

Chapter 2

System-Level and Architectural Trade-offs

This chapter focuses on high level design of ultra-low power wireless nodes. First, different system architectures are compared in order to assess advantages and drawbacks of each architecture from a power consumption point of view. Second, different modulation formats are compared and an optimal data-rate is chosen in order to minimize the average power consumption of the node. Finally, the most common transmitter and receiver architectures are reviewed.

2.1 Modulation Schemes for Ultra-low Power Wireless Nodes

Different radio architectures have been recently studied in order to reduce the power consumption. Some of these architectures comprise Ultra Wide-Band (UWB) transceivers, Back-scattering transceivers, Sub-sampling and Super-Regenerative transceivers, as well as spread-spectrum based transceivers (both frequency hopping and direct sequence).

Though spread spectrum techniques are also ultra-wideband modulation schemes, in this book, “ultra-wideband modulation” is used to refer to a spectrum that is larger than 500 MHz (e.g. impulse radio based schemes). Although a spread-spectrum modulated signal can have a bandwidth larger than 500 MHz, this is not a necessary condition. Therefore, “with spread-spectrum modulated signal”, in this book, we refer to any signal in which the transmitted bandwidth is much larger than the signal bandwidth (e.g. the transmitted bandwidth is larger than 10 times the signal bandwidth).

Looking finally to regulations, Federal Communication Commission (FCC) rules specify UWB technology as any wireless transmission scheme that occupies more than 500 MHz of absolute bandwidth or more than 20% of the carrier frequency.

2.1.1 Impulse Radio Transceivers

Among different architectures suitable for an ultra-low power implementation, UWB based systems are gaining more and more attention.

The most important characteristic of UWB systems is the capability to operate in the power-limited regime. In this regime, the channel capacity increases almost linearly with power, whereas at high Signal to Noise Ratio (SNR) it increases only as the logarithm of the signal power as shown by the Shannon theorem

$$C = BW \times \log_2 \left(1 + \frac{P_S}{P_N} \right) \quad (2.1)$$

where P_S is the average signal power at the receiver, P_N is the average noise power at the receiver and BW is the channel bandwidth. For low data-rate applications (small C), it can be seen from (2.1) that the required SNR can be very small given an available bandwidth in excess of several hundreds MHz. A small SNR translates in a small transmitted power and as a result in a reduction of the overall transmitter power consumption.

Although UWB transceivers can have reduced hardware complexity, they pose several challenges in terms of power consumption. In Fig. 2.1 a schematic block diagram of an UWB transceiver is shown. The biggest challenge in terms of power consumption is the Analog to Digital Converter (ADC). If all the available bandwidth is used, the sampling rate has to be in the order of several Gsamples per second. Furthermore, the ADC should have a very wide dynamic range to resolve the wanted signal from the strong interferers. This implies the use of low-resolution full-flash converters. It can be proven [17] that a 4-bit, 15 GHz flash ADC can easily consume hundreds of milliwatts of power. Even if a 1-bit ADC at 2 Gsample/s is used, the predicted power consumption of the ADC remains around 5 mW [18]. Furthermore, the requirement on the clock generation circuitry can be very demanding in terms of jitter.

Besides these drawbacks, wideband Low Noise Amplifier (LNA) and antenna design are challenging when the used bandwidth is in excess of some Gigahertz. The antenna gain, for example, should be proportional to the frequency [19], but most conventional antennas do not satisfy this requirement. LNA design appears quite challenging when looking at power consumption of state-of-the-art wideband LNAs [20]. A wideband LNA consumes between 9 and 30 mW making it very difficult to fulfill a constraint of maximum 10 mW peak power consumption for the overall transceiver. Although several successful designs are recently published

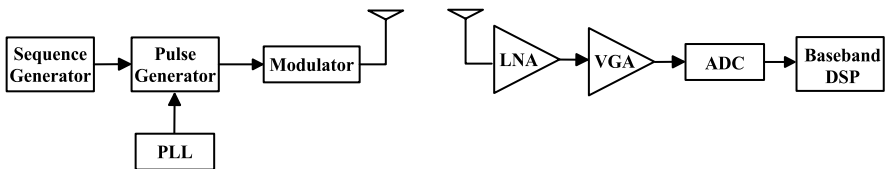


Fig. 2.1 The building blocks of an impulse based UWB transceiver

showing the potential of UWB systems, their power consumption remains too high to be implemented in a “micro-Watt node”. In [20] the total power consumption is around 136 mW at 100% duty cycle. In [21] a power consumption of 2 mW has been reported for the pulse generator only.

2.1.2 Back-scattering for RFID Applications

In the wide arena of low-power architectures, RFIDs represent a good solution when the applications scenario requires an asymmetric network. In this case the “micro-Watt node” needs to transmit data and to receive only a wake-up signal. The required energy is harvested from the RF signal coming from the interrogator. In [22] the interrogator operates at the maximum output power of 4 W, while generating by inductive coupling 2.7 μ W. This power allows a backscattering-based transponder to send On-Off Keying (OOK)-modulated data back to the interrogator in a 12 meters range using the 2.4 GHz ISM band. Unfortunately the limited amount of intelligence at the transmitter side makes this architecture not flexible and only suitable in a highly asymmetric wireless scenario.

There are however other drawbacks in this kind of modulation, mainly shadowed regions. There are two types of such regions. One occurs when the phase of the reflected signal is in opposition with the phase of the RF oscillator. The second occurs due to multiple reflections in an indoor environment. In this situation multiple paths can add destructively at the receiver. Therefore, an increase in the complexity of the receiver is required which can be very severe if the link should be robust enough. Finally the backscattering technique has an increased sensitivity to the fading. The small scale fading observed on the backscattered signal has deeper fades than a conventional modulated signal.

2.1.3 Sub-sampling

The Nyquist theorem has been explored in sub-sampling based receivers in order to reduce the overall power consumption. The power consumption of analog blocks mainly depends on the operating frequency. Applying the theory of bandpass sampling [23], it can be proven that the analog front-end can be considerably simplified reducing the operating frequency. This has the potential to lead to a very low power receiver implementation. Unfortunately due to the noise aliasing, it can be proven that the noise degradation in decibels is:

$$D = 10\text{Log}_{10}\left(1 + \frac{2MN_p}{N_0}\right) \quad (2.2)$$

where M is the ratio between the carrier frequency and the sampling frequency, N_0 is the white noise spectral density and N_p is the Band-Pass Filter (BPF) filtered

version of N_0 . In this sense, the choice of the BPF filter as well as the choice of the sampling frequency become quite critical. Beside this, the phase noise specification of the sampling oscillator becomes quite demanding. Indeed, the phase noise is amplified by M^2 requiring a careful design of the VCO. Accordingly, when interferers are present, a poor phase noise characteristic can degrade the Bit Error Rate (BER) through reciprocal mixing considerably. Consequently, up to now, this architecture has been used mainly in interferer-free scenarios (space applications) [24].

2.1.4 Super-regenerative

Super-regenerative architectures date back to Armstrong, who invented the principle. Despite many years of development, they still suffer from poor selectivity and lack of stability, while having the potential to be low power. Furthermore, they are restricted to OOK modulation techniques only.

In [25] Bulk Acoustic Wave (BAW) resonators are used to reduce the power consumption and to provide selectivity. In spite of achieving an overall power consumption of 450 μ W, it relies on non-standard technologies (BAW resonators), which will increase cost and form factor of the “micro-Watt node”.

In [26] a 1.2 mW super-regenerative receiver has been designed and fabricated in 0.35- μ m CMOS technology. Even though the power consumption is very close to the requirements of a “micro-Watt node”, selectivity is quite poor. Indeed, to demodulate the wanted signal in the presence of a jamming tone placed 4 MHz far from the wanted channel with a BER of 0.1%, the jamming tone has to be no more than 12 dB higher than the desired signal. Generally, to achieve a reliable communication, the receiver should be able to handle interferers which have a power level 40 dB higher than the wanted signal with a BER smaller than 0.1%. This specification is very demanding for a super-regenerative architecture and it requires the use of non-standard components like BAW resonators to achieve a better selectivity.

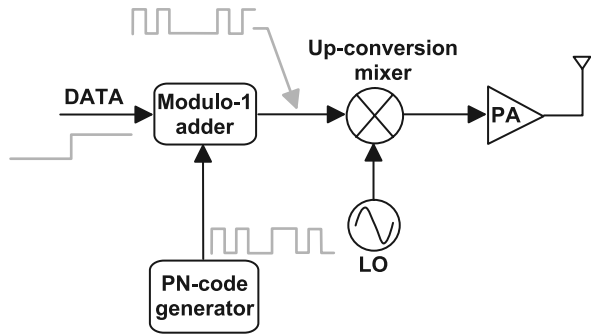
2.1.5 Spread-Spectrum Systems

Any transmission technique in which a Pseudo-random Noise Code (PNC) is used to spread the signal energy over a bandwidth much larger than the information bandwidth is defined as an SS type of transmission. SS techniques are mainly of three types:

- Direct Sequence Spread Spectrum (DSSS)
- Frequency Hopping Spread Spectrum (FHSS)
- Time Hopping Spread Spectrum (THSS)

Sometimes these techniques are combined to form hybrid systems. These hybrid systems are outside the scope of this system given their high complexity. The most widely used systems are of the first two kinds and therefore, this section is restricted to the analysis of both DSSS and FHSS systems.

Fig. 2.2 Schematic block diagram of a DSSS transmitter



Direct Sequence Spread-Spectrum

In DSSS systems, the spreading code is applied to the incoming data. In this way the data symbol is chopped in several parts following a pseudo-random code. Each of these slices within the same symbol period is called a chip. Two quantities are defined in DSSS systems, which are the chip rate $R_c = 1/T_c$ where T_c is the chip duration and the symbol rate $R_s = 1/T_s$, where T_s is the symbol period. The chip rate is an integer multiple of the symbol rate. At every moment, the “instantaneous bandwidth” is equal to the average bandwidth and it is proportional to the chip rate.

A simplified block diagram of a DSSS transmitter is depicted in Fig. 2.2. The same principle is applied on the receiver side where after despreading the data is recovered. To despread the data the receiver must know the PNC sequence and must synchronize in time with it. Therefore, if the receiver does not know the PN sequence, then the received signal will continue to be spread spectrum and the transmitted data cannot be recovered. In this sense a DSSS system is a secure system, which broadens the range of applications in which an ultra-low power node can be used. Another very important parameter in DSSS systems is the so called Processing Gain (PG). The PG is defined as follows:

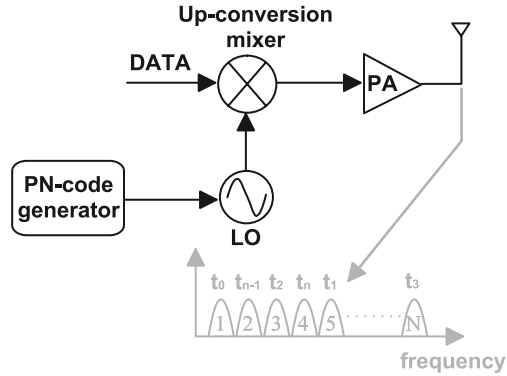
$$G_p = \frac{BW_{ss}}{BW_{info}} = N_c \quad (2.3)$$

where BW_{ss} is the occupied bandwidth after spreading, BW_{info} is the information bandwidth and N_c is the number of chips per symbol period. This parameter has great importance when a DSSS system needs to cope with an in-band interferer usually called a jammer. The effect of interferers will be analyzed later in this section when a comparison between DSSS and FHSS systems will be performed.

Frequency Hopping Spread Spectrum

Differently from DSSS systems, in FHSS systems the spreading code is applied to the frequency domain rather than to the time domain. Therefore, the system hops after a certain amount of time, called dwell time, to another frequency. An FHSS

Fig. 2.3 Schematic block diagram of an FHSS transmitter



system is instantaneously a narrowband system but on the average it is a wideband system.

Important parameters of an FHSS system are the number of channels, the dwell time (T_h) and if the system is a slow hopping or a fast hopping system. The definition of slow hopping or fast hopping is not given in the absolute sense but only in conjunction with the data-rate. A system is considered to be slow hopping if the hopping rate is smaller than the data-rate. When the hopping rate is faster than the data rate the system is called fast hopping.

A simplified block diagram of an FHSS system is given in Fig. 2.3. As in a DSSS system, an FHSS needs a PNC synchronization. To successfully recover the transmitted data the receiver needs to hop coherently with the transmitter. Any FHSS is, therefore, a secure system against intentional jammers trying to steal any information.

DSSS Versus FHSS

In this section the two most common SS techniques are analyzed more in detail and a comparison between them is performed. The reason for such a comparison is driven by the idea to find the optimal SS solution for a power constrained environment. Of course this solution must take into account an interferer scenario composed not only by radios of the same network but also from radios of other standards using the same allocated bandwidth.

Different radio characteristics are further analyzed for both a DSSS radio and an FHSS radio. Those characteristics are the followings:

- Power spectral density and probability of collision
- Susceptibility to the near-far problem
- Radio selectivity
- Robustness to fading conditions
- Robustness to narrowband jammers
- Modulation format and power efficiency
- Acquisition time

A DSSS system is an instantaneously wideband system. Therefore, its transmitted power is spread over a very large bandwidth. Though an FHSS system is on the average a wideband system, instantaneously it operates as a narrowband system. This means that the power spectral density of a DSSS system is lower than that of an FHSS system for the same transmitted power.

In a DSSS system, if two nodes communicate at the same time, they will always interfere with each other. On the other hand, the probability of collision in an FHSS system is the following:

$$P_{\text{coll.,FHSS}} = P_{\text{concurrent-communication}} \times P_{\text{same-frequency}} \quad (2.4)$$

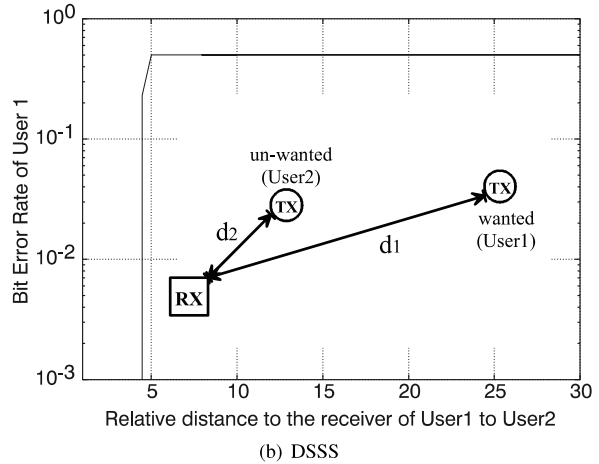
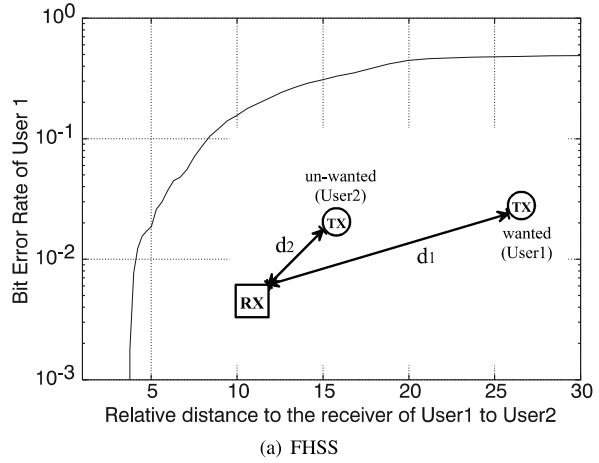
where $P_{\text{concurrent-communication}}$ is the probability that two nodes communicate at the same time (and it is the same for a given network in both DSSS systems and FHSS systems) and $P_{\text{same-frequency}}$ is the probability that two nodes occupy the same frequency bin. This probability is smaller than one and in general it is inversely proportional to the number of available frequency bins in the FHSS system. Therefore, it can be much smaller than one. This translates in a much smaller probability of collision of an FHSS system with respect to a DSSS system.

When a wanted and an unwanted node are communicating at the same time (and on the same frequency bin for an FHSS system) a collision occurs. Depending on the relative received power of the wanted and unwanted node, the data can be retrieved or can be lost because the unwanted node is overwhelming, in terms of received power, the wanted node. This is the so called near-far problem, that mainly appears when the wanted node is much farther than an unwanted node and a collision occurs. Indeed, in this situation the received power of the wanted node can be substantially lower than that of the unwanted node.

Of course the processing gain helps the receiver to distinguish between the wanted node and the unwanted node. For the moment no power control mechanism is supposed. If the wanted node is 3 times further away with respect to the unwanted node then the difference in power equals (see (1.7)) roughly 19 dB. In general DSSS systems use a Binary Phase Shift Keying (BPSK) modulation technique which, for a BER of 1%, requires ideally an $\frac{E_b}{N_0} \simeq 5$ dB where E_b is the energy per bit and N_0 is the noise spectral density. This means, for the above example, that the unwanted signal must be attenuated at least by 24 dB to correctly recover the transmitted data. Supposing now that all the received signals have the same power and considering a processing gain of roughly 12 dB, a maximum of 5 wireless nodes can transmit at the same time without affecting considerably the quality of the link.

FHSS systems, on the other hand, are on the average wideband systems, but instantaneously they are narrowband. Therefore, the susceptibility to the near-far problem is avoided as soon as the data can be transmitted on the next frequency bin and this is not jammed by another interferer. Therefore, while the only way to cope with interferers, for a DSSS radio, is to improve its processing gain, an FHSS system can hop around the problem avoiding it. An increase in the processing gain translates in a much more power hungry digital back-end because of the higher operating speed required. An FHSS system can increase its interferer suppression capability by increasing the number of frequency slots available. This is bounded only by the

Fig. 2.4 Near-far sensitivity comparison between FHSS and DSSS

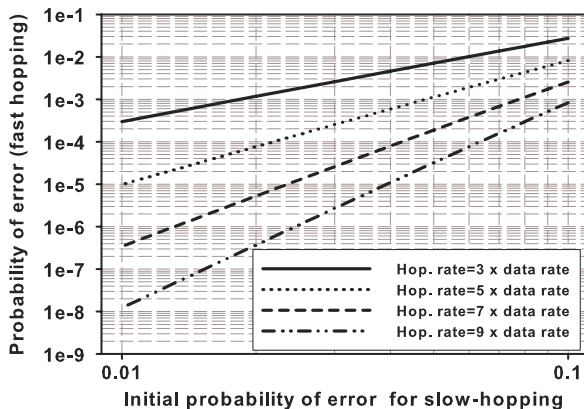


tuning range of the synthesizer, which does not affect power consumption in a first approximation. In Fig. 2.4 two DSSS systems and two FHSS systems are supposed to interfere continuously. The simulation results obtained by using a Simulink model show the clear advantage of an FHSS system over a DSSS system.

Selectivity in an FHSS system is assured by the baseband filter. This means that more nodes can be allocated in the same net. In [27] it is shown that the effective throughput of an FHSS network peaks at a certain number of nodes which is generally larger than in a DSSS system. For example, for the given number of wireless nodes and frequency bins available in the network described in [27], the FHSS network throughput peaks at around 13 nodes. On the other hand, these networks must be placed further away than in the case of a DSSS system. Still, on a single network an FHSS system is more robust than a DSSS system.

SS systems are also known for their capability to cope with fading conditions. A fading condition translates in a very poor SNR, which can be 20 dB below the

Fig. 2.5 Probability of error at variable hopping rate using a majority decision criteria



theoretical SNR calculated in a Additive White Gaussian Noise (AWGN) condition. DSSS suppresses the multipath using again the decorrelation properties of the system. Two PN sequences are uncorrelated only if they are delayed by more than one chip. The delays in a common office environment are in the order of few nanoseconds. This will imply that at low data-rate a very-high processing gain must be applied. Another way is to increase the data-rate but in both cases the digital back-end will suffer from an increase in the power consumption.

An FHSS system copes with the same problem by simply changing the frequency. Because the fading is frequency dependent it is possible by changing frequency to have less adverse fading conditions. This is related to the tuning range and the available bandwidth, more than to power consumption and therefore, it constitutes an advantage of FHSS over DSSS.

The FHSS system has an intrinsic capability to cope against strong narrowband interferers and fading by increasing its hopping rate. Supposing that N hopping channels are available and J out of N are jammed because of strong fading condition or from a large interferer, then if a single bit is transmitted on different channels and a majority decision criteria is used, the probability of error for a given bit is:

$$P_{\text{th}} = \sum_{x=r}^c \frac{c!}{r!(c-r)!} p^x (1-p)^{(c-x)} \quad (2.5)$$

where $p = J/N$, c is the number of bins in which the same bit is sent and r is the number of errors necessary to cause a bit error.¹ In Fig. 2.5 the probability of error versus the initial error probability is shown for different hopping rates. By varying the hopping rate it is possible to decrease by an order of magnitude the BER without increasing the transmitted signal power. In an indoor and interferer crowded scenario this is one of the greatest advantages an FHSS system has.

Another important point of discussion is the modulation format. As it will be shown more into detail in Sect. 2.1.5, an FHSS system generally employs an

¹For example if the hopping rate is three times the data-rate, then 2 errors will cause a bit error. For a hopping rate 5 times the data-rate, 3 errors are necessary to cause a bit error.

FSK modulation while a DSSS system uses a BPSK modulation. The power efficiency of the FSK modulation is larger than that of a BPSK. Therefore, though the BPSK has an advantage over the FSK regarding the robustness against interferers, this advantage is zeroed by the relatively lower power efficiency (see for details Sect. 2.1.5).

As already mentioned, the wake-up time is an important parameter in ultra-low power radios. This wake-up time includes also the synchronization time commonly required in every SS system. When two wireless nodes try to communicate with each other, synchronization of the PNCs must be achieved. At the beginning an offset between the transmitter PN sequence and the receiver PN sequence exists. Therefore, a space of uncertainty exists between the phases of the two PNCs. This space is sliced into “small” pieces called cells. Each cell has a time width so that, when synchronization is achieved, the residual time difference between the transmitter PNC and the receiver PNC code is within half of the chip period or half of the symbol period for a DSSS and an FHSS system respectively. The synchronization algorithm needs to explore those cells in order to find the one for which the phase difference between transmitter PNC and receiver PNC is within the aforementioned range.

It can be proven that the average acquisition time for a SS system, in the case of a serial search synchronization technique with non-coherent detection² is [28]

$$\overline{T_s} = (C - 1)T_{da} \left(\frac{2 - P_d}{2P_d} \right) + \frac{T_i}{P_d} \quad (2.6)$$

where T_i is the integration time for the evaluation of each cell in the time-frequency plane, P_d is the probability of detection when the correct cell is being evaluated, T_{da} is the average dwell time at an incorrect phase cell, C is the total number of cells. This formula can be intuitively explained supposing that the probability of false alarm is zero. In this case the average dwell time at an incorrect phase cell equals the integration time on the cell. The total number of incorrect cells is $C - 1$ and on each one the integration time is spent in the evaluation. Of course, this is just the worst case condition. On the average, there is a certain probability to find the correct cell before all the C cells are evaluated. If the probability of a cell to be the correct one is uniform, and the probability of detecting the correct cell is one (given the fact that the probability of false alarm is zero), then the expression between brackets in (2.6) approaches 0.5. This value makes sense given the fact that an uniform distribution is supposed. This means that, for each cell, there is always a 50% probability that it is the correct one. When the probability of false alarm is non zero, then the average dwell time at an incorrect cell increases. The probability of detection decreases, increasing the term between brackets in (2.6). Therefore, globally the acquisition time increases. The last term in (2.6) reflects the time spent in the evaluation of the correct cell. It is equal to the integration time if

²As it will be proven in Chap. 3 the serial acquisition technique is the most power efficient algorithm for PNC acquisition.

Table 2.1 Summary of the comparison between DSSS and FHSS systems

Radio characteristic	DSSS	FHSS
Power spectral density	Low	High
Probability of collision	High	Very low
Near-far robustness	Low	High
Selectivity	Medium	High
Fading robustness	Medium	High
Narrowband jammer robustness	Medium-low	Very high
Modulation power efficiency	Medium	High
Acquisition time	High	Low

the probability of detection is 1 and larger if there is a chance to skip the correct cell. For this reason, it has to be inverse proportional to the probability of detection.

Now, assuming that no frequency uncertainty is present, there will be a time misalignment between the two PN sequences at the transmitter and receiver side equal to ΔT_i . Therefore, while for a DSSS the system has to be synchronized within $\pm T_c/2$, an FHSS system needs to be synchronized within $\pm T_s/2$.³ Due to the fact that in a DSSS system the processing gain is related to the ratio between the chip rate and the symbol rate, the chip period is at least an order of magnitude smaller than the symbol period. As a result, the number of cells that must be evaluated in a DSSS system is considerably larger than in an FHSS system. From (2.6) the mean DSSS synchronization time is larger than in the case of an FHSS system. This will increase the wake-up time and therefore, the overall system power consumption.

A summary of the comparison between FHSS and DSSS systems is given in Table 2.1. It is clear that the FHSS technique presents a clear advantage over the DSSS technique for most of the characteristics listed in Table 2.1. For this reason, it is more suitable for power constrained, indoor, interferer crowded scenario like the one foreseen in wireless sensor networks.

Modulation Formats

Several modulation formats can be used in digital communications. The relative implementation complexity of various modulation schemes is depicted in Fig. 2.6. Given the power constrained environment it is important to select a low complexity modulation technique. Therefore, three modulation formats are analyzed more in detail:

³The remaining part of the synchronization consists in what is generally called tracking. During tracking, the phase difference between the two PNCs is reduced to virtually zero from a closed-loop system (like a PLL). This system also tracks any instantaneous variation of the phase of the transmitter PNC in order to assure a constantly aligned PN sequences between the transmitter and the receiver when the nodes are communicating with each other.

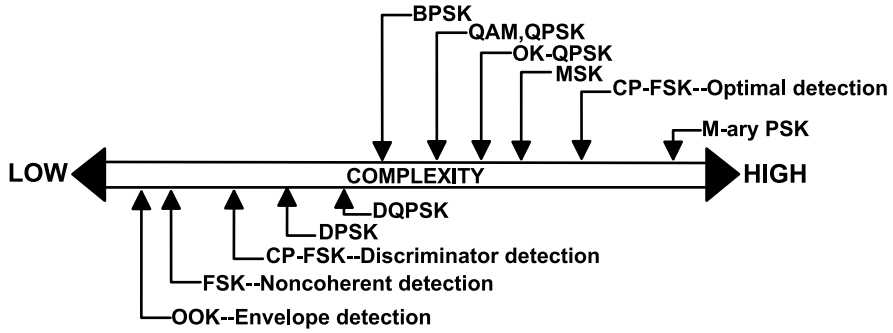


Fig. 2.6 Relative complexity of various modulation scheme (adapted from [29])

- OOK with envelope detection
- FSK with non-coherent detection
- BPSK

Coherent-Phase Frequency Shift Keying (CP-FSK), Differential Phase Shift Keying (DPSK) and Differential Quadrature Phase Shift Keying (DQPSK) are derivatives of those formats and therefore, though their complexity is not high, they will not be further analyzed in this book. OOK, FSK and BPSK modulations cover also all of the most common types of modulation formats. In fact OOK is a type of amplitude modulation, FSK is a frequency modulation and BPSK is a phase modulation. A comparison between these modulation schemes can be done based on an ideal required SNR for a given BER,⁴ signaling speed, robustness against Continuous Wave (CW) interferers, and robustness in a Rayleigh fading channel. The summary is shown in Table 2.2. The performances of OOK and FSK in a AWGN environment are very similar while the BPSK modulation has roughly 4 dB better performance. In a Rayleigh fading environment this gain reaches roughly 6 dB. When interferers are present, as it will happen in the 915 MHz and 2.4 GHz ISM bands, then the OOK modulation happens to be a very weak scheme requiring 5 dB more SNR than FSK and almost 10 dB more than BPSK. Therefore, FSK is more suitable than OOK in an interferer crowded scenario. BPSK has a 4 dB advantage in an interferer dominated scenario over the FSK modulation.

Every frequency modulated signal is a truly constant envelope signal, while BPSK modulated signals contain some amplitude modulation in their modulated envelope. Therefore, while FSK signals can be amplified by a Power Amplifier (PA) operating near the saturation level, BPSK signals require a 3 to 6 dB back-off from this level. This is necessary in order to eliminate the spectral regrowth, which will cause adjacent channel interference. This translates in a smaller power efficiency and therefore, the gain of BPSK modulated signal over FSK signal is practically more than compensated by this drawback.

⁴With the word “ideal” here it is supposed that the channel is AWGN.

Table 2.2 Performance comparison of some low complexity modulation formats for a $BER = 10^{-4}$ (adapted from [29])

Modulation	Speed [bitHz/s]	$\frac{E_b}{N_0}$ (AWGN) [dB]	$\frac{E_b}{N_0}$ ($S/I = 10$ dB) [dB]	$\frac{E_b}{N_0}$ (Rayleigh) [dB]
OOK-Envelope det.	0.8	11.9	20	19 ^a
FSK-Non-coh. ($m = 1$ ^b)	0.8	12.5	14.7	20
BPSK	0.8	8.4	10.5	14

^aWith optimum variable threshold^b m = modulation index

As stated in Sect. 2.1.5 this is a big advantage for an FHSS system over a DSSS system. Implementation of an FSK modulation format on an FHSS system is trivial given the fact that frequency hopping is very close to a frequency modulation scheme.

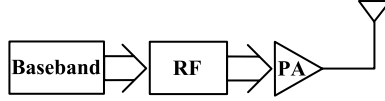
2.2 Optimal Data-Rate

The aim of this section is to find the data-rate, which minimizes the average node power consumption. The transceiver is composed by a receiving section and a transmitting section.

The receiver sensitivity depends on the Noise Figure (NF), noise bandwidth and the required SNR of the demodulator. The NF depends on the receiver architecture and technology used and in an asymmetric scenario can be considered to be smaller than 10 dB. For a chosen modulation scheme and a certain desired BER, the required SNR is fixed, apart from implementation losses in the demodulator. The only parameter left is the noise bandwidth which ultimately affects the system data rate.

A transmitter can be partitioned into three domains in terms of power consumption. This partitioning is shown in Fig. 2.7. The baseband part has a power consumption, which is proportional to the data-rate. The RF portion, is needed to upconvert the baseband information to the wanted high frequency band. Therefore, its power consumption depends on the operation frequency and it is independent of the data rate. Finally, the PA section is needed to transmit the information over the medium. Its power consumption depends on its efficiency and on the required transmission range. The efficiency, is strictly related to the modulation format used and it is optimal for constant envelope formats like FSK. Generally the most power hungry blocks are the PA and the RF section which consists of the synthesizer and a mixer for up-conversion. The first two domains are what is defined later in this section as pre-PA domain.

Fig. 2.7 Transmitter power domains



In general a transmitter first needs a time T_{wu} to wake up, then T_{tx} to transmit, and after transmission it remains T_{idle} in the idle mode till the next transmission cycle is started. So, the time between two consecutive transmissions, T , equals $T_{wu} + T_{tx} + T_{idle}$. The duty cycle of the system, “ d ” can be defined as the ratio between the time required to transmit the data and the time between two consecutive transmissions.

Several parameters are involved in the derivation of the optimal data-rate. Some parameters are fixed. Other parameters, although can be varied and optimized for low power, are also considered fixed. Finally the variable we need to optimize is the data-rate. Parameters, which are fixed are the following:

- $\frac{B_{noise}}{B_{data}}$
- N_0
- $\frac{E_b}{N_0}$
- $L_{path,nat}$

where B_{noise} ⁵ is the noise bandwidth B_{data} ⁶ is the data bandwidth, $L_{path,nat}$ represents the path losses due to propagation expressed in natural units.

Parameters which can be optimized for low power but are considered fixed in this discussion are the following:

- L_{pack}
- P_{diss}
- P_{idle}
- T_{wu}

where L_{pack} is the data packet length, P_{diss} is the power consumption of the remaining transmitter circuitry during wake up and transmission excluding the PA and P_{idle} is the power consumption in the idle mode.

It is possible to suppose that the duty-cycle of the network is constant [30], or that the time between two consecutive transmissions is constant [31]. These two cases will be further analyzed.

⁵The 2-FSK noise bandwidth can be approximated by the Carson rule as $B_{noise} = 2(\Delta f + f_m)$ where $f_m = \frac{2}{T_s}$ with $\frac{1}{T_s}$ the data rate and Δf the frequency deviation.

⁶For a 2-FSK modulated signal it equals four times the data rate.

2.2.1 Constant Duty-Cycle

The average power consumption of the transmitter node can be approximated by the following equation:

$$P_d = P_{tx} \frac{T_{tx}}{T} + P_{diss} \frac{(T_{tx} + T_{wu})}{T} + P_{idle} \frac{(T - T_{tx} - T_{wu})}{T} \quad (2.7)$$

where P_{tx} is the PA power required for the transmission of data. The transmission time depends on the packet length L_{pack} and on the data rate “ D ”:

$$T_{tx} = \frac{L_{pack}}{D} \quad (2.8)$$

The required transmitted power can be written as

$$P_{tx} = N_0 \times \frac{B_{noise}}{B_{data}} \times \frac{E_b}{N_0} \times NF \cdot D \cdot L_{path,nat} = K \times D \quad (2.9)$$

Given the fact that both B_{noise} and B_{data} are proportional to the data rate, their ratio is constant and the transmitted power P_{tx} is direct proportional to the data rate D via the constant K in (2.9). From these considerations, (2.7) can be re-written in the following way:

$$P_d = (K D + P_{diss}) \times d + P_{idle}(1 - d) + (P_{diss} - P_{idle}) \times d \frac{T_{wu}}{L_{pack}} D \quad (2.10)$$

where the duty cycle “ d ” is a constant in this discussion.

From (2.10), it can be seen that for a fixed transmission distance ($\propto P_{tx}$), the power consumption can be reduced by reducing the duty cycle, the data-rate, and by making the wake-up time small compared to the transmission time. This last requirement becomes difficult to achieve in SS systems at high data rates due to PNC synchronization. Therefore, reducing the data rate will help to relax the wake-up time for a given $\frac{T_{wu}}{T_{tx}}$. The transmitter average power consumption as a function of the data-rate for different duty-cycles is plotted in Fig. 2.8. At high data rates, the average power consumption is dominated by the transmitted power. At data rates below a threshold value the average power consumption is dominated by the pre-PA power. This threshold value depends on the pre-PA power dissipation and it is lower for lower values of the pre-PA power dissipation.

From Fig. 2.8, when the pre-PA power dissipation (P_{diss}) is 2 mW, this threshold value is around 100 kbps, while at pre-PA power of 10 mW it is located around 1 Mbps. At higher data rates, the wake-up time has to decrease considerably to keep the contribution to the average power consumption negligible. Therefore, from the previous analysis it is possible to conclude that a good strategy toward the reduction in the average transmitter power consumption consists of reducing the data-rate and decreasing the synchronization time for a given node duty-cycle.

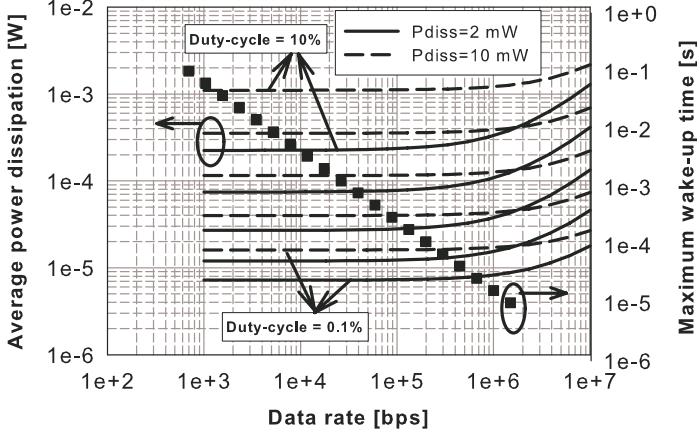


Fig. 2.8 Average transmitter power consumption and maximum wake up time for different value of duty cycle, as a function of the data rate ($L_{\text{pack}} = 1000$ bits, $NF = 10$ dB, $\frac{E_b}{N_0} = 20$ dB, carrier frequency = 915 MHz, $P_{\text{idle}} = 10 \mu\text{W}$) [30]

2.2.2 Constant Time Between Two Consecutive Transmissions

Some applications may require a fixed time between two consecutive transmissions. In this case the time “ T ” is constant while the duty cycle “ d ” varies decreasing by increasing the data rate D . For example, for a temperature sensing inside an apartment a fixed interval between two consecutive transmissions may be sufficient. Equation (2.10) can be rewritten now with the time difference between two consecutive transmissions as a variable instead of the duty cycle:

$$P_d = K \frac{L_{\text{pack}}}{T} + P_{\text{diss}} \left(\frac{L_{\text{pack}}}{D \cdot T} + \frac{T_{\text{wu}}}{T} \right) + P_{\text{idle}} - P_{\text{idle}} \frac{L_{\text{pack}}}{D \cdot T} - P_{\text{idle}} \frac{T_{\text{wu}}}{T} \quad (2.11)$$

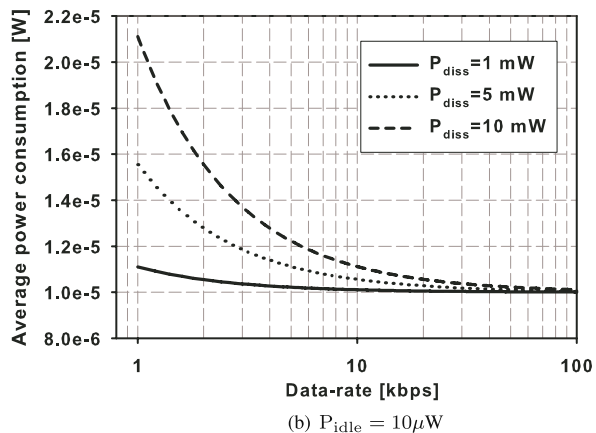
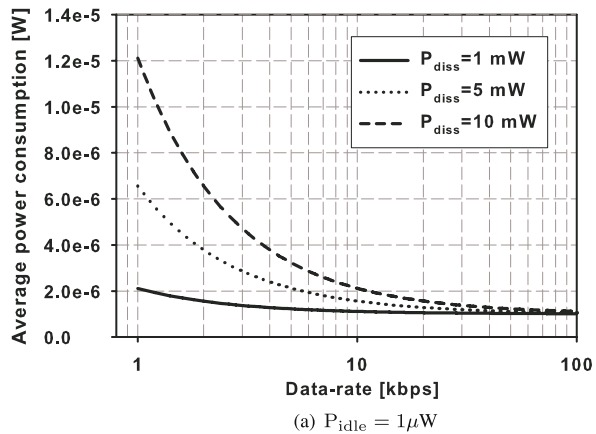
and for T sufficiently large it can be approximated as follows:

$$P_d \approx K \frac{L_{\text{pack}}}{T} + P_{\text{diss}} \left(\frac{L_{\text{pack}}}{D \cdot T} + \frac{T_{\text{wu}}}{T} \right) + P_{\text{idle}} \quad (2.12)$$

The simulation results are given in Fig. 2.9 as a function of the data rate for a given (constant) value for T and T_{wu} . As it can be seen from Fig. 2.9, increasing the data rate will make the average power consumption smaller and smaller. A limit is dictated by the idle power. This is easily understandable by looking at (2.12). When the idle power is $1 \mu\text{W}$, a data rate between 1 and 10 kbps is sufficient to not spoil the average power consumption for P_{diss} ranging between 1 and 10 mW. In this way, in the worst condition ($P_{\text{diss}} = 10$ mW) the average power consumption is determined in equal parts by the idle power and the power used during transmissions. When the idle power increases to $10 \mu\text{W}$, then the data-rate can be relaxed down to around 1 kbps in all the cases.

For high data-rate the pre-PA power consumption can become a strong function of the data-rate and the term P_{diss} cannot be any longer considered constant but

Fig. 2.9 Average power dissipation versus data rate ($L_{\text{pack}} = 1000$ bits, $T_{\text{wu}} = 500 \mu\text{s}$, $NF = 10$ dB, $\frac{E_b}{N_0} = 20$ dB, carrier frequency = 915 MHz, $T = 300$ s) for $1 \mu\text{W}$ and $10 \mu\text{W}$ idle power dissipation



will be a function of the data-rate. In the newest technologies it is reasonable to say that below 100 kbps the transmitter baseband part will consume much less than the RF part and, therefore, the approximation can be considered valid. It should be noticed that in this discussion it has been neglected that, on the receiver side, the power consumption is a function of the data-rate for all the baseband domain. Nevertheless, looking at the receiver, it is a reasonable choice as optimum data-rate the lowest possible, which fulfill the aforementioned considerations. Indeed at the receiver side several analog blocks (like Voltage Gain Amplifier (VGA), filters and ADC) work at baseband frequency and their power consumption is proportional to the data-rate.

While it seems that an idle power in the μW range is too high, it should be noticed that most probably a wake-up type of radio will be used. Some circuitry will be kept on listening the channel for an incoming transmission especially in asynchronous networks. Therefore, a power budget for this circuitry between 1 and 10 μW is a good choice. The considerations regarding the synchronization time for the case of constant duty-cycle networks do apply also to this case. Moving towards lower

data rates helps in relaxing the synchronization constraints. From all the previous considerations it is possible to conclude that a data rate between 1 and 10 kbps is sufficient to optimize the average node power consumption.

2.3 Transmitter Architectures

Generally, the transmitter part is under-estimated in terms of complexity and number of possible system and circuit trade-offs with respect to the receiver part. Any choice on the transmitter side will have a great impact on the receiver specifications as well. The transmitter topology of choice is a result of various trade-offs impacting the maximum level of unwanted frequency components, efficiency of the PA, linearity and maximum power consumption. The transmitter performs some operations before radiating the signal toward the receiver, which can be summarized as follows:

- Modulation
- Up-conversion
- Power amplification

Modulation has been already discussed in the previous section. It has been pointed out that at low data rates the information bandwidth is not a problem. Therefore, the choice has to be directed towards a power efficient modulation scheme rather than a bandwidth efficient scheme. Consequently, a constant envelope modulation scheme like FSK is preferable.

The up-conversion is the action with which the baseband signal is shifted to the wanted frequency band (915 MHz or 2.4 GHz) before transmission. The power amplification needs to increase the power of the radiated signal in such a way that the radiated signal in the worst distance conditions (for example 10 meters indoor) reaches the receiver at a power level equal or above the sensitivity level.

Transmitter architectures can be grouped in three main categories:

- Direct conversion
- Two-step conversion
- Offset PLL

2.3.1 Direct Conversion

A direct conversion transmitter is plotted in Fig. 2.10. The modulated data is directly up-converted to the wanted band. Therefore, the oscillator is running at the carrier

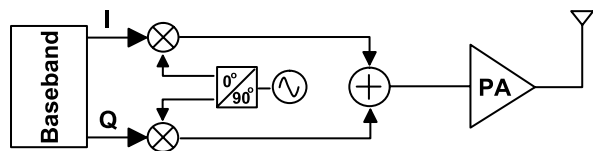


Fig. 2.10 Direct conversion transmitter

frequency. In this type of transmitters can be also included the direct modulated VCO type of transmitters. In this case the baseband data directly modulates the VCO frequency following the data stream. A straightforward implementation of this direct up-conversion scheme employs an FSK modulation type. Direct VCO modulated transmitters merge the modulation and the up-conversion phases into one single operation.

The simplicity of this architectures makes it attractive for a high level of integration, which means lower costs and lower power consumption.⁷ Unfortunately, a big drawback of this architecture is a phenomenon called “Oscillator pulling”. If the output of the PA and the oscillator frequency are very close to each other in frequency, the oscillator is heavily disturbed by the noise coupling back from the PA output. Even a noise level 40 dB below the oscillator level can cause an enormous amount of disturbance on the oscillator output. For this reason generally the two frequencies must be far apart and mostly uncorrelated. Several solution are possible to accomplish this result:

- The up-conversion signal is obtained by an integer division of the oscillator signal.
- The up-conversion signal is obtained by a non-integer division of the oscillator signal.
- The up-conversion signal is obtained by mixing two non divisible oscillator frequencies.

The first option is the most simple one to implement at hardware level. The PA output and the oscillator output are far apart, but the two frequencies are still correlated. This can give still some pulling especially at high noise level. The second option makes the two signals far apart and well uncorrelated. Unfortunately a non integer division requires complex hardware. For example, a multiplication by two followed by a divide-by-three stage can accomplish this result. The last option gives the best results but at the expense of two oscillators, one mixer and one BPF, required to suppress the unwanted harmonics generated from the mixing process.

The most power efficient solution, therefore, is the first option if a good degree of isolation can be guaranteed between the PA output and the oscillator in order to avoid a high level of noise injection in the oscillator.

2.3.2 Two-Step Conversion

A block diagram of a two-step conversion transmitter including the frequency spectrum at various points in the chain is shown in Fig. 2.11. This architecture eliminates the problem of VCO pulling by splitting the up-conversion into two phases.

⁷External components require in general matching to 50 Ω . This means that the stage preceding the external component must be able to drive a 50 Ω impedance. Such a low impedance level will cost a high current consumption.

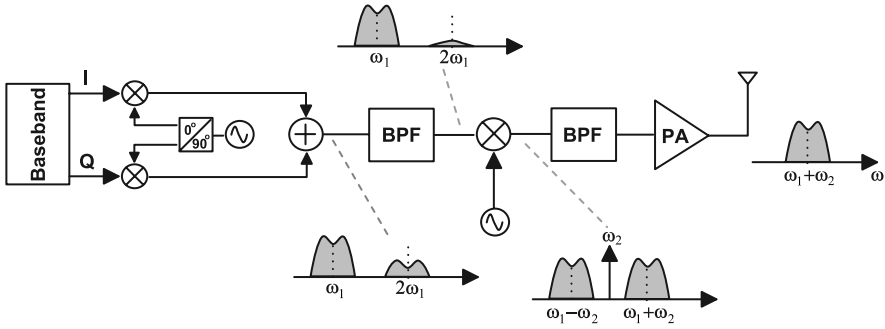


Fig. 2.11 Two-step conversion transmitter (adapted from [32])

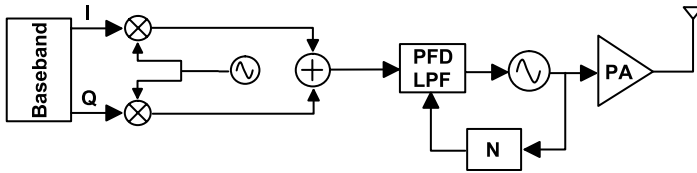


Fig. 2.12 Offset PLL architecture

An advantage of this architecture is that the quadrature up-conversion to the so called intermediate frequency is performed at a lower frequency compared to the direct up-conversion scheme. This greatly improves the matching of the I - Q signals. The main drawbacks of this architecture are first the increased amount of hardware needed and second the BPF before the PA. This BPF needs to attenuate the unwanted sideband more than 40 dB. Therefore, it generally requires an expensive and power hungry off-chip component.

2.3.3 Offset PLL

The schematic block diagram of an offset PLL based transmitter is depicted in Fig. 2.12. The main restriction of this kind of transmitter architecture is that it can be used only with constant envelope type of modulated signal. In this architecture, the PLL acts as a narrowband filter, rejecting all the high frequency noise coming from external sources. Therefore, the high frequency noise is generally determined by the VCO noise but this also happens in all the previously mentioned transmitter architectures. Unfortunately, the divider in the PLL chain, which is used to ease the design of the PFD can cause some severe drawbacks. Any phase modulation at the input of the PLL is amplified by a factor N in amplitude at the output of the PLL. The same happens to any change in the frequency of the baseband VCO. This means that if more channels are used, the baseband synthesizer must have a very fine tuning mechanism, which in general means longer settling times and higher average

power consumption. Therefore, this technique tends to be quite cumbersome to use especially in SS systems.

2.4 Receiver Architectures

The receiving part of a transceiver generally can consume a large amount of power if not optimized for a power constrained environment. Several trade-offs are also present in the choice of a suitable architecture for the receiver. The level of integration is an important aspect, which reduces the costs by eliminating bulky external components while also achieving a decrease in the peak power consumption. The level of achievable integration depends in general by a combination of receiver specifications and receiver architecture. Three main receiver architectures are generally used in modern wireless communication:

- Zero-IF
- Super-heterodyne (or heterodyne)
- Low-IF

2.4.1 Zero-IF

The zero-IF architecture has always been considered very suitable for integration. A schematic block diagram is depicted in Fig. 2.13. Unfortunately several issues can make cumbersome its implementation at transistor level. The first major drawback is the so called DC offset. The signal is down-converted to baseband in a single step. Therefore, any DC component is within the signal bandwidth. Now the self-mixing effect of the oscillator frequency via capacitive coupling to the LNA input creates a DC component at the input which amplified by the LNA can reach the millivolt level. If the gain after the LNA is big enough, the DC component so generated can saturate the receiver. The effect is that the weak signal (which is still in the hundreds of microvolt range) cannot be detected properly.

A simple way to eliminate this problem is again to choose correctly the modulation type in such a way that almost no signal energy is placed at DC. In this way, a simple HPF with a corner frequency of few kilohertz can eliminate the offset. Most of the modulation formats, unfortunately, contain much energy around DC. On the other hand, an FSK modulated signal with a large modulation index ($m > 1$) has almost no energy placed at DC. This is another point in favor of architectures that

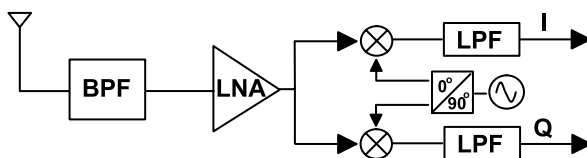


Fig. 2.13 Zero-IF architecture

can easily implement such kind of modulation format within a frequency diversity scheme like it happens for FHSS systems.

Second order distortion is also problematic in zero-IF receivers. A second order distortion in the LNA translates two closely spaced strong interferers to a lower frequency before the mixing process. This component passes through the mixer with finite attenuation due to imperfections in the mixer I - Q matching or Local Oscillator (LO) duty-cycle imperfections. This can again pose a problem to the receiving stage saturating it.

I - Q mismatch between the two quadrature mixers due to the presence of parasitics and due to the high operating frequency is also a big issue. Careful layout is generally needed, especially at high frequencies.

Lastly, a problem which depends on the technology used can come from the noise. For example CMOS transistors are affected by the so called flicker noise. The signal after the LNA is still generally quite low⁸ and therefore, very sensitive to noise. Flicker noise is dominant at low frequencies, which is exactly the frequency region where this architectures tends to directly translate the wanted signal to. Looking at [33] it is possible to see that the flicker noise corner frequency increased from around 1 MHz at 1.2 μm channel length to more than 100 MHz for 30 nm channel length. This means that technology scaling, while helping in reducing the power consumption especially in the digital domain gives more and more drawbacks when an analog block needs to be designed. This is a point which must be greatly taken into account when conceiving an architecture for ultra-low power wireless nodes in CMOS technology.

2.4.2 Super-heterodyne

A schematic block diagram of a super-heterodyne receiver architecture is depicted in Fig. 2.14. The heterodyne principle first down-converts the signal to an intermediate frequency called IF, and after a BPF and a further signal amplification it down-converts the IF signal to baseband. In the case of digital modulation the I and Q signal components are generated in the latest down-conversion stage.

This architecture alleviates some common drawbacks seen in the zero-IF architecture. For example the DC offset coming from the first two stages is filtered out by the BPF, while the one of the last stage is negligible thanks to the high gain of the first two stages. Because the I - Q signals are generated only in the last stage where the frequency is lower, the I - Q mismatch can be easily controlled and reduced to a very low level. Finally, this architecture has a very good selectivity achieved by splitting the filtering among different stages at progressively lower frequency.

Despite all these advantages, the heterodyne architecture suffers of some major drawbacks as well. The biggest problem is the image rejection problem. This issue

⁸Considering a required sensitivity level of -76.5 dBm at 2.4 GHz, then the signal at the output of an LNA with gain equal to 15 dB is generally smaller than 1 mV rms.

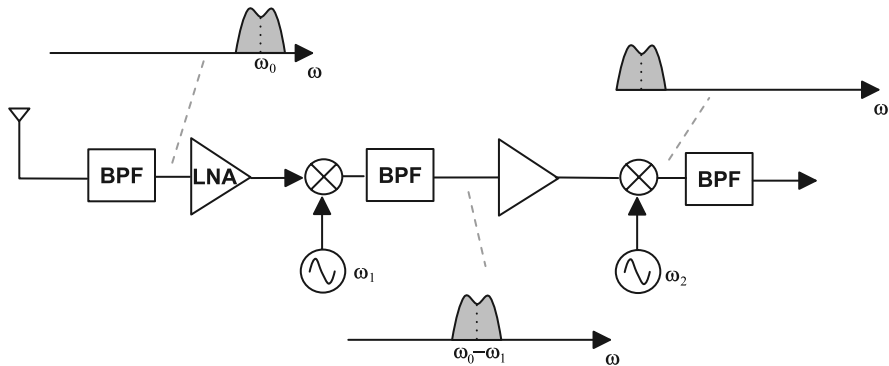
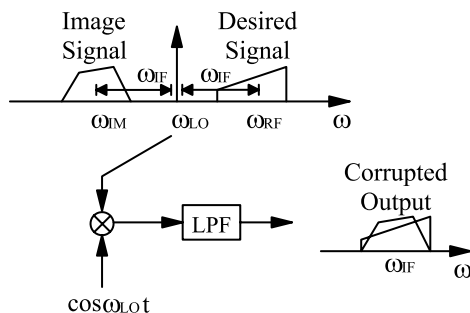


Fig. 2.14 Super-heterodyne architecture

Fig. 2.15 Image frequency problem



can be easily understood noting that all the signals at a distance equal to ω_{IF} from the LO frequency will be down-converted to the same IF frequency. This is illustrated in Fig. 2.15. To suppress the image frequency, an image reject filter is often used. The choice of the IF is not trivial and it entails a trade-off between sensitivity and selectivity. Indeed, to reduce the image noise, the IF frequency has to be chosen as large as possible. On the other hand, this will increase the constraints on the band selection filter coming after the first mixer. The higher the IF frequency, the greater the required Q for a given attenuation.

Generally, the image reject filter precedes the mixer and it is realized using external components. This translates in a worse transceiver form factor and it requires the LNA to drive a $50\ \Omega$ impedance. This, in turn, translates in higher power consumption due to a larger required bias current.

2.4.3 Low-IF

The low-IF architecture is closely related to the zero-IF topology and it tries to minimize the major drawbacks of the zero-IF topology by operating near the DC but not at DC. In a low-IF topology, the problem of the image suppression, which

burdens the heterodyne receiver, can be shifted to the IF stage. In this way, given the lower operating frequency the high-frequency BPFs can be integrated given the lower required Q .

On the other hand, this topology tends to shift the demanding specifications to the ADC preceding the Digital Signal Processor (DSP) in which the final demodulation is performed. This can be seen as a reasonable way to cope with the power reduction problem. Indeed, the ADC converter, is a mixed signal block in which lots of functions are anyhow performed in the digital domain. Therefore, it is more prone to power scaling with technology advances than a purely analog block.

Concluding, the low-IF topology is a good alternative to the zero-IF topology especially when problems like flicker noise and DC offset become a show stopper for further power reduction.

2.5 Conclusions

Though ultra-low power wireless nodes have very tight constraints in terms of power consumption, they should allow for a robust wireless link in the harsh indoor environment. Moreover, because it is necessary to offer to the end user a very low cost solution, ISM bands are generally used because they are license free. This option, however, presents the inconvenience of a very interferer crowded environment for the wireless radio.

It has been shown, in this chapter, that to cope with those strong non-idealities, while keeping the power consumption very low, a combination of modulation format, transmitter and receiver architectures and wideband techniques can be used. It has been proven, indeed, that spread spectrum techniques are an optimal choice to have a robust wireless link. Among several SS techniques, it has been shown that an FHSS system can offer a very robust wireless link while having the potential to be low power especially if it is combined with an FSK modulation. The possibility to trade between hopping rate and transmitted power is an unique advantage of this system, which allows for a not negligible reduction of the transmitted power without affecting the reliability of the link.

The choice of the data-rate affects also the average power consumption of the wireless node. It has been shown that increasing the data-rate can help in reducing the average power consumption. On the other hand, it has been demonstrated that it is useless to increase the data-rate above a certain level dictated by the idle power, because at this point the average power consumption is always dominated by the idle power. Increasing the data-rate above this threshold level, however, costs an increase in the receiver power consumption because several baseband blocks have a bandwidth, which is a function of the data-rate.

On the transmitter side, it has been proven that a single up-conversion scheme is the most appropriate choice to minimize the average power consumption of the wireless node. On the receiver side two options are foreseen. The best option in terms of integrability and power consumption is a zero-IF architecture. However, especially in the newest CMOS technology, flicker noise and other second order

effects, can cause severe difficulties in its physical implementation. For this reason, a low-IF topology, though it can be more power hungry, results in a good choice when the zero-IF becomes too problematic to be implemented at transistor level.

Concluding, this chapter has found the optimal system level choices for an ultra low power wireless node for wireless sensor area networks: an FHSS system with Binary Frequency Shift Keying (BFSK) modulation and a data-rate below 10 kbps, a single up-conversion transmitter architecture and a zero-IF (or a low-IF) receiver architecture. The following chapter focuses on FHSS systems and the limitations existing in the state-of-the-art solutions in terms of power consumption.

Architectures and Synthesizers for Ultra-low Power Fast
Frequency-Hopping WSN Radios

Lopelli, E.; van der Tang, J.; van Roermund, A.H.M.

2011, XII, 236 p., Hardcover

ISBN: 978-94-007-0182-3