevic, D., Roermund, A.H.M. van, "A 60 GHz Ultra Low-Power Wake-up Radio," in *IEEE Radio and Wireless Symposium (RWS 2011) ,* 2011.

Li, X., Baltus, P.G.M., Zeijl, P.T.M. van, Milosevic, D., Roermund, A.H.M. van, "A 73 to 83 GHz 9-mW Injection-Locked Oscillator in 65-nm CMOS Technology," in *IEEE 11th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF 2011) ,* 2011 (the honorable-mentioned paper award).