Optimizing RF front ends for low power

Citation for published version (APA):

DOI:
10.1109/5.888994

Document status and date:
Published: 01/01/2000

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Optimizing RF Front Ends for Low Power

PETER G. M. BALTUS, MEMBER, IEEE, AND RONALD DEKKER

This paper discusses optimizations for the power dissipation of RF front ends in portable wireless devices. A breakthrough in power dissipation can be achieved by simultaneously optimizing the antenna interface, circuits, and IC technology of such devices. A model that predicts the minimum power dissipation of a front end for both short-range and long-range connections will be introduced. Using these models, the impact of the antenna interface on the power dissipation will be assessed. Using two antennas with equal gain combining, a typical power dissipation reduction of 2.5 to 30 times can be achieved. Using high-impedance circuits for short-range systems in combination with silicon-on-anything technology, a further reduction of power dissipation by up to one order of magnitude can be realized.

Keywords—Antenna diversity, design methods, low power, RF front end, silicon-on-anything (SOA), silicon-on-insulator (SOI).

I. INTRODUCTION

In this paper, the design of low-power receiver front ends will be discussed. In the first section, the function and relevance of the front end within a transceiver will be defined. In many applications, the front end is responsible for more than half of the power dissipation. Power dissipation of the transmitter and receiver front end are quite different, and the receiver front end in particular is still far away from the ultimate goal of 100% power efficiency.

Improvements of more than an order of magnitude can be achieved by simultaneously optimizing the antenna interface, and the circuits and IC technologies used to implement the front-end electronics. Optimizing the antenna interface requires a method of predicting the influence of changes in the antenna interface on the power dissipation. Models that can be used for this purpose are introduced in Section II. From the developments in technology and limitations in propagation of radio signals, it is to be expected that a split will occur between short- and long-range wireless connections. Since the power dissipation is limited in different ways for short- and long-range systems, different models for these systems are introduced.

These models are used in Section III to evaluate the impact of the antenna interface on power dissipation. Optimizations at circuit and IC technology level are also discussed. In Section IV the silicon-on-anything (SOA) technology and its impact on power dissipation is discussed.

To demonstrate the impact of the combination of these methods on the power dissipation of an RF front end, a design example of a short-range 2.5-GHz antenna diversity receiver in SOA is shown in Section V.

A. RF Front End

A major and rapidly growing use for RF front ends is in portable wireless devices, such as cellular and cordless phones. This paper will therefore concentrate on front ends in such portable wireless devices. They typically run from a battery that is often the largest and most expensive component in the device. The main design problem is to keep such devices small and cheap while achieving the desired functionality and performance.

The function of an RF front end is to provide a convenient interface between electromagnetic fields and (often digital) signal processing. For most telecommunications systems, this interface is bidirectional since it consists of both a transmitter and receiver front end. A front end in a transceiver context is shown in Fig. 1.

These models are used in Section III to evaluate the impact of the antenna interface on power dissipation. Optimizations at circuit and IC technology level are also discussed. In Section IV the silicon-on-anything (SOA) technology and its impact on power dissipation is discussed.

To demonstrate the impact of the combination of these methods on the power dissipation of an RF front end, a design example of a short-range 2.5-GHz antenna diversity receiver in SOA is shown in Section V.

A. RF Front End

A major and rapidly growing use for RF front ends is in portable wireless devices, such as cellular and cordless phones. This paper will therefore concentrate on front ends in such portable wireless devices. They typically run from a battery that is often the largest and most expensive component in the device. The main design problem is to keep such devices small and cheap while achieving the desired functionality and performance.

The function of an RF front end is to provide a convenient interface between electromagnetic fields and (often digital) signal processing. For most telecommunications systems, this interface is bidirectional since it consists of both a transmitter and receiver front end. A front end in a transceiver context is shown in Fig. 1.

The main signal processing functions of an RF front end are gain (to convert the usually weak signals to convenient amplitude levels for further processing) and frequency conversion (to convert signals to convenient frequencies for further processing). In the receive path, selecting the desired channel among (many) other channels, and extracting the information that is applied through modulation to the radio
Fig. 2. Inside of GSM handset.

signal, is usually carried out in the IF signal processing circuits. In the transmit path, modulating the information to be transmitted onto a radio signal is often also carried out in the IF circuits.

B. Relevance of the RF Front End

An RF front end consists of a very limited number of active devices, often in the order of hundreds or thousands. The silicon area of the front end decreases with improvements in IC technology that causes devices to shrink in physical dimensions, and currently measures just a few square millimeters. Also, the desired functionality and therefore the number of active devices decreases with advanced transceiver architectures such as single-conversion receivers, and further digitization of the radio. Compared to the millions of transistors in the digital parts of a handset and the millions of bytes in the software, it might be assumed that the design of an RF front end is a trivial job, with little relevance for the overall handset. From such a viewpoint, it would be hard to justify any research in this area. Fortunately for RF researchers, the RF design problem is highly relevant. Fig. 2 shows the inside of a modern cellular (GSM) terminal.

As can be seen from the “inside” picture (Fig. 2), the front-end part contributes significantly to the size of the total handset. Also, the total power consumption is dominated by the radio rather than the digital circuits, which has an additional impact on size and cost through the battery size: both in transmit and in receive mode, more than 60% of the total power is consumed by the radio in a typical cellular handset. Therefore, the RF front end has a major impact on achieving the design goals of a small and cheap phone.

Much effort has already gone into reducing the power dissipation of RF front ends for this type of application, and significant improvements have been made over the years [1]–[10]. On the other hand, there seems to be no reason why, ultimately, the power dissipation should be constrained by other limits than the law of conservation of energy and imperfections in implementation technology.1

C. Power Dissipation of the Transmitter Front End

Transmitter efficiencies of over 40% are currently found in publications [11], [12], suggesting that there is little room for improvement in this area. However, when taking into account losses in the antenna matching, filtering, and switching circuits, and losses between battery and power amplifier, the overall efficiency quickly drops below 25%. Even worse, the best efficiencies are usually achieved for maximum power output levels. When power levels are decreased to more typical output levels, efficiencies often drop below 10%. In the future, we will see the emergence of more variable-envelop modulation schemes especially for the long range systems in order to use the available spectrum more efficiently. Also, CDMA (code division multiple access) access methods in addition to FDMA (frequency division multiple access) and TDMA (time division multiple access) will be used to allocate the available spectrum flexibly and efficiently. Variable-envelope modulation schemes and CDMA access methods will increase the requirements for linearity, power control range and accuracy of the power amplifier, resulting in overall efficiencies below 5% when the output power level is reduced by 20 dB or more [13]. Future improvements in the area of the transmitter are therefore to be expected in the area of improving efficiency at low power levels and high linearity, and in the antenna interface.

D. Power Dissipation of the Receiver Front End

The power dissipation of the receiver tends to be much lower than the power dissipation of the transmitter. However, since the receiver will in many cases be switched on for a much larger part of the time than the transmitter, e.g., to check whether new calls or messages are coming in, or whether a handover to a different base station is needed, the power dissipation of the receiver is nevertheless very relevant.

Receiver front ends tend to consume power far in excess of the bound set by conservation of power. The output of a typical front end might drive a sigma delta ADC with an input impedance that has a real part of, e.g., 100 kΩ and needs to be driven at signal levels at or below 1 Vpp. The maximum power delivered by the front end is then 0.5 V^2 / 100 kΩ = 1.25 μW. A well-optimized front end consumes around 50 mW, resulting in an efficiency below 0.01% in best case conditions. For lower input signals, resulting in lower output power, the efficiency further decreases by several orders of magnitude. Since the power efficiency of current receiver front ends is that low, there should be many opportunities for power reduction in this area.

1Although it might be that the technology needed to achieve this limit is physically impossible.
II. SIMPLE MODEL FOR RF POWER DISSIPATION

As power dissipation in current receiver front ends is not limited by conservation of power, there are other causes for the power dissipation of a receiver front end. These causes are not always the same, but depend on the type of radio connection that is being used. Before introducing models for power dissipation, trends in radio frequencies for portable wireless devices will be discussed. Afterward, their impact on the power dissipation in radio front ends will be assessed.

A. Trends in Radio Frequencies

In the past, the maximum radio frequency of portable wireless devices was mainly limited by the availability of (relatively) cheap technology in which the radio could be implemented. The minimum radio frequency was limited by the availability of unused spectrum.

Increased application of wireless devices has resulted in a gradual increase of the use of higher radio frequencies for wireless devices, which are currently concentrated in the 1–2 GHz region. Especially the explosive growth of the cellular phone, and the resulting move from bands below 1 GHz to the 2-GHz region, has been driving the development of cheap IC technologies with increased bandwidths.

In the near future this is likely to change: Cellular and other wireless systems that need to work over longer distances will be limited by the decreasing link budget at high frequencies when using omnidirectional antennas. In the radio transmission equation below [(1)], \( P_{RX} \) is the available signal power from the receive antenna, \( P_{TX} \) is the power applied to the transmit antenna, \( G_{RX} \) and \( G_{TX} \) are the gains of the receive and transmit antennas, \( \lambda \) is the wavelength, \( r \) is the distance between receive and transmit antenna, and \( \alpha \) is the propagation constant. This equation holds in the far field of the antennas assuming that the polarizations of both antennas are perfectly matched.

The propagation constant \( \alpha \) depends on the environment of the antennas. In free space, this constant is 2, whereas inside buildings it varies between 1.81 and 5.22 [14]:

\[
P_{RX} = P_{TX}G_{RX}G_{TX} \left( \frac{\lambda}{4\pi r} \right)^{\alpha}, \tag{1}
\]

From (1), it becomes obvious that the link budget for omnidirectional antennas that have constant gains scales with frequency \( f \) as

\[
\frac{P_{RX}}{P_{TX}} \sim \frac{1}{f^{\alpha}}. \tag{2}
\]

Increasing transmit power to make up for this reduction in the link budget becomes impractical in the frequency region above roughly 5 GHz. Another solution would be to use high gain antennas, but since such antennas will also be highly directional, they are not practical for portable wireless devices unless advanced beam steering approaches are applied and a line-of-sight path is available.

Since neither of these remedies is very practical, it seems likely that there will be a split in the frequency bands for short-range, high-bandwidth, high-bit-rate connections, and long-range (including outdoor) connections. Short-range applications will continue to move to higher frequency bands to satisfy the need for increased bandwidths, whereas long-range systems will be limited to frequencies below 5 GHz. These long-range systems will concentrate on efficient use of the increasingly scarce spectrum below 5 GHz through efficient modulation schemes, efficient access mechanisms, and adaptive use of the available resources. This split between short-distance high-bandwidth systems and long-range systems is shown in Fig. 3.

Front ends in long-range systems will require high linearity, low noise figures and high power efficiency for most efficient use and reuse of the spectrum. Since the frequencies in these systems will not increase significantly in the future, but the bandwidths of the devices in future IC technologies will continue to increase for many years to come, the power dissipation is typically limited by the dynamic range rather than the frequency.

Front ends in high-bit-rate, high-bandwidth, short-range systems will probably continue to move to higher frequencies. The high bandwidths for these systems will only be available at increasingly high frequencies in, e.g., the 5- and 17-GHz bands. The power dissipation in front ends for these systems will be limited by the high frequencies at which the front-end circuits will need to operate. Such systems will have relaxed requirements for sensitivity and linearity compared to long-range systems, and these requirements will typically be met at the current levels that are needed to achieve the high-frequency operation of the front-end circuits. Therefore, the power dissipation is typically limited by the operating frequency rather than the dynamic range.

This divergence of specifications can already be found in current systems such as GSM cellular system (for long range) and the Bluetooth system (for short range). Table 1 shows the noise figure and IP3 requirements defined through the standard.

Please note that competitive implementations typically need to perform significantly better than the performance set in a standard. Still a significant difference between both types of system remains.

Future portable wireless devices for systems such as UMTS/IMT 2000 will be able set up both short-range high-bit-rate links and long-range connections, depending on the environment and the needs of the user. In such multimode devices, the properties of the transceiver front...
end will depend on the particular mode in which that the device is being operated, and the optimizations discussed in this paper will apply to each of these modes.

B. Power Dissipation Limits for Short-Range Systems

Short-range systems are mostly bandwidth limited because the radio frequencies tend to be at the high end of feasible bandwidths in low-cost IC technologies. A simple model for the achievable bandwidth can be derived when making the following assumptions.

• The impedance of circuit inputs and outputs scales with current consumption at fixed voltage levels. This is reasonable to assume, since within reasonable limits, device sizes can be scaled with the current, and passive element impedance levels can be scaled with 1/current, resulting in fixed voltages throughout the circuit and little effect on noise behavior or voltage distortion.

• The main bandwidth limitation is a parasitic capacitance that is not tuned out and causes a first-order pole in combination with the impedance level of some signal internal to, or between, circuit blocks.

Under these assumptions, the bandwidth scales with current for fixed gain and noise figure. The linearity, expressed as third-order intercept power level (see Section II-C1), also scales with current. For high frequencies, the front end will be operating beyond the frequency of this dominant pole. That implies that there is a fixed gain-bandwidth product proportional to the power dissipation:

\[ P_{DC} = \gamma G f \]  
(3)

where \( G \) is the gain of a circuit, \( P_{DC} \) is the power dissipation, \( f \) is the signal frequency, and \( \gamma \) is the dissipation per gain-bandwidth product. The parameter \( \gamma \) depends on the circuit implementation. At this level of abstraction, there is no universal tradeoff between noise figure of a block and its power dissipation.

1) Power Sensitivity Tradeoff: A front end is usually a cascade of circuit blocks such as amplifiers, mixers, filters, etc. (Fig. 4).

A tradeoff between power dissipation and noise figure in such a cascade of circuit blocks can be achieved through trading higher noise figures of a stage by higher gain in the initial stages of the cascade. The higher gain will result in lower overall noise figure through Friis’ formula [15]:

\[ F_{tot} = 1 + \sum_{j=1}^{n} \frac{F_j - 1}{\prod_{j=1}^{n} G_j} \]  
(4)

It then follows that the lowest power dissipation that can be achieved for a cascade of two circuit blocks for short-range applications is

\[ P_{tot} = f \left( \gamma_1 \frac{F_2 - 1}{F_1 + \gamma_2 F_2} \right) \]  
(5)

Friis’ equation demonstrates that the closer the overall noise factor \( F_{tot} \) and the noise figure \( F_1 \) of the first circuit are, the lower the contribution of subsequent stages to the overall noise figure are allowed to be. This implies that a noise figure of the first stage close to the overall noise figure has to be compensated by a large gain, and therefore larger power dissipation, in the first stage, which is compensated by lower gain and power dissipation in subsequent stages. The total effect as predicted by (5) depends on the relative power efficiencies of the first and second stage. A graph based on a typical example is shown in Fig. 5.

In this example, we have assumed a typical cascade of a 2.5-GHz low noise amplifier (LNA) and a mixer. For the LNA, we have assumed a noise factor \( F_1 \) of 2. This corresponds to a noise figure \( NF_1 \) of approximately 3 dB, (using \( NF = 10 \log_{10}(F) \)). For the mixer, a noise factor \( F_2 \) of 30 was assumed (NF = 15 dB). For both the mixer and the LNA, a \( \gamma \) parameter with a value of 0.2 pW/Hz was assumed, which corresponds to about 5 mW for 10 dB of gain at 2.5 GHz. Total gain target \( G_{tot} \) was assumed to be 100, corresponding to 20 dB, and total noise factor target \( F_{tot} \) was swept from 2.1 to 10, corresponding to total noise figures from approximately 3–10 dB. As can be seen from this graph, the impact of the total noise figure is especially dramatic once the target is within 2 dB of the noise figure of the first stage. The slight increase of the power dissipation for high total noise figures is a consequence of the model used. The model finds a solution that matches the target gain and noise figure exactly. For high target noise figures, the gain of
the LNA has to drop close to unity, resulting in huge gains for the mixer. Since power dissipation of each block scales with gain, the total power dissipation scales with a weighted sum of the gains. On the other hand, the total gain is the product of the gains of the individual blocks. Therefore, it is more power efficient to realize the total gain with similar contributions of each of the subblocks, rather than with high contributions from one block and negligible contributions from others. In practice, this situation is usually avoided by allowing the noise figure to be lower than the target.

C. Power Dissipation Limits for Long-Range Systems

Power dissipation for long-range systems depends on both linearity and sensitivity. The linearity requirements for a front end depend on both the system in which the front end needs to operate and the implementation of the front end. In most cases, third-order distortion is the dominant source of signal impairment, because the large majority of portable wireless systems employs a frequency band that is just a few percent of the radio frequency. Therefore, second-order distortion components (which are the sum or difference of two signal frequencies) will end up far outside the frequency band of interest, and can be easily dealt with through filtering. Third-order distortion components of signals with frequencies \( f_1 \) and \( f_2 \) will show up at (among others) the frequencies \( 2f_1 - f_2 \) and \( 2f_2 - f_1 \), which can be well within the frequency band of a narrow-band system. Therefore, they cannot easily be suppressed through filtering. Higher order distortion tends to be weaker than the third-order distortion, causing the third-order distortion components to be dominant. The main exception to this rule occurs for direct conversion zero IF receivers. In such receivers, the input signal is multiplied by a local oscillator signal, resulting in an IF frequency of (approximately) 0 Hz. In such a receiver, the second order distortion of the switching stage of the mixers and subsequent circuits is relevant, since one second-order distortion component of any strong interferer will have a 0–Hz frequency and will therefore end up exactly on top of the wanted signal. Such receivers therefore require special attention to second-order linearity (e.g., through carefully balanced circuits) in addition to the usual third-order linearity targets.

1) Third-Order Distortion: Third-order distortion for front ends is often quantified through IP3, the third-order intercept point. This is the input power level at which the extrapolated curve of output power vs. input power intersects with the extrapolated curve of output third order distortion vs. input power. This is demonstrated in the graph in Fig. 6.

The third-order intercept point can therefore be interpreted as the input power level at which the third-order distortion components would have been as large as the wanted signal at the output, assuming that the wanted signal would continue to increase linearly with the input signal, and the distortion components would continue to increase proportionally to the third power of the input signal. In a similar way, the second-order intercept point (IP2) can be defined. Note that the definition of IP3 is based on extrapolation from the linear region of the output power vs. input power curves. This implies that the signal levels are small compared to the level at which gain compression occurs. Since most systems are defined through one or more interferers that a receiver should be able to cope with, the IP3 specification that is derived from this should also include a minimum interferer level at which this IP3 should still be achieved. Often, this is not explicitly stated, and 1 dB compression is assumed to be about 10 dB below the IP3. However, this does not always need to be the case, especially if linearization through distortion cancellation is applied in a circuit.

2) Power Linearity Tradeoff: The tradeoff between gain, noise, linearity, and power dissipation for long-range systems depends on IC technology, circuit topology, desired specifications, power supply voltage, and probably many more parameters. In order to deal with this complexity at a more abstract level, it can be shown [16] that it is possible for any circuit to trade linearity, gain and power dissipation through structure independent transforms, independent of circuit topology, desired specifications, power supply voltage etc. Assuming unilateral gains and matching, the power dissipation can be approximated through the following equation:

\[
P_{DC} = \kappa G_p \text{IP3}
\]

where \( P_{DC} \) is the power dissipation of a circuit, \( G_p \) is the power gain of this circuit, and IP3 is the third order input intercept point. The power linearity parameter \( \kappa \) depends on circuit topology, IC technology, etc. This parameter allows us to rate different circuit topologies in the same technology on their power efficiency in achieving linearity and gain at lower power dissipation. From Fig. 6, it follows that \( G_p \text{IP3} \) equals OIP3, the output third-order intercept point. Therefore, we can also rewrite (6) as

\[
P_{DC} = \kappa \text{OIP3}.
\]

3) Power Sensitivity Tradeoff: Sensitivity depends on the noise figure of the front end, and is mainly dependent on the noise figure of the active devices and losses incurred in passive components in matching networks. Assuming that some gain is required, there is no way to improve the noise

\[
\text{Authorized licensed use limited to: Eindhoven University of Technology. Downloaded on August 19, 2009 at 09:51 from IEEE Xplore. Restrictions apply.}
\]
figure of a circuit beyond the noise figure of the active device, no matter how much current is available. Also, the structure-independent transforms don’t allow for a direct tradeoff between the power dissipation and noise figure of a single block. As is the case for short-range systems, the noise figure can be traded off with power dissipation in a cascade of circuit blocks. Increasing the gain of a stage will reduce the noise contribution of subsequent stages to the total noise figure in accordance with Friis’ equation (4).

However, there is an important difference with short-range systems: the larger gain in the initial stages will also require larger linearity of later stages, since they have to deal with larger signals. This larger linearity will result in higher power dissipation of the later stages, and therefore there will be an indirect tradeoff between gain and noise figure as well. This tradeoff does depend on the target for the total front end noise figure, and the noise figure of the initial stage.

It can be shown that the minimum power dissipation that can be achieved for two cascaded stages takes the following form:

$$P_{tot} = P_{3dB} \left( \kappa_1 \frac{F_2 - 1}{F_1} + 2 \sqrt{\kappa_1 \kappa_2} \frac{\sqrt{G_{tot}(F_{tot} - F_1)(F_2 - 1)}}{F_{tot} - F_1} + \kappa_2 \frac{G_{tot}(F_{tot} - F_1)}{F_{tot} - F_1} \right) .$$

(8)

As was the case for short-range systems, the equation indicates a fast increase in power dissipation that results from an overall noise factor $F_{tot}$ close to the noise factor $F_1$ of the first circuit. This time, the effect is more pronounced since the larger gain in the first stage results in larger signals in subsequent stages. These larger signals result in higher linearity requirements for subsequent stages, and therefore higher power dissipation. Fig. 7 shows the impact of the total noise figure specification on the power dissipation for such a simple two-stage front end.

The same parameters for the circuit blocks have been used as for the short-range example, with the additional constraint that the total IP3 should be 0.1 mW (or –10 dBm), and that $\kappa_1 = 5, \kappa_2 = 15$. From Fig. 7 it becomes apparent that the impact of the noise figure on the total power dissipation is most dramatic when the specification for the total noise figure approaches the noise figure of the initial stage, especially when the difference becomes less than about 2 dB. The reason for this behavior is that the higher gain of the LNA causes higher IP3 requirements in the mixer, resulting in even higher overall power dissipation. In contrast to short-range systems, the noise figure of the first stage is never “good enough,” because at higher total noise figures, a better tradeoff between gain and linearity can be achieved.

D. Limits to Scaling

Structure independent transforms are limited by the scaling available in IC technologies. The relations described in this section break down when the linearity requirements become so low that the transforms point to device sizes that are either not available in an IC technology, or cannot be achieved with (approximately) scaled electrical behavior. This typically occurs when scaling a transistor down to minimum size, when parasitic capacitances of the device itself and the surrounding interconnect don’t scale proportionally with the current through the transistor. This is the main cause for front ends in which the power dissipation is limited by the signal frequencies rather than the dynamic range.

E. Other Circuits

The models introduced thus far apply only to circuits that process the antenna signal directly, providing, e.g., gain, filtering, or frequency translation. There are other circuits that do not directly process the antenna signal, e.g., for the generation of local oscillator (LO) signals that are applied to mixers in which the antenna signal is shifted in frequency. For the purpose of this paper, these circuits are considered to be part of the mixer and scale in principle with the scaling of the mixer. However, they do pose different demands on the IC technology than signal processing circuits. For example, linearity is typically less relevant for LO signal generation than for antenna signal processing. In terms of technology requirements, these circuits will resemble much more the gain and bandwidth limited circuits in short-range systems, even when they are part of long-range systems. Especially for the voltage controlled oscillator (VCO) that is usually part of the LO signal generation circuits, the quality of the passive components in the resonator circuit is of crucial importance to the power dissipation of this circuit [17], [18]. Leeson’s equation [19] predicts that the power dissipation will scale proportionally to the inverse of the quality factor of the tank circuit squared. Therefore, the quality of passive components is especially relevant for the power dissipation of LO generation circuits in both long- and short-range systems.
III. MINIMIZING RF FRONT-END POWER CONSUMPTION

The power dissipation of an RF front end depends on many different parameters, including:
- antenna interface;
- circuit design;
- IC technology levels.

Achieving minimum power consumption for a front end requires systematic optimization at all three levels. Each of these levels will be discussed in a separate part of this section.

A. Antenna Interface

The antenna interface is a much neglected area that significantly affects the power dissipation of a front end. Any losses in the interface to the antenna have to be compensated by an equal decrease in noise figure of the front end in order to maintain the same overall noise figure (and therefore sensitivity), as indicated by Friis’ formula \((\text{4})\). In addition, the gain has to increase by the same amount as the loss to keep overall signal levels the same. On the other hand, the input IP3 of long-range systems can decrease proportionally to the losses because of the reduced input signal levels. Typical losses in the antenna interface include losses in the antenna filter, duplexer and/or switches, and can add up to about 2 dB. Using again the same simple LNA/mixer cascade example from the previous section, we can estimate the impact of such losses on the power consumption of a front end for both short- and long-range systems (Fig. 8).

In this example, we have assumed that the target for the overall noise figure is 7.8 dB. The relative increase in power dissipation with antenna interface losses is in this case lower for long-range systems than for short-range systems, although of course the absolute power dissipation is higher. The explanation for this behavior is that the extra power required for improved sensitivity is partially compensated by the reduction in power due to lower linearity requirements for the first stage. This does not occur as a general rule, but depends on the specific parameters of the system and the circuits.

In the same way, the power dissipation of a system with difficult sensitivity specifications compared to the noise figures achievable in the IC technology can be significantly reduced by using antenna diversity schemes. The received antenna power predicted by the radio propagation equation \([\text{(1)}]\) is an estimate of the average power that will be received at some distance between transmit and receive antenna. The received antenna power will vary significantly around this average over small distances due to multipath interference. Typical indoor values in the 2.5-GHz range show variations in the order of 10 dB over distances of a few centimeters. A graph of a typical indoor signal strength distribution is shown in Fig. 9.

Because of the digital coding of the information in modern portable wireless devices, it is not the average received power that is relevant, but whether the power level is high enough to allow perfect reconstruction of the information after error correction. This is possible if the signal level exceeds a specified threshold. The fraction of the area inside (or outside) a building in which the received antenna power exceeds this threshold is called coverage \([\text{20}]\). For example, a coverage of 99% indicates that in 1% of the area the signal is too weak for reliable information transfer, but in the remaining 99% of the area perfect information transfer can be achieved. For portable wireless devices that transmit and receive digital information, coverage is a more relevant performance parameter than average received antenna power. It is possible to significantly reduce the effect of multipath by antenna diversity schemes, in which the signals of multiple antennas are combined. These antennas can, e.g., be at different locations, or they can have different polarization or different antenna patterns, etc. The signal from one of these antennas can be selected as input for the receiver, depending on the signal quality on each of the antennas, or the signals from several antennas can be combined in ways to further improve the total signal quality.

Even relatively simple schemes such as equal gain combining of two antennas can result in an improvement of about 10 dB when targeting equal coverages in the 98% range indoors \([\text{21}]\) (Fig. 10).
In equal gain combining antenna diversity, the signals from multiple antennas are added with the same gain, but with an optimized phase difference. This concept is shown in Fig. 11. In angle scanning diversity, the same approach is taken, but the optimized phase difference is determined periodically by scanning all possible settings of the phase difference. For speeds up to 5 km/h and update frequencies of 1 kHz, angle scanning performs as well as ideal equal gain combining.

One way to interpret the function of an equal gain combiner is to model the output signal from the adder as if it was generated by a single antenna with a complex antenna pattern. In that case, the effective antenna pattern of this virtual antenna depends on the value of the variable phase shifter. Fig. 12 demonstrates some of the antenna patterns that can be generated this way.

When distributing 10 dB of diversity link budget improvement over 5 dB reduced transmit power and 5 dB reduced sensitivity of the receiver, approximately three times power dissipation reduction can be achieved both in transmit and receive mode. Fig. 10 also shows that the improvement depends on the target coverage. For higher coverage targets, the improvement is substantially more than 10 dB. In Fig. 13, the relative power dissipation of the receiver is shown for both long- and short-range applications assuming that half of the improvement in link budget is allocated to reducing transmit power, and the other half to reducing receiver sensitivity. Complementary to the situation with antenna losses, the increased signal levels provided by antenna diversity can be used to optimize the front end for a proportionally increased noise figure and proportionally decreased gain to maintain overall sensitivity and signal levels. In addition, the input IP3 of long-range systems has to increase proportionally to the signal levels to maintain overall linearity. The result is shown in Fig. 13. In this figure, the relative power dissipation is the power dissipation of a receiver with equal gain antenna diversity relative to a receiver without antenna diversity. A short-range system benefits more from antenna diversity since there is no penalty (in terms of linearity) in dealing with the larger combined signals, whereas in a long-range system extra linearity is required for dealing with larger signal levels. The situation for long range systems can be improved by using a diversity control algorithm that reduces interfering signals. This can be achieved by optimizing for a tradeoff between received antenna power of the wanted signal and suppression of unwanted signals. The impact of such a scheme on power dissipation is more complex and cannot be dealt with using the simple models in this paper.
In this simple model, we have not taken into account the power dissipation of the overhead circuitry needed to implement antenna diversity. In principle, the diversity combination can be carried out by passive components between the antennas and the receiver, without any additional signal processing circuits. Also, with the predicted power dissipation savings of 2.5 to 30 times for this example, there would still be appreciable power savings in the extreme case of doubling the complete receiver circuit.

B. Circuit Level

In short-range, high-bandwidth systems, the main problem will be in designing circuits with sufficient gain at high frequencies and low power dissipation. This can be achieved by reducing currents in the circuits (to reduce the power) while decreasing transistor sizes (to keep the available bandwidth constant) and increasing impedance levels proportionally to the current (to keep the gain constant). This is possible since the current required to achieve sufficient gain at these high frequencies usually results in higher linearrings than needed.

This approach is typically limited by technology that does not allow further scaling of the transistors and impedances with proportionally scaled parasitics. In that case, parasitics of transistors, passive devices, and/or interconnect start to dominate the achievable gain at the desired signal frequency. In the next section, we will investigate technology measures that can improve this situation. When this approach is successful, currents can be scaled down to the point where the linearity is just good enough, and further power savings will then require approaches similar to those for long-range systems as discussed in the second part of this section. The result of this approach is a circuit with high internal impedance levels.

Circuits with high internal impedance levels are often difficult to interface to external signals and components, since the impedance levels on a PCB tend to be limited to a few hundred ohms for practical reasons. This makes interfacing with the outside world awkward and is an additional impetus for highly integrated transceivers with a minimum number of interfaces to the outside world. At these interfaces, the impedance of the signals has to be adjusted to external impedance levels either electronically or by passive components. Electronic matching will result in increased power dissipation as well as additional noise and distortion. Matching through passive components that are on-chip is costly because of the required chip area, and if the on-chip passive components are of limited quality (as, e.g., on-chip inductors often are), signal losses are incurred. Matching through passive off-chip components is most common but results in a higher overall component count. A better solution could be achieved with a low-cost IC technology that allows integration of high-quality passive components.

Low power circuits for long-range systems need to meet three main design goals:

- low noise figures at high frequencies;
- good power/linearity (κ) values at high frequencies;
- efficient output power generation.

The noise figure is mostly limited by the properties of the active device and the availability of high quality passive components from which to build the matching networks.

Good power/linearity (κ) values require high gain and/or good linearity at high frequencies and low currents. One way of achieving this is through distortion cancellation, which can be thought of as another structure-independent transform. It is based on the principle that the ratio between distortion component and signal component is generally dependent on the signal level. By processing (at least two) different levels of the same signal through identical circuits, it is possible to obtain two distorted signals with different signal-to-distortion ratios. By adding these two distorted signals with appropriate gain factors, it is possible to cancel a distortion component while preserving (part of) the signal.

In principle, a perfect cancellation can be obtained for a single distortion component. In practice, the reduction is limited by the matching that can be achieved between the branches. By recursively applying this procedure, in principle all distortion components can be eliminated. Obviously, this is impractical for more than two or three components because of the cost and power dissipation of the circuits involved.

In order to achieve good linearity, many RF front-end circuits operate in class A. The efficiency of the generated output power for class A circuits is rather poor. Even at maximum output levels, it does not exceed 25% for a resistive load. The lowest signal levels in a front end are often five to eight orders of magnitude lower than the maximum signals, resulting in efficiencies as low as 25×10−10 for such signals. Achieving better power efficiency for low signal levels requires circuit topologies with at least class A/B outputs or class A circuits with adaptive biasing dependent on the signal levels [22].

C. IC Technology

In this paper, we will concentrate on bipolar IC technologies since they are still the most widely used technologies for portable wireless devices. When optimizing an IC process for RF low power circuits [23], [24], the following points need to be considered:

- achieving high gain-bandwidth products at low currents requires both small emitter areas and parasitics that are scaled proportionally to the current;
- building low power RF circuits with these devices requires interconnect with parasitic capacitances that are scaled proportionally to the current, and high-Q passive components.

1) IC Technologies with High Gain-Bandwidth Products at Low Currents: IC technologies for low-power short-range systems will need to provide much gain at high frequencies while consuming low currents. Long-range systems need a high product of linearity and gain at low currents. This can be achieved by improving linearity at the same gain and current, or improving the gain at the same linearity and current. Since improving the gain at low currents is beneficial for both long- and short-range systems,
this improvement will make an IC process attractive for more applications.

The gain of a transistor at high frequencies is typically expressed as some type of bandwidth vs. current parameter. Such a parameter can be determined in many ways. A common figure of merit for bandwidth is $f_T$. This parameter indicates the frequency at which the current gain of a common emitter stage, driven by a current source at the base, and shorted at the collector, has dropped to one. For a given device, this frequency is a function of the current through the device as shown in Fig. 14.

In typical RF circuits, the transistor is used in a different environment than the conditions in which $f_T$ is determined. The input is typically driven from a finite impedance, and the output is not shorted, making $f_T$ a less relevant predictor of RF transistor performance. For that reason, common emitter input ($f_0$) and output ($f_{out}$) bandwidths are often used to characterize the bandwidth of a transistor [25]. Input bandwidth is the 3-dB transconductance bandwidth of a transistor driven from a voltage source and shorted at the collector. The output bandwidth is the 3-dB voltage gain bandwidth of a common emitter stage with a voltage gain of ten. The total bandwidth in common emitter configuration is mostly limited by these two bandwidths, and is called available bandwidth ($f_a$) [26]. What we are then looking for is a method for achieving high $f_a$ values at low current. At low currents, $f_a$ tends to be limited by $f_{out}$. To improve $f_{out}$, it is necessary to decrease the parasitics at the transistor output: collector (or drain) to substrate capacitance and collector-base (or drain-gate) capacitance.

In addition to a low-power transistor, the parasitics of the interconnect also have to scale with the current reduction, in order to allow for the higher impedance levels in the circuits. There are two ways in which the capacitance of interconnect can be reduced by:

1) reducing the dielectric constant of the material between the interconnect and the ground;
2) increasing the distance of the interconnect to the ground relative to the line width of the interconnect.

The first approach is limited since the relative dielectric constant of SiO$_2$, the most common intermetal dielectric, is around 4. A decrease to values around 2 is possible using low-$K$ dielectric materials, but further improvements are highly unlikely. The second approach is limited because the capacitance of interconnect does not scale proportionally any more if the width of the interconnect becomes much smaller than its height over the substrate. This is demonstrated in Fig. 15.

The increase in oxide thickness required to achieve an interconnect capacitance reduction by an order of magnitude is very impractical using standard IC processing techniques. Therefore, a different approach is needed. An IC process that solves this interconnect problem as well as many other low-power related problems will be discussed in the next section.

IV. SILICON-ON-ANYTHING TECHNOLOGY

Both the scaled transistor and scaled interconnect parasitics have been addressed simultaneously in the SOA process [27]. This is a bipolar IC process in which the active layer of an (upside-down) SOI wafer is glued onto a substrate of choice (the “anything”) after processing, as shown in Fig. 16.

The substrate of the SOI wafer is now facing upwards. This silicon substrate is removed completely, using the buried oxide as an etch stop. For this design, glass was selected for the substrate because it is cheap and has low losses over a wide frequency range.

From a design point of view, this process offers five important advantages compared to a conventional bipolar silicon process:

- a lateral NPN transistor with 0.1 $\mu$m$^2$ emitter area using 0.5-$\mu$m lithography. A common emitter stage provides 2.4 GHz of bandwidth at 20 dB of gain with only 10 $\mu$A current;
- interconnect (including bondpads) with 5 to 20 times lower parasitic capacitance to ground (depending on the line width);
- almost perfect isolation between circuit blocks;
• integrated inductors with $Q$ values up to 40;
• an individual tradeoff between $f_T$, base resistance, early voltage, and breakdown voltages for every individual transistor in the design by changing the collector drift region (parameter $L_{ceh}$ in Table 1).

The process also provides 15 kΩ/poly resistors, 1.5 nF/mm² capacitors, PIN diodes, varicaps, PNP s, and JFETS, and 2.5 metal layers, all in 14 mask steps, which together with the 0.5 μm lithography result in low fabrication costs. This allows cost-efficient integration of high-performance input and output matching circuits. Table 2 summarizes some transistor parameters for two different values of $L_{ceh}$, and Fig. 17 shows the layout and cross section of an SOA transistor connected to a polysilicon resistor.

To demonstrate that the goal of increasing gain at low power levels has been achieved in SOA, Fig. 18 shows the $f_T$ of both a minimum size SOA transistor and a standard 1 μm BiCMOS transistor.

The high bandwidth at low currents is a direct result of the scaling of the emitter area and proportional scaling of the transistor parasitics.

V. FITTING IT ALL TOGETHER: A LOW-POWER FRONT END

Architectures have recently been targeting integration level and cost reduction. This has resulted in zero-IF and low-IF architectures with integrated active filters for channel selectivity. Although such filters require more power than passive filters with similar specifications, this is partially compensated by not having to go off-chip for the filters. This is especially important for low-power short-range systems that use high internal impedance levels in the RF circuit.

Although multiple IF branches in zero IF and low IF receivers might seem to increase power dissipation, this is not the case for long-range receivers, at least according to the simple power dissipation model used throughout this paper. The signal levels in each of the branches of a dual branch receiver is only half of the total signal power, and therefore linearity requirements in each of the branches is 3 dB lower than the linearity requirements for an identical receiver with a single IF branch. Since in a long range receiver the power dissipation scales with linearity, the total power dissipation of both receivers will be the same. Nevertheless, integrated high-quality passive filters might offer the best long-term solution for performance, integration level, and power dissipation [29], but the required high-quality passive resonators are not yet commonly available in low-cost IC processes.

A. Designing a Low-Power Short-Range Receiver

To verify the approach to low power presented in this paper, an equal gain antenna diversity receiver for 2.5-GHz wireless data transfer has been designed, built, and evaluated in the SOA process. This is an example of a short-range link with emphasis on high gain at high frequencies and low power consumption. The receiver has been designed to minimize the need to go off-chip by using a zero-IF architecture. The front end includes an on-chip antenna matching network and filter to interface the high on-chip impedance levels to the external antenna impedance.

The equal gain diversity combiner can be implemented in several ways, trading flexibility for low cost and low power dissipation.

- Using passive components between the antenna and the LNA to achieve a continuous phase shift will probably result in the lowest power dissipation and highest flexibility. However, the sensitivity will be degraded because of the extra components in the antenna signal path.
- If only two antenna patterns are required, a simple implementation is possible using separate LNAs for each antenna with balanced outputs that can be added with zero degrees or 180° phase shift by crossing the wires of the balanced output signal of one LNA. The overhead in circuit area and power dissipation is only one additional LNA.
If multiple, but not more than, e.g., eight antenna patterns are required, the phase shift can easily be implemented as resistive interpolation between the quadrature outputs of the mixers in a quadrature (e.g., zero IF) receiver. The appropriate phase shift can then be selected through low-frequency switches. This might result in a very interesting tradeoff between power, cost, performance, and flexibility.

In this design, the equal gain diversity combiner has been implemented for maximum flexibility, at the cost of additional circuits and power dissipation. It is positioned after down conversion to allow easy implementation at low frequencies. The quadrature down converter already generates IF signals with 90° phase difference. Using weighted summation of the IF signals of two antennas, any phase difference can be implemented up to the accuracy of the amplitude weighing networks (Fig. 19).

The receiver operates in the 2.5-GHz ISM band, with a spacing between the antennas of 1/2 wavelength (6 cm). Two 8-bit digital-to-analog converters (DACs) drive the weighing amplifiers to enable very fine steps in the phase setting of the diversity receiver. Also, dual phase shifters and dual IF circuits have been implemented to be able to simultaneously monitor the reception at two different settings of the phase shifter. These features are intended for further study of diversity schemes. As we will see, even this very flexible implementation results in very low power dissipation when implemented in SOA using high impedance circuits.

### B. Implementation of the Low-Power Short-Range Receiver

The RF input of the LNA (Fig. 20) uses a combination of LC circuits and electronic buffers to achieve the required impedance transform. It is implemented with an integrated LC type transformer that feeds into a common base stage ($Q_1$ and $Q_2$, which are $2 \times 7$ transistors running at $24 \mu A$ each) with 300-Ω input impedance. The output of this stage is connected to two differential pairs ($Q_3/Q_4$ and $Q_5/Q_6$) which provide variable gain operation. These feed into the RF inputs of the mixers. The LNA provides 20 dB of gain at 336 $\mu A$ with a 4-dB noise figure and $-20$ dBm IP3.

![Diagram of the low-noise amplifier at the input of the low-power front end.](image)

**Fig. 20.**

The LO inputs are implemented with electronic buffers. In a single chip version of this receiver, the LO will obviously be an internal connection, which does not need these buffers. The buffers provide 8.5 dB of gain. The mixers themselves provide 17 dB of gain, whereas the phase shifters have a maximum gain of 0 dB.

The performance of this LNA is compared to the performance of an earlier LNA in a standard 1-$\mu$m BiCMOS process [30] in Table 3.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>SOA</th>
<th>BICMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>2.5GHz</td>
<td>1.9GHz</td>
</tr>
<tr>
<td>Gain</td>
<td>20dB</td>
<td>20dB</td>
</tr>
<tr>
<td>NF</td>
<td>4dB</td>
<td>3.5dB</td>
</tr>
<tr>
<td>IP3</td>
<td>-20dBm</td>
<td>-20dBm</td>
</tr>
<tr>
<td>Pdc</td>
<td>1mW</td>
<td>10mW</td>
</tr>
</tbody>
</table>

The LO inputs are implemented with electronic buffers. In a single chip version of this receiver, the LO will obviously be an internal connection, which does not need these buffers. The buffers provide 8.5 dB of gain. The mixers themselves provide 17 dB of gain, whereas the phase shifters have a maximum gain of 0 dB.

The performance of this LNA is compared to the performance of an earlier LNA in a standard 1-$\mu$m BiCMOS process [30] in Table 3.

This table demonstrates that the high impedance circuit in combination with the SOA process already results in about one order of magnitude power savings with similar specifications. What is not immediately obvious from this comparison is that the SOA front end actually performs with a 10-dB higher apparent sensitivity than the BiCMOS front end because of the integrated equal gain antenna diversity. If this would be taken into account, the power savings would be even higher as predicted in Fig. 13. The die with the complete front end is shown in Fig. 21. The inductors and capacitors of the input transformer are clearly visible slightly above the center of the die.
Total die area of the front end IC is 9 mm². The measured performance shows an overall gain of 35 dB, with 6.2-dB noise figure and -22 dBm IP3. The IP3 measurement result is shown in Fig. 22.

The supply current is 1.0 mA at 3.0 V, with about half that current going into the LNA and mixers, and the other half into the DACs and phase shifters, that can be powered down independently. Biasing is provided by on-chip bandgap reference sources. Compared with a similar front end without diversity and implemented with conventional circuits in a standard 1 μm BiCMOS process, the power dissipation improvement is even more impressive than for the LNA by itself (Table 4).

VI. CONCLUSION

To achieve very low power dissipation in RF front ends, the front end antenna interface, circuits, and IC technology all need to be optimized. Depending on the type of application (long-range radio connection or short range), different optimizations are required. The antenna interface offers much potential for reducing power dissipation, e.g., through antenna diversity. At the circuit level, high impedance levels, linearization, and class A/B circuits offer possibilities for power dissipation reduction depending on the type of system. Finally, the IC process has a big impact, especially the scaling of the device size with proportionally scaled parasitics and interconnect. This requires dedicated low-power process optimization such as has been achieved in silicon-on-anything.

Combining all of these ideas has been shown to result in 10-dB improved sensitivity in combination with a power dissipation reduction by one order of magnitude.

ACKNOWLEDGMENT

The authors would like to thank H. Maas, L. Leyten, T. Wagemans, G. Dolmans, G. Janssen, J. van Sinderen, A. Tombeur, H. Visser, L. de Maaijer, and other colleagues in Philips Research for their contributions to this paper.

REFERENCES


Table 4

<table>
<thead>
<tr>
<th></th>
<th>SOA</th>
<th>BiCMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frequency</strong></td>
<td>2.5GHz</td>
<td>1.9GHz</td>
</tr>
<tr>
<td><strong>Gain</strong></td>
<td>35dB</td>
<td>37dB</td>
</tr>
<tr>
<td><strong>NF</strong></td>
<td>6.2dB</td>
<td>6dB</td>
</tr>
<tr>
<td><strong>IP3</strong></td>
<td>-22dBm</td>
<td>-25dBm</td>
</tr>
<tr>
<td><strong>Pdc</strong></td>
<td>3.5mW</td>
<td>45mW</td>
</tr>
</tbody>
</table>

Peter G. M. Baltus (Member, IEEE) was born in Sittard, The Netherlands, on July 5, 1960. He received the M.S. degree in electrical engineering from the University of Technology, Eindhoven, The Netherlands, in 1985. For five years, he worked at Philips Research Laboratories, Sunnyvale, CA, in the areas of data converters, micro controllers, and computer-aided design. In 1990, he joined Philips Research Laboratories, Eindhoven, where he worked on telecom RF transceiver front ends, concentrating on low power system and circuit design, and the relation between power dissipation and IC process and device parameters. He is currently with Philips Semiconductors, Sagamihara, Japan.

Ronald Dekker was born in Rotterdam, The Netherlands, in 1961. He received the M.Sc degree in electrical engineering from the University of Eindhoven, Eindhoven, The Netherlands. He joined Philips Research Laboratories, Eindhoven, in 1988, where he works on IC process development for both low-power and high-power bipolar RF applications.