CHAPTER 2   LITERATURE REVIEW

2.1   INTRODUCTION

Chapter 1 dealt with the introduction to the research performed in this thesis. In this chapter, a literature review is presented to place this into context of available knowledge. A few key concepts seminal to this study are identified and presented. Firstly, basic concepts of power amplification are listed, including some examples of good PA designs from the second part of the first decade of the twenty-first century. Secondly, a very important component of any PA, the inductor, is explored in depth, because a large number of different implementations and models available in literature make identifying a suitable inductor a time-consuming task. Thirdly, available fabrication technologies are investigated because understanding of the technology available for testing forms the foundation of this study. The fourth section of this review deals with the gap in knowledge when it comes to automating the task of rapid PA and passive inductor design. Finally, the conclusion of the chapter is drawn.

2.2   POWER AMPLIFICATION

PAs are used on the transmit side of RF circuit, typically to drive antennas with high power [8]. In such a scenario, the scaled-up versions of small-signal amplifiers used in the low-power low-frequency CMOS circuits are fundamentally incapable of high frequency, and other approaches must be considered [30]. In the monolithic PA implementations, the trade-off between efficiency and linearity is very significant [8]. In addition to these two determining factors for the PA integration, some other challenges, such as the limited supply voltage and heat generation, are also present. These factors often lead to the discrete PA designs or integration separate from the rest of the front end of the RF circuit.

Figure 2.1 shows a general single-ended PA model [31]. In this model, V_{DD} is the voltage supply, R_L is the load and RFC is the RF choke, and inductor is large enough to ensure the substantially constant current through the drain. In some designs, this inductor can be replaced by a finite one, if it forms a part of the output filter, also shown in this figure [32]. The output filter can include harmonic tuning and wave shaping, impedance matching or any other passive circuitry. The transistor T_1 is shown as an n-channel MOS (NMOS)

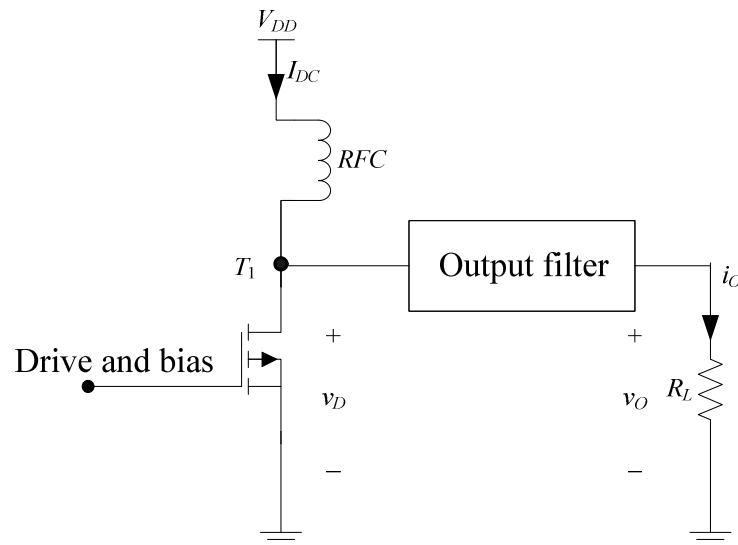transistor, but it can be any power transistor (HBT, MOS, BJT or other) used in a particular PA application.

Figure 2.1. General model of PA [31].

2.2.1   Power capability

The task of a PA is to deliver a given power into the load [1]. This power is determined by the power supply voltage and the load. Maximum power that can be delivered is

$$P = \frac{V_{DD}^{2}}{2R_{L}} \tag{2.1}$$

In case of Class-E and Class-F PAs, somewhat larger power can be delivered to the load. This will be detailed in Chapter 4.

2.2.2   Power consumption

Total DC power consumption is an important quantity in PA design, especially for battery powered portable devices. DC input power of a PA is the current drawn from the voltage supply over a period of time, or

$$P_{dc} = \frac{1}{T}\int_{0}^{T} V_{DD} i_{D} dt = \frac{V_{DD}}{T}\int_{0}^{T} i_{D} dt = V_{DD} I_{DC}, \tag{2.2}$$

where i_{D} is the drain current and I_{DC} is the DC component of the current waveform.

2.2.3   Power efficiency

Efficiency is a measure of performance of a PA. The performance of a PA will be better if efficiency is higher, irrespective of its definition. There are several definitions of efficiency commonly used with the PAs [8, 30, 31]. Drain (or collector) efficiency η is defined as the ratio of RF output power (P_{out}) to DC input power (P_{dc}), or

$$\eta = \frac{P_{out}}{P_{dc}} \tag{2.3}$$

Assuming sinusoidal voltage and current, the RF output power is given by

$$P_{out} = v_{eff} i_{eff} = \frac{i_1 v_1}{2} = \frac{i_1^2 R_L}{2}, \tag{2.4}$$

where i_{eff} and v_{eff} are effective and i_1 and v_1 are the peak fundamental components of current and voltage respectively, and the DC input power is given by Equation (2.2).

Power added efficiency (PAE) takes into account the input power (P_{in}) by subtracting it from the output power:

$$PAE = \frac{P_{out} - P_{in}}{P_{dc}} = \frac{P_{out} - \dfrac{P_{out}}{G}}{P_{dc}} = \eta\left(1 - \frac{1}{G}\right), \tag{2.5}$$

where $G = P_{out}/P_{in}$ is the gain of the PA. The PAE will give a good indication of the performance of a PA for high amplifier gains but it can even become negative for low gains. This relationship is shown in Figure 2.2.

The third type of efficiency is overall efficiency, the ratio of output power to the sum of input power and DC input power:

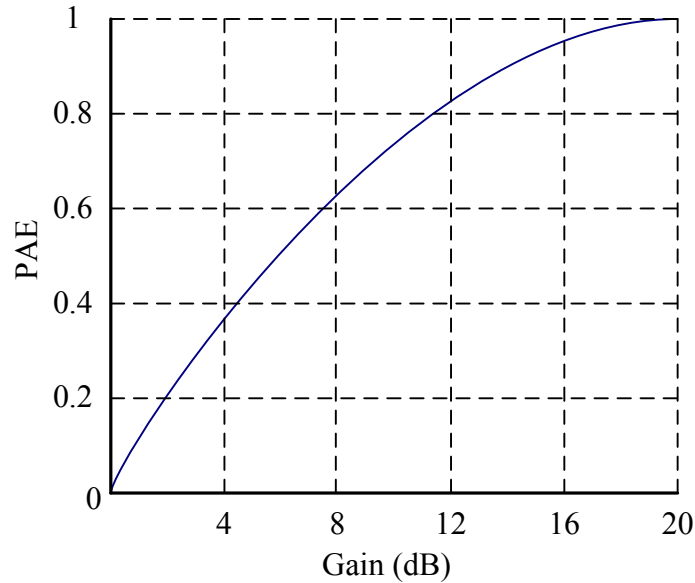$$OE = \frac{P_{out}}{P_{dc} + P_{in}} \tag{2.6}$$

Figure 2.2. Normalized PAE versus the PA gain [1].

The above efficiencies are instantaneous quantities. The instantaneous efficiency is highest at the peak output power. Instead of instantaneous efficiencies, average efficiency can be used for the signals with time varying amplitudes:

$$\eta_{AVG} = \frac{\overline{P}_{out}}{\overline{P}_{dc}},$$
(2.7)

where $\overline{P}_{out}$ is average output power and $\overline{P}_{dc}$ is average DC input power.

2.2.4   Matching for desired power

The maximum power transfer theorem, commonly used in RF electronics for optimal power transfer by means of conjugate matching, can only be used at the input side of a PA [1, 30], but it does not apply at the output side. At the input side, the PA has to be designed so that correct current and voltage waveforms are delivered at the gate or base of the transistor (biasing). At the output side load has to be correctly chosen. From Equation (2.1), it is obvious that the only two parameters influencing the output power are the voltage supply, V_{DD} or V_{CC} and the load impedance, R_L. The supply is normally fixed for a given application, so that the only degree of freedom left to the designer is the impedance of the load. Looking for correct impedance for maximum power output is called load pull. This impedance will often differ from standard impedances of 50 or 75 Ω.

Impedance matching networks are used to convert standard impedances to required impedances. At microwave frequencies, where wavelengths are correspondingly small, this matching can be accomplished with microstrip lines [33]. At low gigahertz frequencies, the microstrip lines are impractically long to be used on a chip, so that impedance matching using discrete components is employed. The two-component networks (L networks) and three-component networks (T and Π networks) are commonly used. Eight L-network configurations are possible, as shown in Figure 2.3 (a) and (b), where X_1 and X_2 can be any combination of inductors and capacitors, Z_S is the source impedance and Z_L is the load impedance. Such an L network is a broadband (high-pass or low-pass) network. Conversely, the T and Π networks with passives X_1, X_2 and X_3, shown in Figure 2.4 (a) and (b), are narrowband networks.

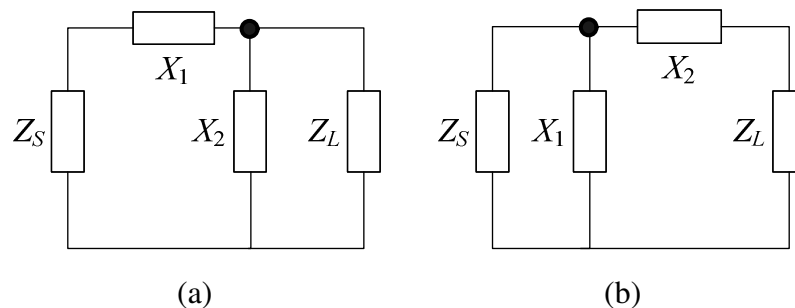(a)   (b)

Figure 2.3. Two-component matching networks where passive component is parallel to (a) load and (b) source [33].
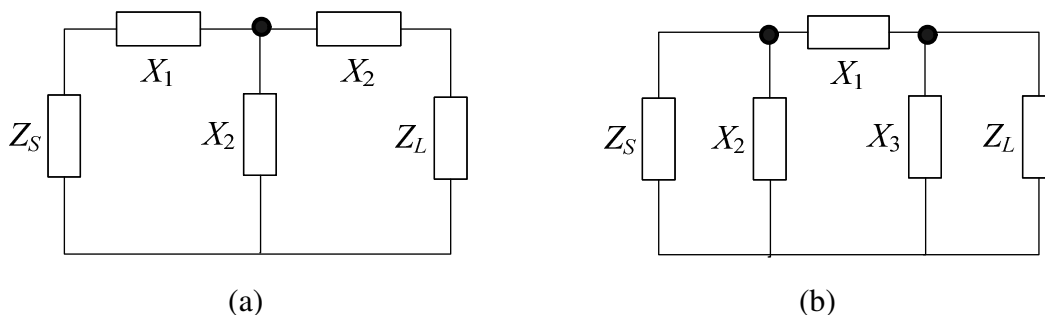

(a)   (b)

Figure 2.4. Three component matching networks: (a) T network and (b) Π network [33].

2.2.5   Classification of PAs

PAs can be divided into several categories. They are commonly grouped into broadband and narrowband amplifiers. Alternatively, they can be grouped depending on whether they are intended for the linear or constant envelope operation [30]. However, the most common grouping of PAs is grouping into classes according to the nature of their voltage and current waveforms. The variety of PA classes reflects the inability of any single circuit to satisfy stringent requirements for linearity, power gain, output power and efficiency.

Each class is designated with a letter or a combination of letters of the alphabet. The following classes are commonly used for different applications: A, B, AB, C, D, DE, E, F, FE, G, H and S [1, 8, 34]. Inverse classes, where the shape of voltage and current waveforms across the power transistors are inverted, are also possible. Common examples are inverse Class-C (C^{-1}) and inverse Class-F (F^{-1}) amplifiers [35]. Most of the real-life PAs are operating with current and voltage waveforms that lie between two different classes.

Not all of the classes are suitable for use at RF. For example, Class-D amplifiers are the switching-mode PAs generally used in low-frequency applications [36]. One such PA and its waveforms is shown in Figure 2.5 [37]. Ideally, only one of the two transistors in this figure is switched at a time and efficiency is 100% [31]. However, the use of this PA at high frequencies is limited due to prominent parasitic reactances that lead to substantial losses. This high-frequency performance degradation is less severe in current-mode Class-D implementations [36]. Class-S amplifiers are similar to those of Class D, but their input is driven with pulse-width modulated (PWM) waveform. This class needs transistors Q_1 and Q_2 to be switched with frequencies much higher than the signal frequency, thus making them impractical in RF range [1]. Class-G and Class-H amplifiers are also commonly used for audio applications, with some limited use in digital telephony and for code divison multiple access (CDMA) at low megahertz frequencies.

Traditionally, power amplification at RF was done with amplifier classes A through C, often termed classic amplifiers [8]. Class-E and Class-F amplifiers are considered modern amplifiers since they can be used in many hi-end applications. Because of their importance, all these amplifiers will be presented in separate sections.
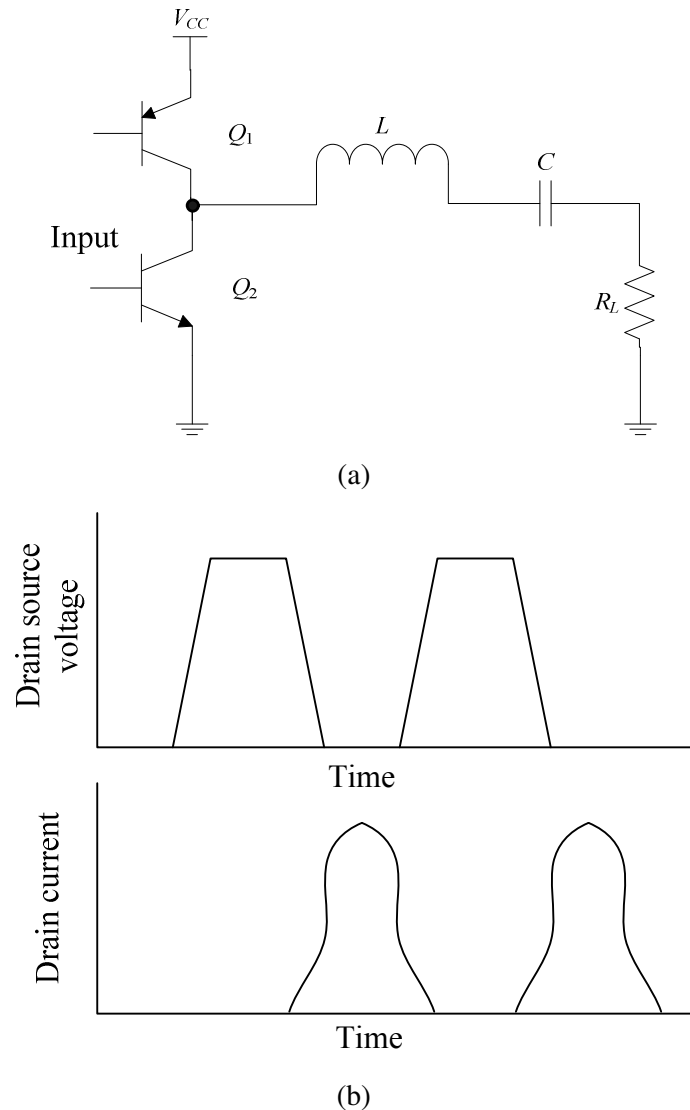
Figure 2.5. General Class-D amplifier: (a) circuit, (b) voltage and current waveforms.

2.2.5.1 Classes A through C

Class-A, B, AB and C amplifiers, or classic amplifiers, are usually analyzed together because of the minimal differences in their waveforms. Figure 2.6 shows a circuit operating as either Class-A, Class-AB, Class-B or Class-C amplifier [8]. Depending on the biasing, the fraction of the cycle during which the current through the PA is flowing (defined as conduction angle or 2θ) is different resulting in different class operation, as analyzed in Table 2.1, with efficiency to be explained later in this section. The voltage (v_c) and current (i_c) waveforms marked in Figure 2.6 are shown in Figure 2.7 for each class. In this figure, current i_c is sinusoidal waveform in all cases, with its bottom part cut off by biasing (V_b) for classes B and C. The waveform for the Class-AB amplifier, shaped between the Class-A and B waveforms, is not shown in this figure.
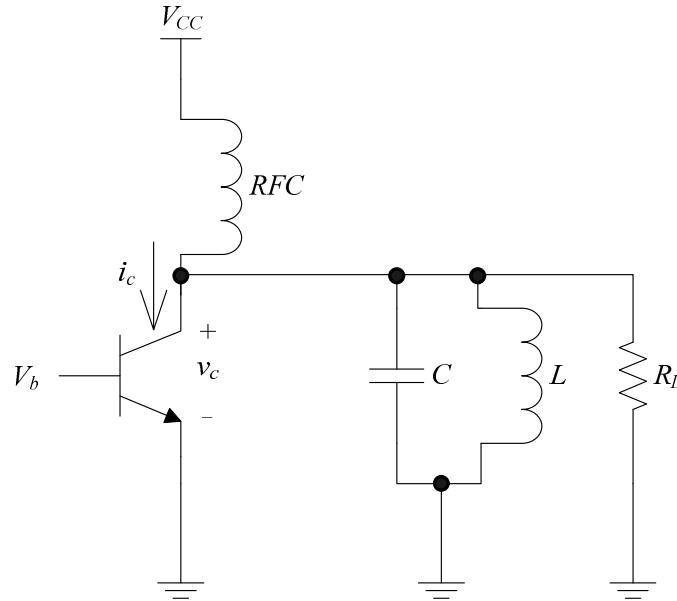
Figure 2.6. Circuit schematic of a single ended Class-A, AB, B or C PA used for theoretical analysis [8].

Table 2.1. Comparison of conduction angles and efficiencies for Class-A through C PAs [1].

Class	Conduction angle (°)	Maximum theoretical efficiency (%)	Normalized output power
A	360	50	1
AB	360 – 180	50 – 78.5	1
B	180	78.5	1
C	180 – 0	78.5 – 100	1 – 0

Classic PAs can be analyzed for efficiency as follows [1]. The DC component of the current, I_{DC}, can be derived as

$$I_{DC} = \frac{I_{CC}}{\pi}\left(\sin\theta - \theta\cos\theta\right), \tag{2.8}$$

where θ is a half of the conduction angle and I_{CC} is the peak of the current waveform across the transistor if it was operating in the Class-A mode (Figure 2.7 (b)). Power from the power supply is given by

$$P_{dc} = V_{CC}I_{DC} = \frac{V_{CC}I_{CC}}{\pi}\left(\sin\theta - \theta\cos\theta\right) \tag{2.9}$$

Fundamental current i_1 is derived as

$$i_1 = \frac{I_{CC}}{2\pi}\left(2\theta - \sin 2\theta\right) \tag{2.10}$$
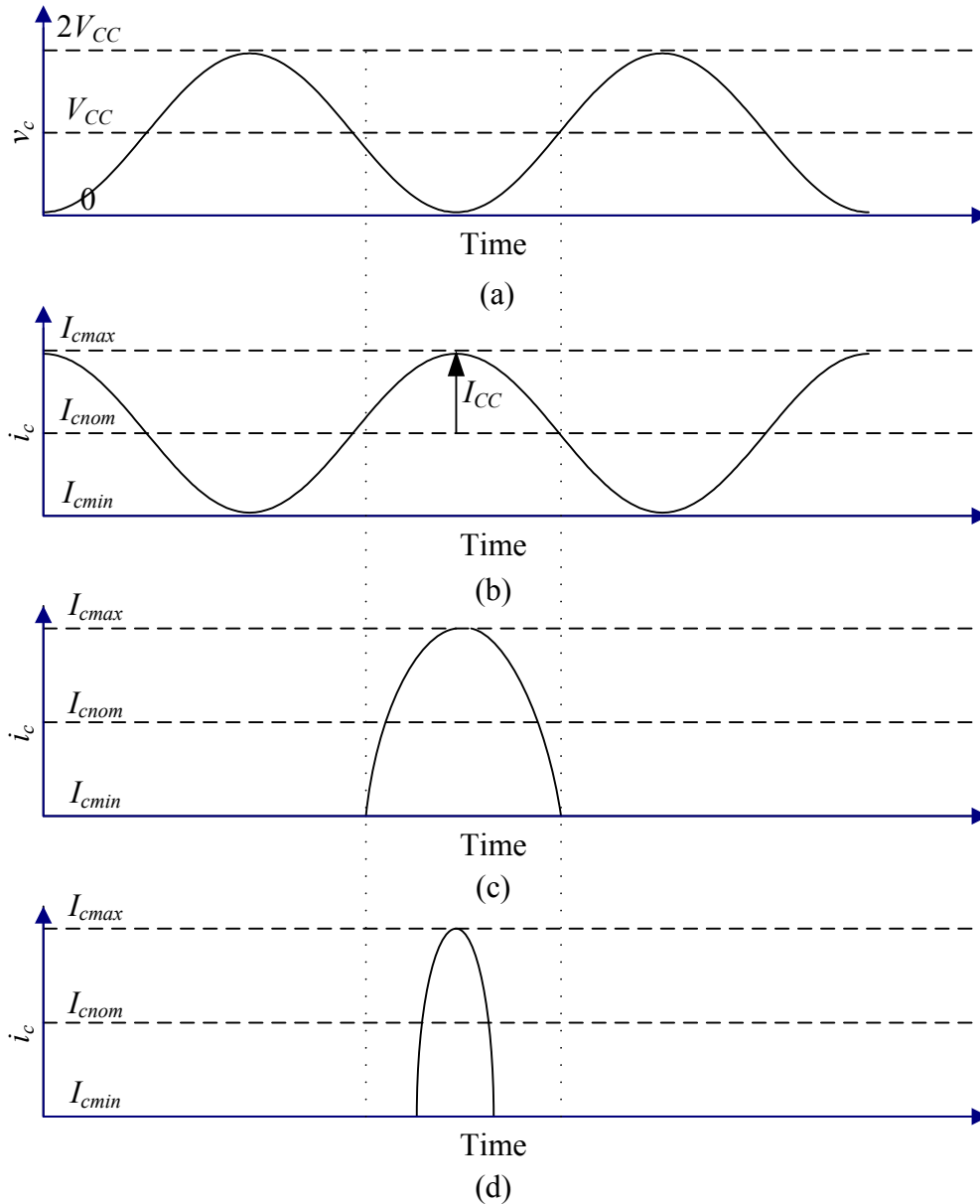
Figure 2.7. Voltage and current waveforms for classic amplifiers: (a) voltage for classes A, B and C, (b) current for Class A, (b) current for Class B and (c) current for Class C [8].

Output power is

$$P_{out} = \frac{v_{peak} i_{peak}}{2} \tag{2.11}$$

Maximum power will be reached if $v_{peak} = V_{CC}$. Hence (assuming $V_{CEsat} = 0$),

$$P_{out,\max} = \frac{V_{CC} i_{peak}}{2} = \frac{V_{CC} I_{CC}}{4\pi} (2\theta - \sin 2\theta) \tag{2.12}$$

The maximum possible efficiency is then

$$\eta = \frac{P_{out,\max}}{P_{DC}} = \frac{2\theta - \sin 2\theta}{4(\sin \theta - \theta \cos 2\theta)} \quad\quad (2.13)$$

Figure 2.8 shows the plot of maximum normalized output power versus θ. Maximum possible efficiencies and regions of operation are also marked on the figure. From this figure, it is clear that maximum efficiency of 100% can be obtained with the Class-C operation, but with zero output power, which makes such configuration impractical. Usually, the compromise between efficiency and output power is considered optimum design solution, resulting in θ lying in 30 to 60 degrees range. It should be noted that linearity of this amplifier is highest with the Class-A operation, while it decreases for the Class-AB, B and C operation. The passives C and L in Figure 2.6 are not explicitly defined by the PA formulae as in the case of the Class-E amplifiers described in Section 2.2.5.2, but they must be chosen for the amplifier to operate at correct resonant frequency given by

$$f_o = \frac{1}{2\pi\sqrt{LC}}, \quad\quad (2.14)$$

with acceptable Q-factor

$$Q = \frac{2\pi f L}{R_L} \qu\quad (2.15)$$

Efficiencies for monolithic Class-A through C PAs are much lower than ideal values due to a number of factors: low quality factor of inductors, saturation voltage in the transistors, tuning errors and temperature variations. Because high efficiencies are possible only at low conduction angles which are difficult in ICs, fully integrated classic PAs are rare, but not impossible. One example is a Class-AB amplifier for Bluetooth applications [38].

2.2.5.2   Class E

Class-E amplifiers were proposed by Sokal and Sokal in 1975 [39] and together with the Class F amplifiers [40] have been used in communication ever since. They are classified as switching amplifiers and as such they can exhibit efficiencies close to 100%. A single ended Class-E PA is shown in Figure 2.9 [37]. A simple PA analysis can be performed if the following is assumed:
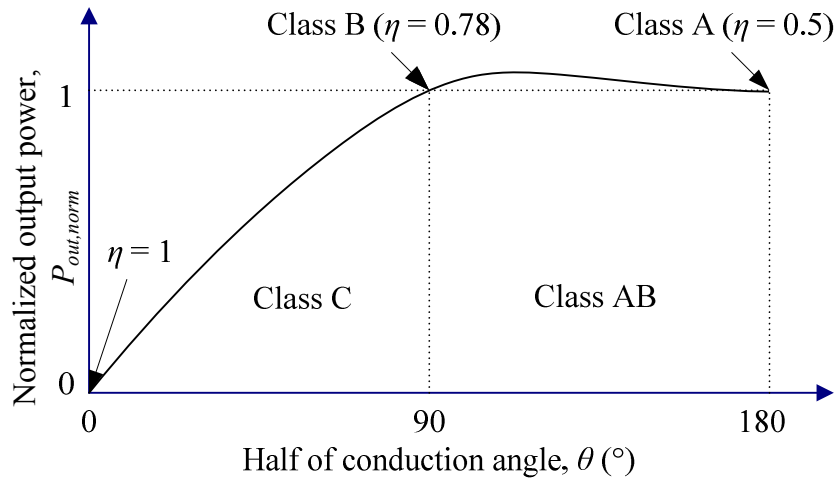
Figure 2.8. Maximum normalized output power versus half of conduction angle for the classic PAs [1].

- Inductance of the RF choke (L_1) is very high;
- Output capacitance of the transistor is independent of the switching voltage;
- Transistor is an ideal switch with zero resistance and zero switching time, open for half of the signal period.

From [39], the value of the optimum load resistance to deliver the highest power to the load with $v_{peak} = V_{CC}$ and non-zero V_{CEsat} is

$$R_L = \frac{2}{\frac{\pi^2}{4}+1}\frac{\left(V_{CC}-V_{CEsat}\right)^2}{P} = 0.577\frac{\left(V_{CC}-V_{CEsat}\right)^2}{P} \tag{2.16}$$

For desired Q-factor of the output resonant tank, inductance L_2 can be calculated:

$$L_2 = \frac{Q_L R_L}{2\pi f} \tag{2.17}$$

Shunt capacitance C_1 is given by

$$C_1 = \frac{1}{2\pi f R_L \left(\frac{\pi^2}{4}+1\right)\left(\frac{\pi}{2}\right)} = \frac{1}{5.447(2\pi f R_L)} \tag{2.18}$$

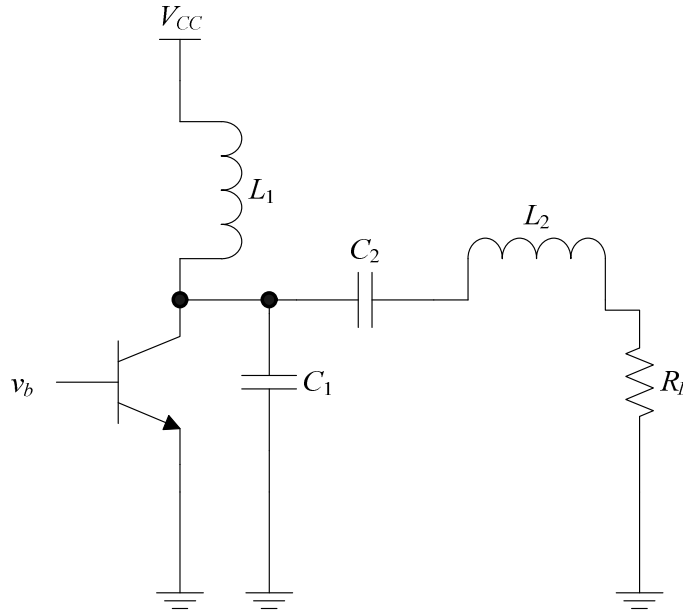and resonant capacitance C_2 is given by

Figure 2.9. Circuit diagram of a single ended Class-E PA used for theoretical analysis [37].

$$C_2 = \frac{1}{(2\pi f)^2 L_2} \cdot \left(1 + \frac{1.42}{Q_L - 2.08}\right) = C_1 \left(\frac{5.447}{Q_L}\right) \cdot \left(1 + \frac{1.42}{Q_L - 2.08}\right) \tag{2.19}$$

Capacitance C_1 includes parasitic capacitances of the transistor.

Efficiency of Class E is

$$\eta = \frac{1 - \frac{(2\pi A)^2}{6} - \frac{V_{CEsat}}{V_{CC}} \left[1 + A - \frac{(2\pi A)^2}{6}\right]}{1 - \frac{(2\pi A)^2}{12}}, \tag{2.20}$$

where

$$A = \left(1 + \frac{0.82}{Q_L}\right) f t_f \tag{2.21}$$

and t_f is the collector current fall time. For the ideal PA t_f and V_{CEsat} are zero, and efficiency is 100%. This efficiency is obtained at 78.5% of the Class-AB output power [1]. Practical waveforms for this topology with transistor turned on and off are shown in Figure 2.10.
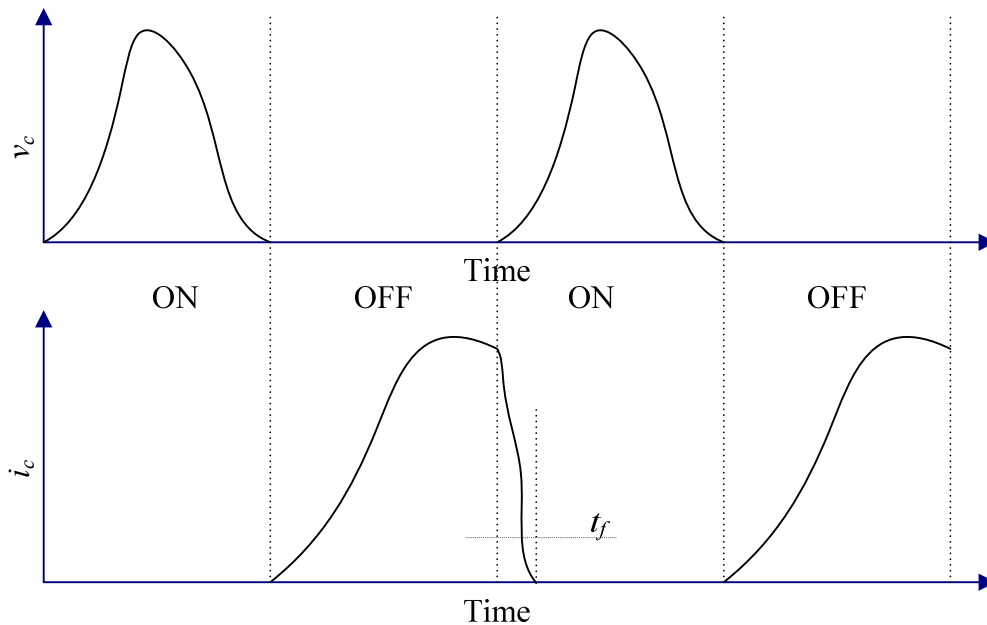
Figure 2.10. Voltage and current waveforms for the Class-E amplifiers when transistor is ON and OFF [39].

Although the Class-E amplifiers have output powers typically a few dB less than their Class-AB counterparts, they are used because of their high efficiency and swing that is more than three times the supply voltage. A recent example of the Class-E amplifier design can be found in [41].

2.2.5.3   Class F and F^{-1}

The Class-F amplifier is similar to that of the Class E in a sense that it can also result in 100% efficiency. However, at the same time, the output power can be similar to that of the Class-AB amplifiers [42]. This efficiency and power boost is a result of the presence of harmonic resonators in the output networks that shape drain (collector) waveforms in such a way that load appears to be short at even harmonics and open at odd harmonics, as shown in the general model in Figure 2.11. As a result, the ideal collector voltage waveform approximates a square wave, while collector current waveform approximates a half-sine wave. There is no overlap between the two waveforms. In case of the Class-F^{-1} amplifier resonators are swapped around, and the collector voltage is shaped as a half-sinusoid and collector current is shaped as a square wave.
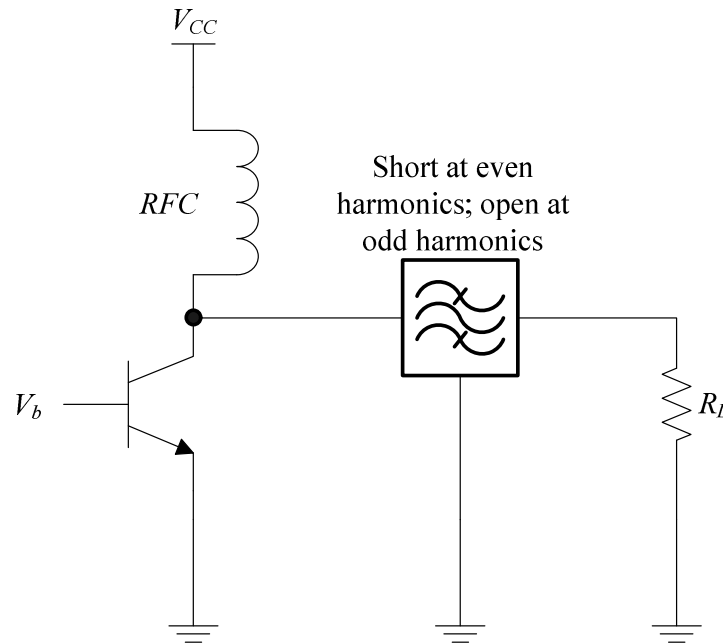
Figure 2.11. General Class-F amplifier model [42].

Shaping of waveforms can be done by means of transmission lines, as the case is with the Class-F amplifier with the quarter-wave transmission line. However, this is not a practical implementation for the low-gigahertz integration. Instead, passive resonators are used. Monolithic implementations of the Class-F amplifiers would require an infinite number of resonators to correctly shape output waveforms. Figure 2.12 shows efficiencies that can be obtained with increasing number of harmonic traps [43]. If only first harmonic trap is implemented, amplifier behaves as a Class-A amplifier, with maximum efficiency of 50%.

Most real life integrated Class-F amplifier implementations consider only a few harmonics, usually two or three. Figure 2.13 shows the third harmonic peaking circuit [42]. In this circuit, the tank at $3f_o$ provides an open circuit at $3f_o$ and short circuit at $2f_o$. A resonant tank at f_o ensures optimum load at f_o.

If fundamental frequency components of voltage and current are V_{om} and I_{om}, waveform factors that relate them to the fundamental frequency components V_{CC} and I_{DC} can be defined as [42]

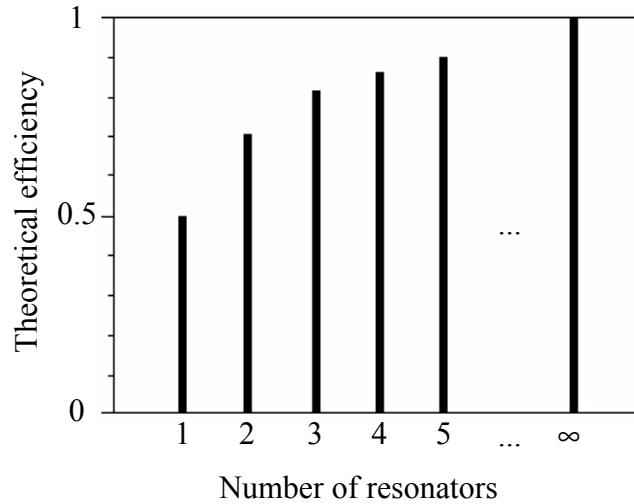$$V_{om} = \gamma_V V_{CC} \tag{2.22}$$

and

Figure 2.12. Efficiency of a Class-F amplifier versus a number of resonators [43].
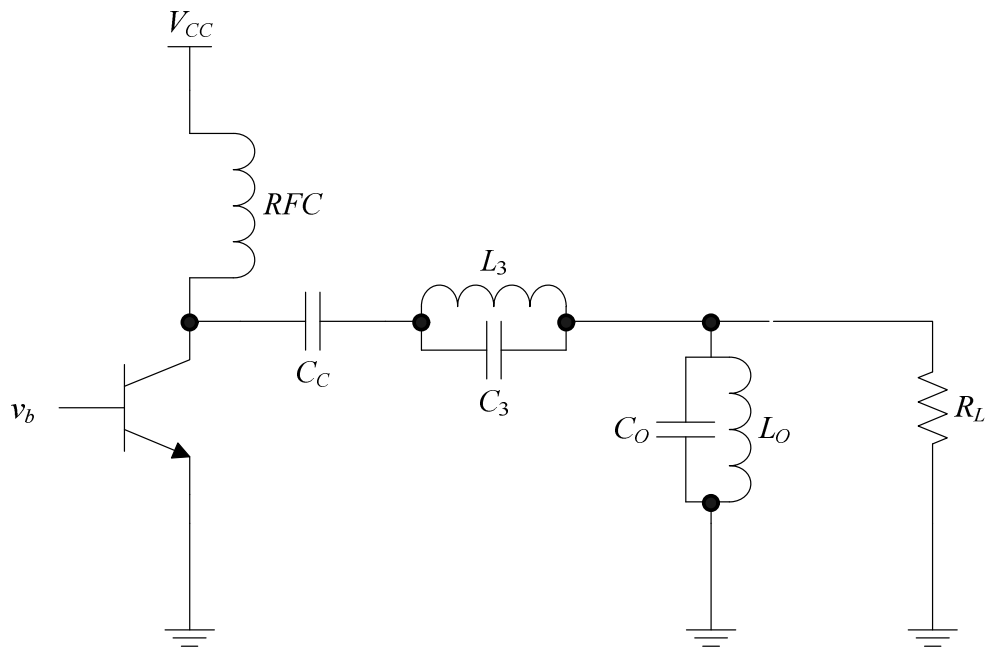

Figure 2.13. Class-F third harmonic peaking circuit [43].

$$I_{om} = \gamma_I I_{DC}, \tag{2.23}$$

whilst the efficiency of a Class-F circuit is

$$\eta = \frac{\gamma_V \gamma_I}{2} \tag{2.24}$$

For the third harmonic peaking circuit, γ_V and γ_I are calculated in [43] for the highest possible efficiency as 1.1574 and 1.4142 respectively, resulting in an efficiency of $\eta = 81.65\%$.

As it is the case with other classes, the Class-F amplifiers also exhibit efficiencies lower than predicted, mostly because of the presence of a large number of resonant circuits that suffer a loss because of the low Q-factors possible for on-chip inductors. A recent monolithic Class-AB/Class-F implementation in the InGaP/GaAs HBT process can be found in [44].

2.2.6   Techniques for performance improvement of PAs

Quality of the PA output is usually improved by various linearization techniques [45-47]. These include feedback, feedforward, pre- and postdistortion outphasing, as well as PWM and supply modulation. Efficiency can be boosted by means of adaptive bias and Doherty techniques [48]. Power combining is usually used to increase the total output power of a PA system. This can be performed on- and off-chip [1]. A good example of a high power amplifier that employs power combining can be found in [49]. The circuit uses distributed active transformer (DAT) architecture to deliver 35 dBm of power from an IC fabricated in 130 nm CMOS technology. Further two power combining transformer configurations are described in [50].

2.2.7   Temperature aspects of PAs

Efficiency, discussed in Section 2.2.3, is the ability of the PA to convert the electrical energy into output power. Excess power is converted to heat, which can limit the performance of an amplifier [51]. With increased efficiency, the amount of heat generated decreases. However, even with high efficiency, high output power configurations such as DAT architecture [49] will dissipate significant amounts of heat. The amount of heat generated and the way in which that heat is dissipated in a PA system depends on the technology in which the IC is fabricated and on the type of the active device (transistor) used for power amplification.

As detailed later in Section 2.4, SiGe technology is better choice than the rival GaAs technology from perspective of heating because the high conductive silicon substrate [52] facilitates heat distribution around the chip.

MOS transistor is less prone to heating than a BJT (HBT) because of its negative temperature coefficient [53]. This negative coefficient causes the drain current of the

MOSFET to decrease with temperature, which allows multiple MOSFETs to be connected in parallel. In HBTs, the situation is reversed. If multiple emitter fingers are used and are not perfectly balanced, the emitter with higher current will tend to take up even higher proportion of current, which gives rise to local hot spots and possibly permanent device damage [51]. Because current is flowing only in a small hot region, the total gain of the device decreases [54]. Normally, to compensate for this current gain decrease, ballast resistors are added in series with the emitter or base electrodes. This acts as negative feedback, since with the increase in current, the voltage across ballast resistors will also increase, counter-effecting the current increase. Including ballasting resistors, however, decreases the output power, power gain and the efficiency of the amplifier.

Typically, if heating poses a problem, the excess heat can be removed by means of heat sinks. The ground bondings should be adopted to favour heat sinks [55]. Usually, the chip is glued to a gold- or nickel-plated copper heat sink using a conductive adhesive [56].

2.2.8   Biasing

In preceding sections, it has always been assumed that the correct drive level is applied at the input of the PA. Input impedance matching, described in Section 2.2.4, is used to ensure that correct amplitudes of AC signals appear at input. Biasing, described in this section, provides the appropriate quiescent point for the PA [33].

Biasing point should remain constant irrespective of transistor parameter variations or temperature fluctuations. Active and passive biasing networks are possible. Figure 2.14 shows two passive biasing networks commonly used with BJT PAs.
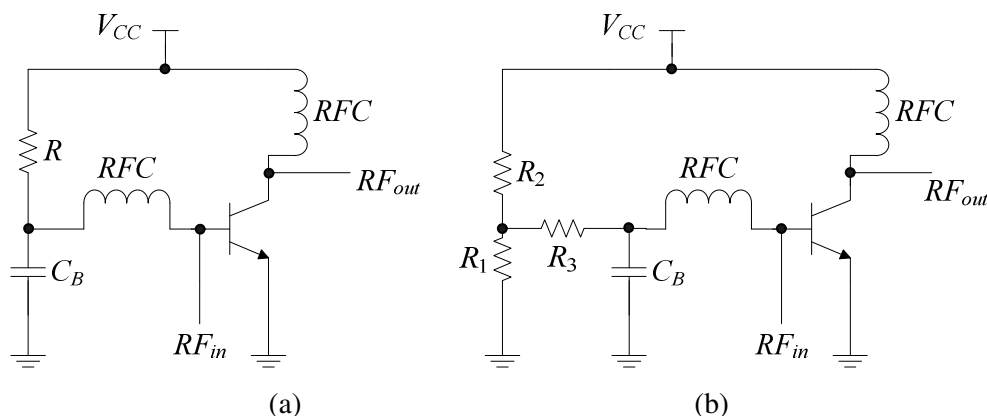

Figure 2.14. Passive biasing networks for a BJT PA: (a) One-resistor configuration, (b) Three-resistor configuration [33].

2.3   INDUCTORS FOR POWER AMPLIFIER IMPLEMENTATIONS

The demand for low-cost ICs has generated a high interest in integrated passive components. Passive components include resistors, capacitors and inductors.

A real inductor is usually modelled as an ideal inductor L_S in series with a resistor R_S, both in parallel with capacitor C_S, as shown in Figure 2.15 [33]. Inclusion of the series resistor and parallel capacitor is necessary to model the losses of the inductor even at frequencies below RF because Q-factor is in general much lower for the inductor then for other passive components. The Q-factor of the inductor is defined as 2π times the ratio of energy stored in the device and energy lost in one oscillation cycle. If Z is defined as impedance of an inductor, then the Q-factor is given by equation

$$Q = \frac{\mathrm{Im}(Z)}{\mathrm{Re}(Z)} \tag{2.25}$$

For the simple circuit in Figure 2.15, Equation (2.25) reduces to

$$Q = \frac{X}{R_S}, \tag{2.26}$$

where X is total reactance of the inductor. Q-factor is heavily dependent on frequency and exhibits a peak Q_{max}.
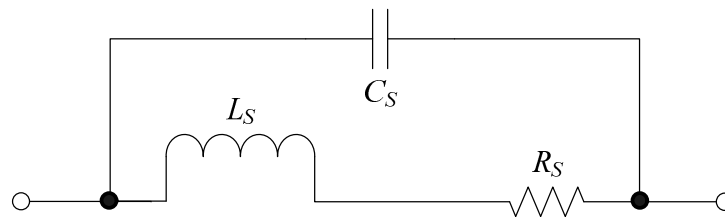

Figure 2.15. General high frequency model of an inductor [33].

While an ideal inductor exhibits constant impedance value for all frequencies, every non-ideal inductor exhibits impedance value dependent on frequency, as shown in Figure 2.16. Frequency where magnitude of impedance ($|Z|$) peaks is called the resonant frequency of an inductor. The resonant frequency, $f_r = \dfrac{1}{2\pi\sqrt{L_S C_S}}$ should ideally peak at infinity, but the finite value of the peak is due to the resistance R_S. Similarly, capacitance

C_S is the reason the inductor exhibits capacitive instead of inductive behaviour at frequencies above the resonance.
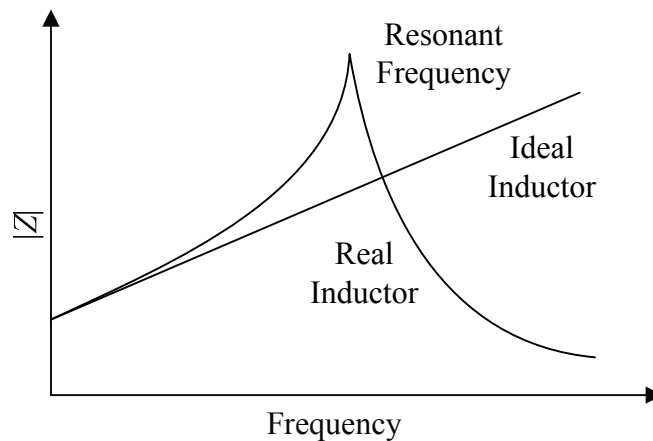

Figure 2.16. Frequency response of the impedance of ideal and real inductors [33].

2.3.1   Inductor implementation options

Traditionally, capacitor and resistor implementations are easily accomplished in CMOS and BiCMOS and are almost exclusively fabricated on-chip. In case of inductors, integrated passive (spiral) inductors are not such an obvious choice. Various factors, such as inductor size and low Q-factor of integrated passive inductors often result in one of the following inductor alternative implementations:

- External inductors,
- Active integrated inductors,
- Microelectro-mechanical systems (MEMS) inductors,
- Bond wires, or
- Other on-chip implementations.

Each of the above possibilities is considered as an alternative to the passive inductor implementation presented in this thesis and is discussed in more detail in the subsections that follow.

2.3.1.1   External inductors

External or off-chip inductors are connected to a system outside of the IC package. They are usually implemented as a solenoidal coil or a toroid, with a structure and a photo of a solenoid shown in Figure 2.17 [9]. Their usage at high frequencies also implies careful printed circuit board (PCB) modelling and design. Although high quality inductors are

widely obtainable from suppliers [57], inductance values available are usually limited to standard values of 10 nH and higher. Frequency of the Q-factor peak (typically in range of hundreds) is also predefined and is usually located in either high megahertz or low gigahertz range. Another drawback for integrated design is that the value obtained in real-life will differ from the anticipated value due to parasitics involving PCB tracks, IC bonding, and others.
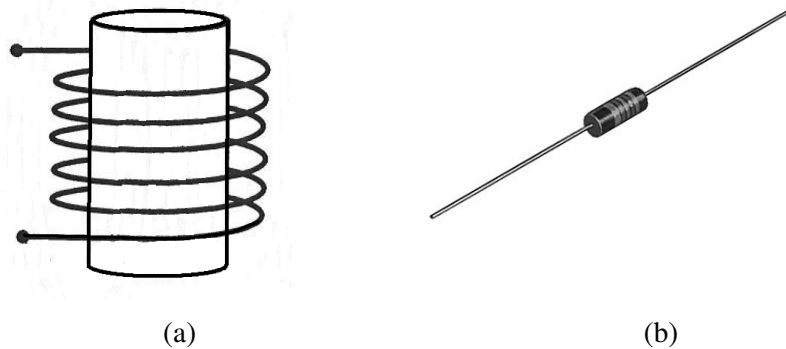

<center>(a)   (b)</center>

Figure 2.17. A solenoidal off-chip inductor: (a) structure, (b) photo [9].

2.3.1.2   Active integrated inductors

Active inductors are a good alternative to integrated passive inductors because of their higher Q-factor. Typical Q-factors that can be obtained for this configuration are between 10 and 100, which is up to ten times that of spiral inductors [58]. Active inductors also take up much smaller area on the chip than spiral inductors. The main disadvantages of active inductors are increased power consumption, presence of electrical noise from active devices and limited dynamic range. A design requiring only six transistors has been proposed [59], which makes active inductor much more compact compared to traditional designs requiring ten transistors or more. Active die area of this topology is only 30 µm by 65 µm, and it achieves a Q-factor of about 20 at 1 GHz whilst exhibiting differential inductance of 20 nH.

Many active inductors employ a gyrator approach [60]. An ideal gyrator is a linear two-port network that neither absorbs nor dissipates energy. Linearized model of an ideal gyrator is shown in Figure 2.18.
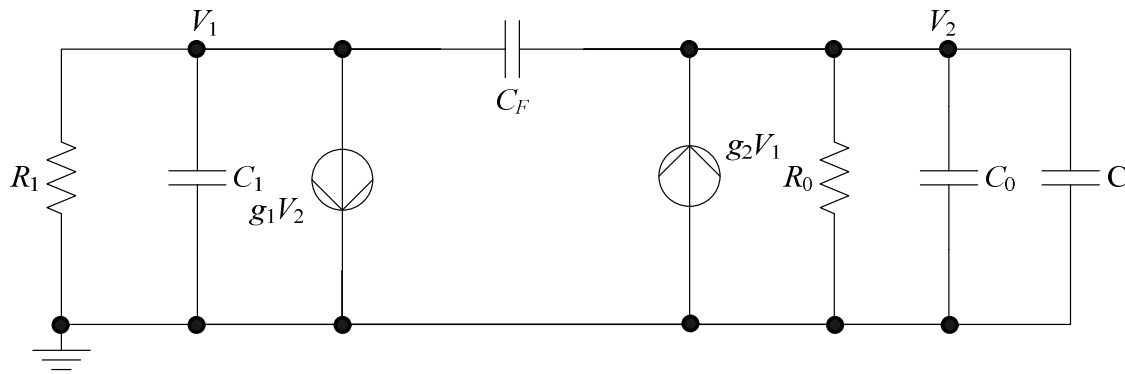
Figure 2.18. High frequency two-port equivalent model of an active inductor [60].

At low frequencies the effective inductance and series resistance can be accurately modelled in terms of capacitances (C_1, C_0 and C), output resistance (R_0) and transconductances (g_1 and g_2) represented in Figure 2.18. At higher frequencies, the parasitics become more prominent and the linearized model cannot be used any longer. In this case, the analysis of active inductors becomes specific to a given configuration. Real-life examples of active inductor implementations are included in references [58] and [59].

2.3.1.3   MEMS inductors

MEMS is an IC fabrication technique that empowers conventional two-dimensional (2-D) circuits to expand into the third dimension (3-D) [61]. This principle becomes particularly useful in inductor fabrication, because substrate parasitics can be reduced significantly. In later sections of this chapter it will be discussed that substrate losses are the main contributor to the low Q-factor of integrated passive inductors. A number of techniques for producing MEMS devices exist, of which bulk and surface micromachining are most commonly used for inductors. In bulk micromachining, a 3-D structure is created on the wafer by etching different atomic crystallographic planes in the wafer. In surface micromachining, the 3-D structure is created by the sequential addition and removal of thin film structural and sacrificial layers to and from the wafer while preserving the integrity of structural layers [61]. Effectively, silicon below the inductor is replaced by air which has the lowest possible relative permittivity ($\varepsilon_r = 1$), which allows the Q-factor and resonant frequency to approach the values of off-chip inductors. Typical obtainable Qs range from 10 to 30 for 1 nH inductors at multi-gigahertz frequencies.

An example of a high-Q micromachined inductor can be found in [62]. In this paper, a robust micromachined spiral inductor with a cross-shaped sandwich membrane support is

fabricated for RFIC applications. This inductor achieves a Q-factor of 50 for 4 nH inductor at 5 GHz. The photograph in Figure 2.19 illustrates the substrate removal at the corner of this robust inductor.
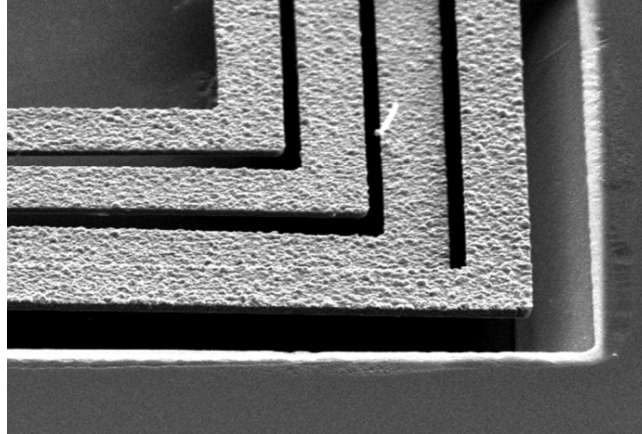

Figure 2.19. Photograph of the enlarged view of the corner region of the robust inductor illustrating removal of the substrate [62].

Alternative to spiral MEMS inductors, solenoidal inductors suspended on-chip can be used with various degrees of chip stability [63]. Several advantages over conventional (or MEMS) spiral inductors can be identified, which include a lower stray capacitance due to the fact that only a part of the inductor is lying on the silicon substrate, a simple design equation and greater possibilities for flexible layout. Advanced micromachining techniques for solenoidal inductors have been proposed, including the 3-D laser lithography, multiple-trenched sidewalls, the U-shaped solenoidal shape and a concave-suspending MEMS process.

The second alternative to the MEMS spiral inductors are out-of-plane inductors [64]. Coils are fabricated using stress-engineered thin films. Stress gradient is induced by changing the ambient pressure during film deposition. When released, a stress-graded film curls up in a circular trajectory as photographed in Figure 2.20 [64]. Typical Q-factor of this configuration is over 70 at 1 GHz.
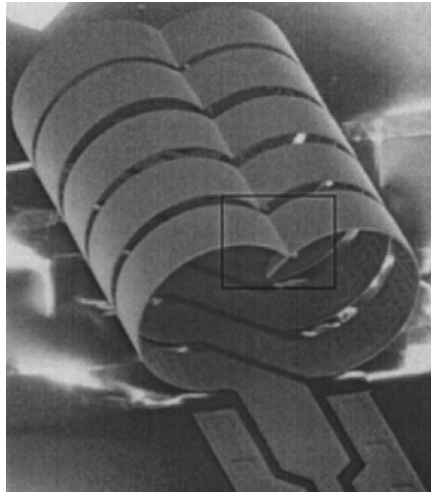
Figure 2.20. Out-of-plane inductor [64].

Although MEMS devices present an attractive alternative to conventional passive inductors, particularly because of the high Q-factors, the micromachining and suspending require modifications to standard fabrication procedures and/or modifications to the wafer after fabrication. Because after such modifications the repeatability [65], important for mass production of these devices, is not assured, MEMS will not be considered any further in this study.

2.3.1.4   Bond wires

Bond wires, which usually present a parasitic quantity for signals communicating between systems on the packaged chip and the outside world, reflect inductive behaviour that impacts module performance [66]. Electrical characteristics of bond wires depend on the material of which they are made of and its cross-section, the height above the die-plane, horizontal length and the pitch between the adjacent wires [67]. These characteristics can be used accurately to determine the Q-factor and inductance of bond wires in order to use them in RF applications. Although bond wires with Q-factors of 50 have been reported, their inductances will typically be less than 1 nH [67]. This limits their usability in gigahertz range where well controlled inductances of 1 nH and more are often needed.

2.3.1.5   Other on-chip implementations

Other inductor implementations are worth mentioning in a separate section since some fundamental differences from standard inductor types can be identified.

Masu, Okada and Ito [68] discuss two types of inductors not commonly found in literature. First type of inductor is a meander inductor. It is a flat passive inductor consisting of a long piece of metal that is not wound as in the case of the spiral inductor which will be described in detail later, but meanders as shown in Figure 2.21 (a). This inductor occupies a small area, but its measured Q-factor is quite low (about 2.1 for inductance of 1.3 nH). Such trade-off between the area and Q-factor is acceptable for matching network applications. The second type of inductor is a snake inductor similar to the previously mentioned MEMS solenoidal inductor, but wound as shown in Figure 2.21 (b). It occupies even smaller area than the meander inductor and its Q-factor lies in the same range.

Vroubel *et al.* discusses electrically tuneable solenoidal on-chip inductors [69]. Other tuneable inductors are commonly seen as implemented in active configuration, such as in the case of the inductor in [70]. The toroid inductors can also be implemented on-chip by means of micromachining. One such integrated inductor is shown in Figure 2.22 [71].
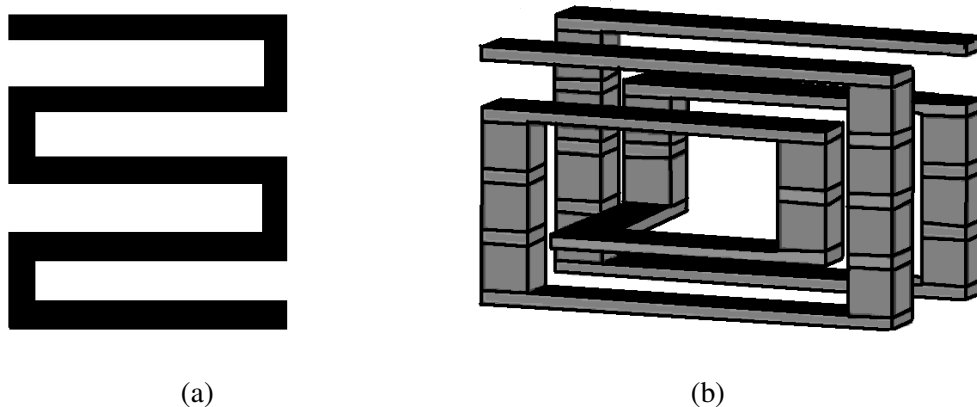

(a)   (b)

Figure 2.21. The meander (a) and snake (b) inductors [68].
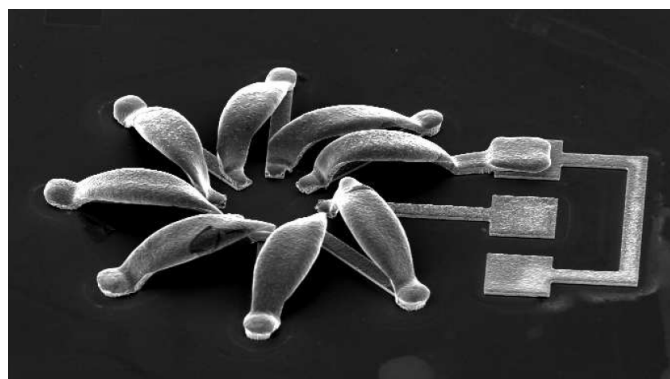

Figure 2.22. Photograph of an integrated toroid inductor [71].

2.3.2   Spiral inductors on silicon and SiGe

Although inductor implementations described in the previous section are widely used due to their advantages over passive integrated inductors, they are normally too complex to implement and make the PA devices too bulky and expensive. This leaves passive spiral inductors as a reasonable choice for PAs.

2.3.2.1   Common spiral inductor geometries

There are several spiral inductor geometries commonly used. These include square and circular inductors as well as various polygons [72]. The square spiral has traditionally been more popular since some IC processes constrain all angles to 90° [9], but it generally has lower Q-factor than the circular spiral, which most closely resembles the common off-chip solenoidal inductors but is difficult to lay out. A polygon spiral is a compromise between the two. Drawings of these geometries are shown in Figure 2.23.

The geometries shown in Figure 2.23 are asymmetric, and require only a single metal layer for fabrication. Additional layers are only needed to bring the signal lines to the outside of an inductor and are universally known as underpasses. Symmetrical inductors are also possible, but they require more than one underpass, in this case known as metal-level interchange. Such geometry is presented in Figure 2.24 [9].

The second metal layer can be used as part of the core of the inductors. An example of such multilayer geometry is a two-layer square inductor shown in Figure 2.25 [73]. The multi-layer geometries can deliver higher quality factors than a single layer inductor due to their more significant inductance coupling.

Another common geometry is a taper geometry, where inner spirals of inductors decrease in width in respect to the outer spirals [74] (Figure 2.26). Tapering is done to suppress eddy current losses in the inner turns in order to increase the Q-factor [74], but it is most effective when substrate losses are negligible [9].
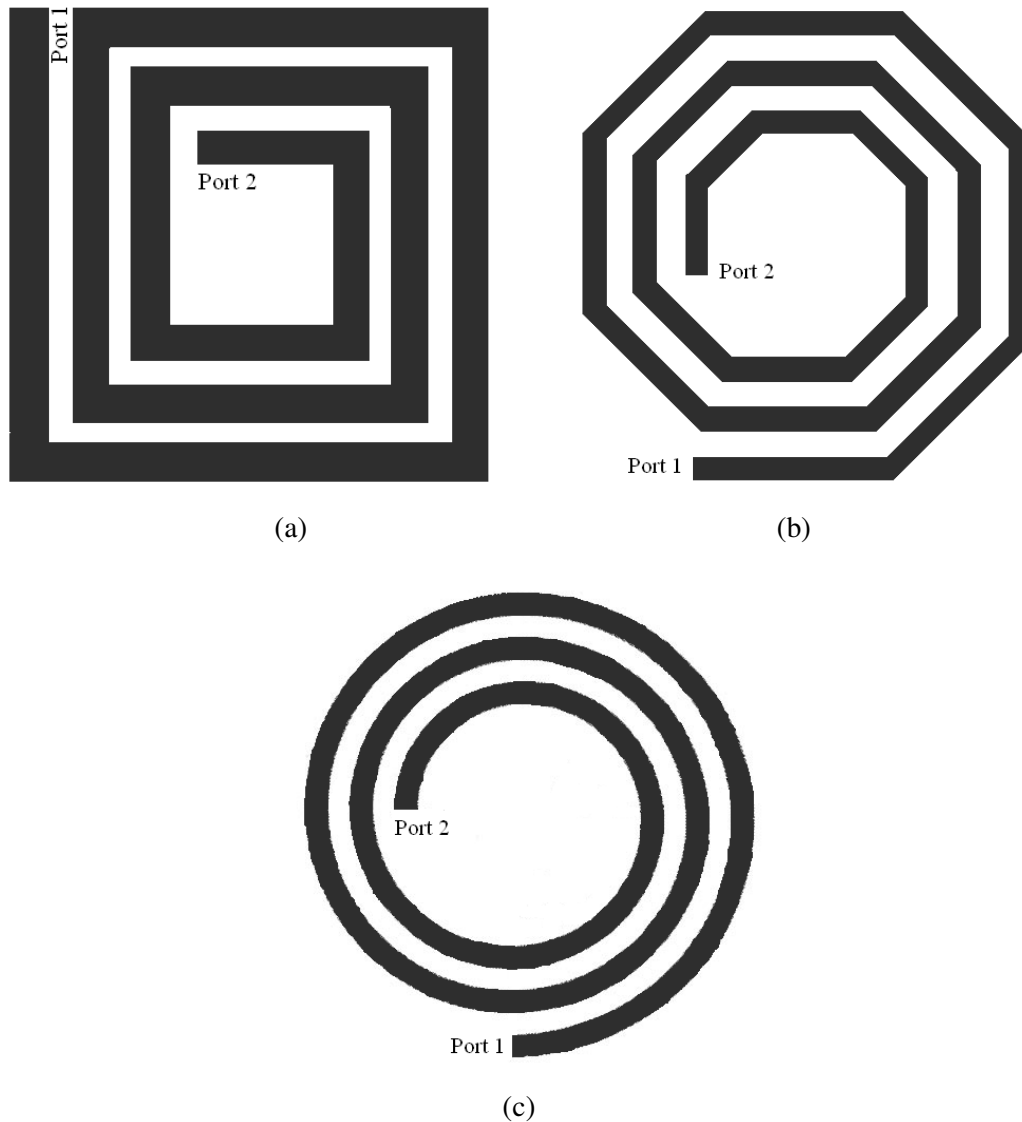
Figure 2.23. The square (a), polygonal (octagonal) (b) and circular (c) spiral inductors [9].

2.3.2.2   Spiral inductor geometry parameters

For a given geometry, a spiral inductor is fully specified by the number of turns (n), the turn width (w) and two of the following: inner, outer or average diameter (d_{in}, d_{out} or $d_{avg} = (d_{in} + d_{out})/2$), as shown in Figure 2.27 for the square and circular inductors. Spacing between turns, s, can be calculated from other geometry parameters. Another commonly used geometry parameter is the fill ratio, defined as

$$\rho = \frac{d_{out} - d_{in}}{d_{out} + d_{in}} \tag{2.27}$$

Total length of a spiral is also important for calculations. It is dependent on inductor geometry. For a square inductor it can be calculated as

Figure 2.24. The symmetrical spiral inductor [9].
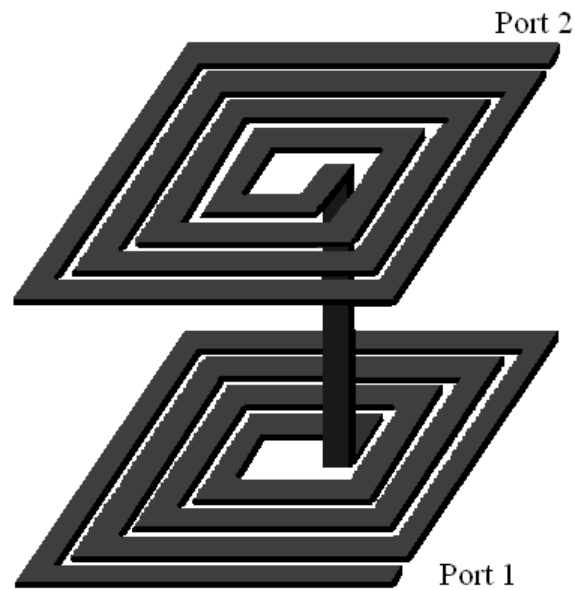
Figure 2.25. A multilayer (two-layer) spiral inductor [73].

$$l = 4(d_{in} + w) + 2n(2n-1)(s+w) \qquad (2.28)$$
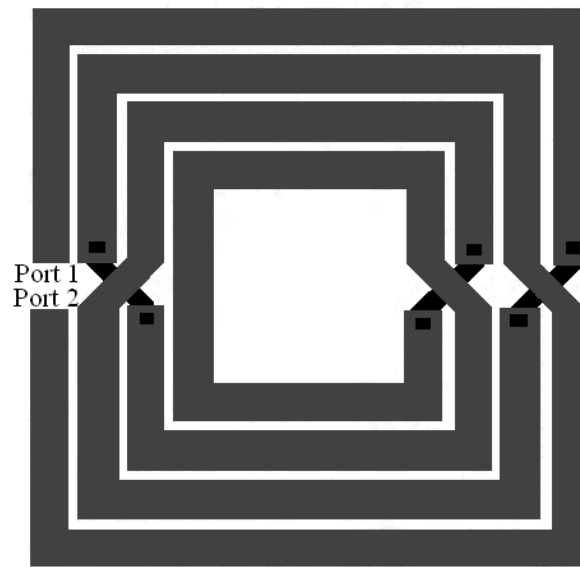
2.3.2.3   Spiral inductor models

Several spiral inductor models are widely used. In this section, a single-π, segmented, double-π and third-order models will be described.
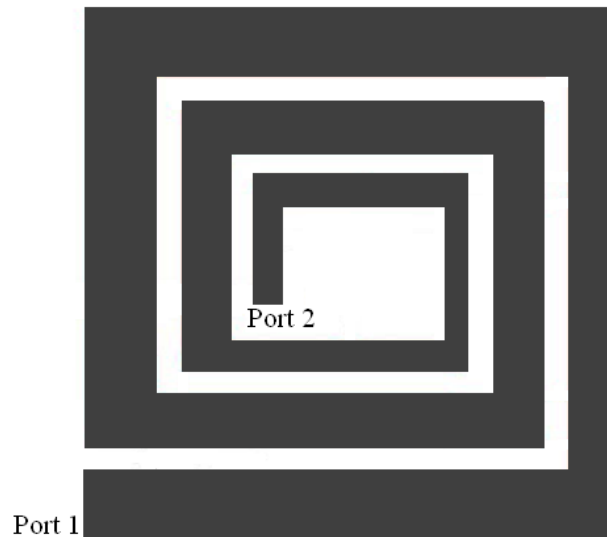
Figure 2.26. A taper spiral inductor [74].

(a)   (b)
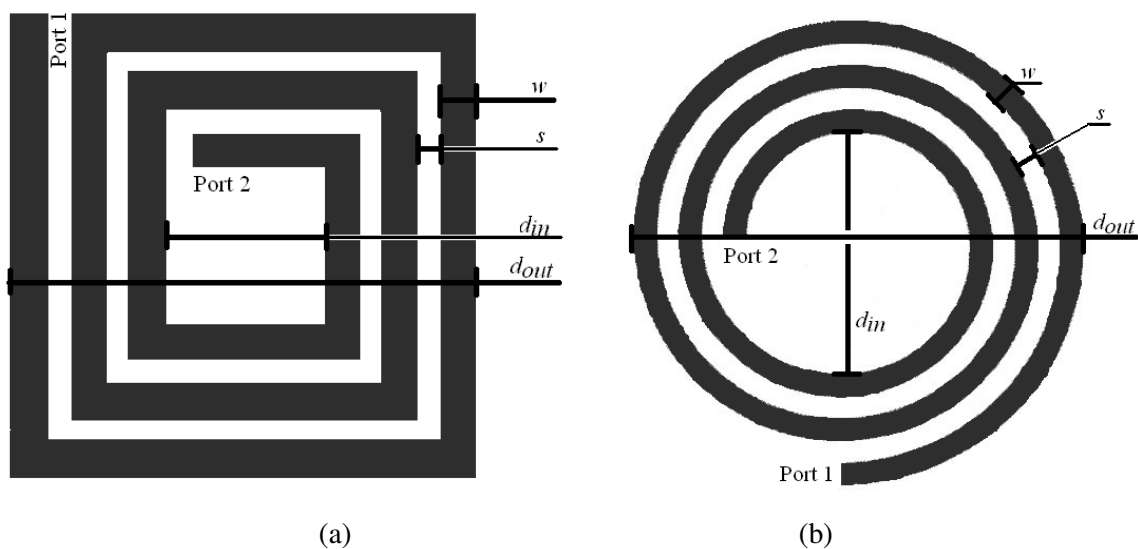
Figure 2.27. Geometry parameters of the (a) square and (b) circular spiral inductors.

Single-π model

Most commonly used model is a lumped single-π nine-component configuration shown in Figure 2.28 [10, 72]. In this model, L_S is the inductance at the given frequency, R_S is the parasitic resistance and C_S is the parasitic capacitance of the spiral inductor structure. C_{ox} is the parasitic capacitance due to oxide layers directly under the metal inductor spiral. Finally, C_{Si} and R_{Si} represent the parasitic resistance and capacitance due to the silicon substrate. This topology does not model the distributive capacitive effects, but it models correctly for parasitic effects of the metal spiral and the oxide below the spiral, as well as for substrate effects.
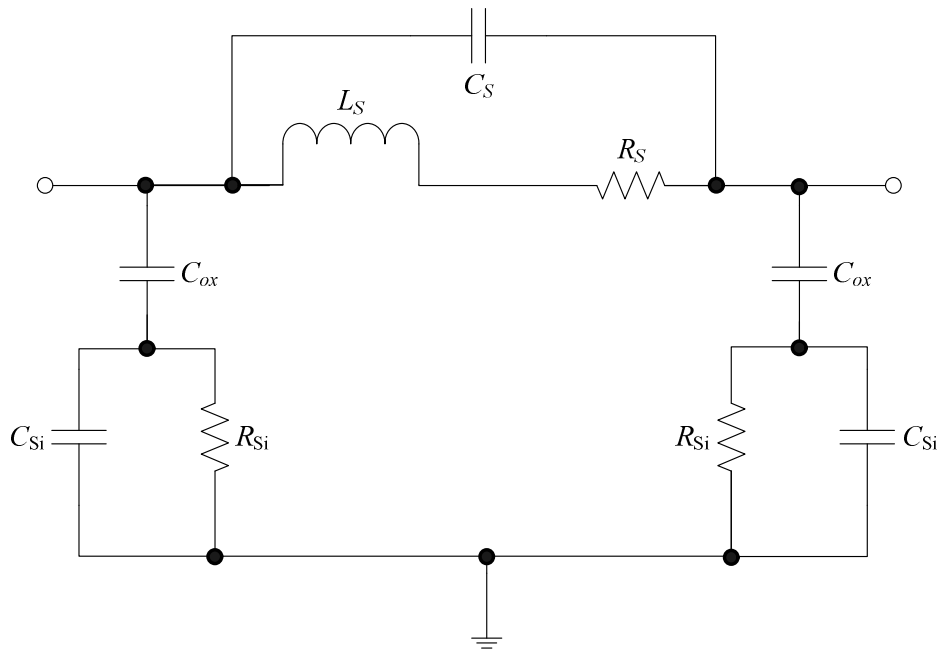
Figure 2.28. A commonly used nine-component spiral inductor model [72].

Segmented model

Somewhat more complicated model is the model presented in [75]. Each segment of the inductor is modelled separately with a circuit given in Figure 2.29. In this model, parasitics C_{ox}, C_{Si} and R_{Si} represent parasitics of only one inductor segment, L_S and R_S represent inductance and parasitic capacitance of one segment coupled to all segments, whilst capacitances C_{f1} and C_{f2} are added to represent coupling to adjacent segment nodes.

Double-π distributed model

The standard single-π model can also be extended into distributed double-π model shown in Figure 2.30 [11, 76]. A second order ladder (with third grounded branch) is used to model the distributive characteristics of metal windings. The interwinding capacitance (C_w) is included to model the capacitive effects between metal windings of the inductor. The transformer loops (M_{S1} and M_{S2}) represent the frequency-dependent series loss effects.

Third order transmission-line model

The second order model shown in Figure 2.30 is valid for the inductor up to the first resonance frequency. If a third order model is used, it is possible to accurately predict inductor behaviour even beyond the resonant frequency. One such model is presented in [12]. An equivalent circuit diagram for this configuration is given in Figure 2.31. Extrinsic admittances are used and all circuit components are self-explanatory from this figure.
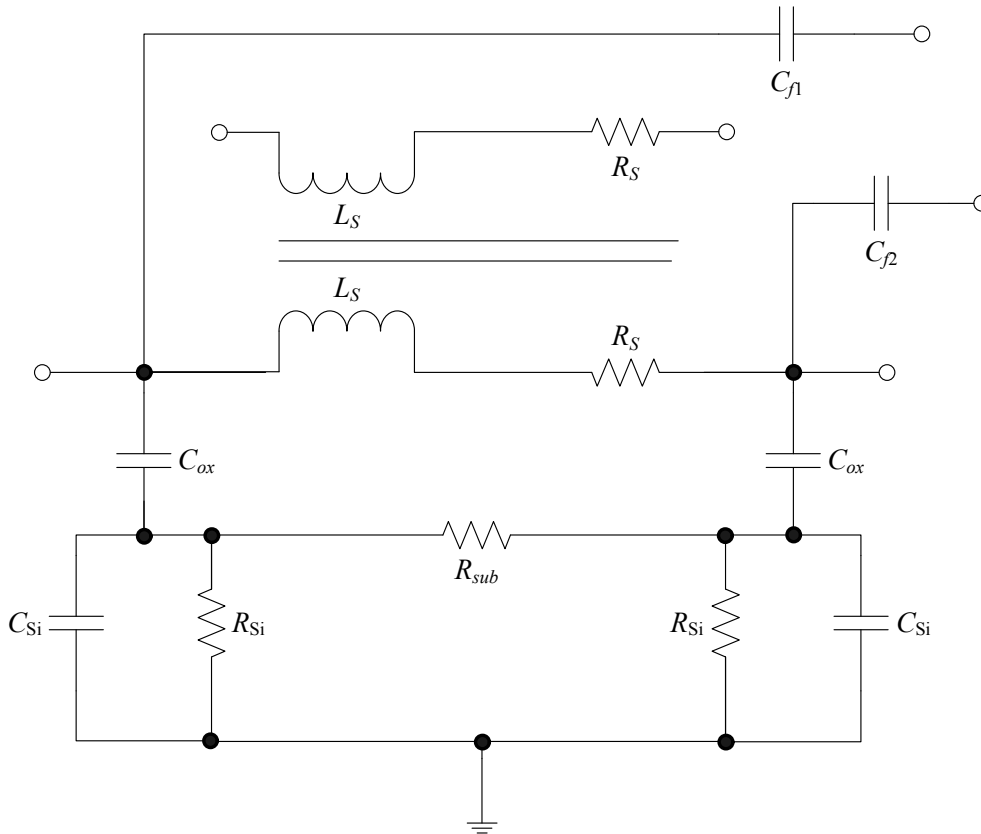
Figure 2.29. An equivalent two-port model for one segment of a spiral inductor [75].
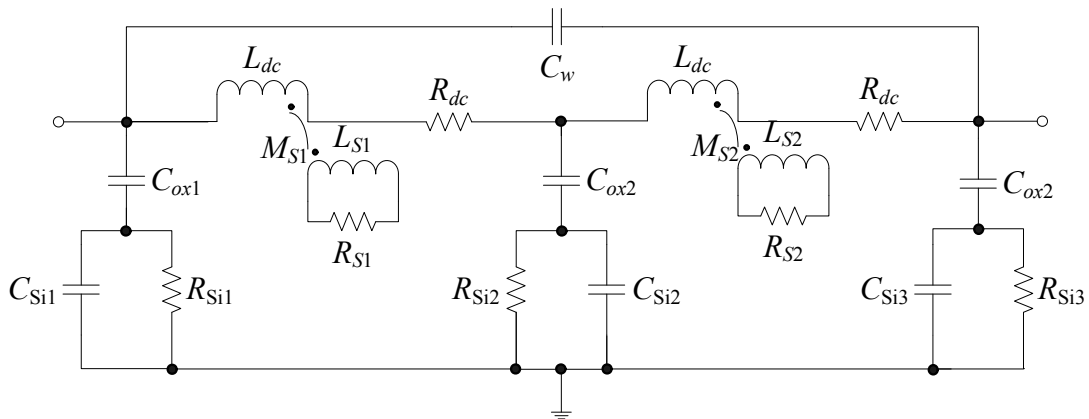
Figure 2.30. A double-π distributed inductor model [76].

2.3.2.4   Calculation of series inductance and parasitics for single-π inductor model

The single-π inductor model of Figure 2.28 is sufficient to accurately model spiral inductors for frequencies below resonance. In this section, series inductance L_S as well as parasitic capacitances and resistances shown in this figure are explained.
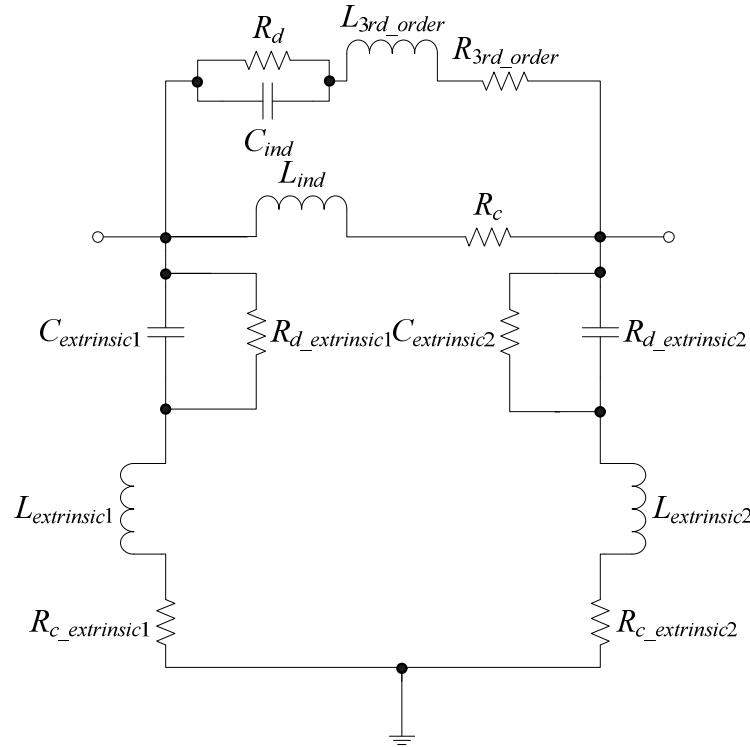
Figure 2.31. A complete third-order inductor model [12].

Series inductance (L_S)

Various equations are commonly used in literature to represent the series inductance of spiral inductors. These include the modified Wheeler equation, expression based on current sheet approximation [72], Bryan equation, as well as the data-fitted monomial expression [77].

The modified Wheeler equation is based on the equation derived by Wheeler in 1928 [72]:

$$L_{mw} = K_1 \mu \frac{n^2 d_{avg}}{1 + K_2 \rho}, \tag{2.29}$$

Here, *n* is the number of turns, d_{avg} is the average diameter of the spiral and ρ is the fill ratio, all defined in Section 2.3.2.2 . K_1 and K_2 are geometry-dependent coefficients with values defined in Table 2.2 and μ is magnetic permeability of the metal layer.

Table 2.2. Coefficients for the modified Wheeler expression [72].

Layout	K_1	K_2
Square	2.34	2.75
Octagonal	2.33	3.82
Hexagonal	2.25	3.55

Another expression can be obtained by approximating the sides of the spiral by symmetrical current sheets of equivalent current densities as described in [72]:

$$L_{gmd} = \mu \frac{n^2 d_{avg} c_1}{2} \left[\ln \frac{c_2}{\rho} + c_3 \rho + c_4 \rho^2 \right] \qquad (2.30)$$

Here, c_1, c_2, c_3 and c_4 are geometry dependent coefficients with values defined in Table 2.3. This expression exhibits a maximum error of 8% for $s \leq 3w$.

Table 2.3. Coefficients for the current sheet expression [72].

Layout	c_1	c_2	c_3	c_4
Square	1.27	2.07	0.18	0.13
Hexagonal	1.09	2.23	0.00	0.17
Octagonal	1.07	2.29	0.00	0.19
Circle	1.00	2.46	0.00	0.20

Bryan's equation is another popular expression for the square spiral inductance [77]:

$$L = 0.00241 \left(\frac{d_{out} + d_{in}}{4} \right) n^{\frac{5}{3}} \ln \left(\frac{4}{\rho} \right) \qquad (2.31)$$

The data-fitted monomial expression is an expression that results in an error smaller than seen in the expressions given above (typically less than 3%). It is based on a data-fitting technique. Inductance in nanohenries (nH) is calculated as [72, 77]:

$$L_{mon} = \beta d_{out}^{\alpha_1} w^{\alpha_2} d_{avg}^{\alpha_3} n^{\alpha_4} s^{\alpha_5} , \qquad (2.32)$$

where coefficients β, α_1, α_2, α_3, α_4 and α_5 are once again geometry dependent, as presented in Table 2.4.

Table 2.4. Coefficients for the spiral inductor inductance calculation [72]

Layout	β	α_1 (d_{out})	α_2 (w)	α_3 (d_{avg})	α_4 (n)	α_5 (s)
Square	$1.62 \cdot 10^{-3}$	-1.21	-0.147	2.40	1.78	-0.030
Hexagonal	$1.28 \cdot 10^{-3}$	-1.24	-0.174	2.47	1.77	-0.049
Octagonal	$1.33 \cdot 10^{-3}$	-1.21	-0.163	2.43	1.75	-0.049

The monomial expression has been developed by curve fitting over a family of 19000 inductors [72]. It has better accuracy and higher simplicity than the equations described above, making it useful for this thesis.

Parasitic resistance (R_S)

Parasitic resistance is dependent on the frequency of operation. At DC, this value is mostly dependent on the sheet resistance of the material from which the wire is made. At high frequencies, this is overshadowed by the resistance that arises due to formation of eddy currents. It depends on resistivity of metal layer in which the inductor is laid out (ρ), total length of all inductor segments (l), width of the inductor (w) and its effective thickness (t_{eff}) [10]:

$$R_S = \frac{\rho l}{w t_{eff}} \tag{2.33}$$

Effective thickness, t_{eff}, is dependent on the actual thickness of the metal layer, t:

$$t_{eff} = \delta(1 - e^{-t/\delta}), \tag{2.34}$$

where δ is skin depth dependent on frequency f via relation

$$\delta = \sqrt{\frac{\rho}{\pi \mu f}} \tag{2.35}$$

Parasitic capacitance (C_S)

Parasitic capacitance is the sum of all overlap capacitances created between the spiral and the underpass. If there is only one underpass and it has the same width as the spiral, then the capacitance is equal to [10]

$$C_S = n w^2 \frac{\varepsilon_{ox}}{t_{oxM1-M2}} \tag{2.36}$$

where $t_{oxM1-M2}$ is the oxide thickness between the spiral and the underpass and ε_{ox} is the dielectric constant of the oxide layer between the two metals.

Oxide and substrate parasitics (C_{ox}, C_{Si} and R_{Si})

Oxide and substrate parasitics are approximately proportional to the area of the inductor spiral ($l \cdot w$), but are also highly dependent on the conductivity of the substrate and the operating frequency. In order to calculate the oxide capacitance C_{ox} and substrate capacitance C_{Si}, the effective thickness (t_{eff}) and effective dielectric constant (ε_{eff}) of either oxide or substrate must be determined. Effective thickness is calculated as [78]

$$t_{eff} = w\left[\frac{w}{t} + 2.42 - 0.44\frac{t}{w} + \left(1 - \frac{t}{w}\right)^6\right]^{-1}, \text{ for } \frac{t}{w} \leq 1, \tag{2.37}$$

or

$$t_{eff} = \frac{w}{2\pi}\ln\left(\frac{8t}{w} + \frac{4w}{t}\right), \text{ for } \frac{t}{w} \geq 1 \tag{2.38}$$

for both oxide and substrate. Effective dielectric constant is determined as

$$\varepsilon_{eff} = \frac{1+\varepsilon}{2} + \frac{\varepsilon-1}{2}\left(1 + \frac{10t}{w}\right)^{-\frac{1}{2}} \tag{2.39}$$

Then,

$$C_{ox} = \frac{wl\varepsilon_0\varepsilon_{effox}}{t_{effox}} \tag{2.40}$$

and

$$C_{Si} = \frac{wl\varepsilon_0\varepsilon_{effSi}}{t_{effoxSi}} \tag{2.41}$$

Similarly, to calculate R_{Si}, effective thickness (t_{eff}) and effective conductivity (σ_{eff}) of substrate are needed. As for the capacitance, effective thickness is given by Equation (2.38), and effective conductivity can be obtained from

$$\sigma_{eff} = \sigma\left[\frac{1}{2} + \frac{1}{2}\left(1 + \frac{10t}{w}\right)^{-\frac{1}{2}}\right], \tag{2.42}$$

where $\sigma = \frac{1}{\rho}$ represents the substrate conductivity.

Therefore,

$$R_{Si} = \frac{t_{effSi}}{\sigma_{eff}wl} \tag{2.43}$$

2.3.2.5 Quality factor and resonance frequency for single-π inductor model

The quality factor, discussed in Section 2.3, is the basic characterisation technique for inductors. For the single-π model, Q-factor can be calculated as [79]

$$Q = \frac{\omega L_S}{R_S} \cdot \frac{R_P}{R_P + \left[\left(\frac{\omega L_S}{R_S}\right)^2 + 1\right]R_S} \cdot \left[1 - (C_P + C_S) \cdot \left(\omega^2 L_S + \frac{R_S^{\,2}}{L_S}\right)\right], \qquad (2.44)$$

where

$$R_P = \frac{1}{\omega^2 C_{ox}^{\,2} R_{Si}} + \frac{R_{Si}(C_{ox} + C_{Si})^2}{C_{ox}^{\,2}}, \qquad (2.45)$$

$$C_P = C_{ox} \cdot \frac{1 + \omega^2 (C_{ox} + C_{Si}) C_{Si} R_{Si}^{\,2}}{1 + \omega^2 (C_{ox} + C_{Si})^2 R_{Si}^{\,2}} \qquad (2.46)$$

and $\omega = 2\pi f$. Three different factors can be isolated in Equation (2.44) [28]. The first factor, $F_1 = \omega L_S/R_S$, is the intrinsic (nominal) Q-factor of the overall inductance. Second

factor, $F_2 = \dfrac{R_P}{R_P + \left[(\omega L_S/R_S)^2 + 1\right]R_S}$, models the substrate loss in the semiconducting

silicon substrate. The last factor, $F_3 = 1 - (C_P + C_S) \cdot (\omega^2 L_S + R_S^{\,2}/L_S)$, models the self resonance loss due to total capacitance $C_P + C_S$. This resonant frequency can be isolated by equating the last factor to zero, and solving for ω. This results in the formula for self resonance frequency of the spiral inductor:

$$\omega_o = \sqrt{\frac{1}{L_S \cdot (C_P + C_S)} - \left(\frac{R_S}{L_S}\right)^2} \qquad (2.47)$$

Substrate losses

At low frequencies, the loss of metal line (F_1) restricts the performance of inductors [80]. In high frequency ranges, the loss of substrate (F_2) prevails as the restricting factor. F_2 is greatly dependent on the conductivity of the substrate. As conductivity increases at a fixed frequency, the skin depth of the substrate also increases, leading to an increase of eddy currents in the substrate resulting in a decrease of the Q-factor of the inductor. Heavily doped substrates are usually used in a sub-micron process, with substrate resistivity usually

lying in the range of 10 $\Omega \cdot$cm to 30 $\Omega \cdot$cm. As a result, in the traditional (Bi)CMOS process, the performance of spiral inductors is limited by the substrate. The MEMS processes, mentioned earlier in this thesis, strive to rectify this dependence.

Figure 2.32 shows the analysis of factors F_1, F_2 and F_3 defined in (2.44) for 1 nH and 5 nH sample spiral inductors optimised at different frequencies for their highest quality operation. It can be observed that, although the nominal Q-factor (F_1) increases with frequency, F_2 and F_3 decrease in the same range, resulting in the decrease of the overall Q-factor (Q) at frequencies above 1 GHz.
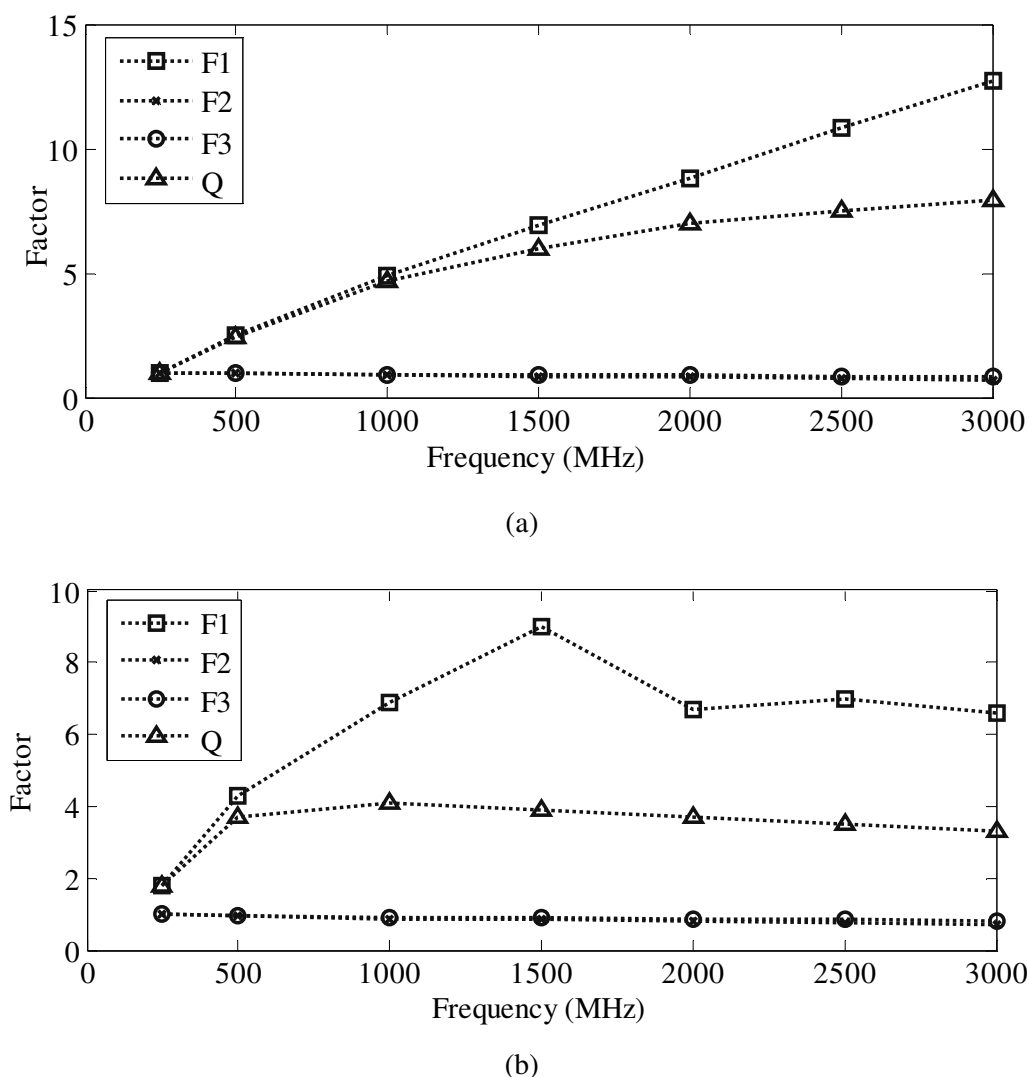

(a)

(b)

Figure 2.32. Analysis of the determining factors of the Q-factor equation for (a) 1 nH inductor and (b) 5 nH inductor.

2.3.2.6 Guidelines for integrating spiral inductors

As detailed in Section 2.3.2, although spiral inductors are a good choice for exclusively on-chip PAs, their use is not as straight forward. They occupy large areas on the chip, suffer from the low quality factors, and are difficult to design for the low tolerance. Capacitors and resistors, on the other hand, do not suffer from these problems.

When designing an integrated capacitor, a designer can simply increase or decrease the area of the component until the required capacitance is obtained. Although capacitance of the parallel plate capacitor does not solely depend on the area of its plates, but also on other factors such as fringing effects, a nearly linear relationship between the two is retained. Similar relationship between the length and total resistance holds for resistors. By modifying the length of a part of the process layer used for fabrication of the resistor, a designer can obtain the wanted value of its resistance. However, this does not apply to the spiral inductors. Contrary to the common sense, one cannot just simply increase the number of turns or width of a single turn to change the inductance. The complicated inductance relationship given in Equation (2.32) or any other can illustrate this interdependency. This complexity of spiral inductor models is one of the reasons why it is a common practice to choose an inductor from a library of inductors supplied with the process rather than to design an inductor for a needed inductance value. If a standard inductor cannot be used in the application of interest, then the iterative process is needed, where one guesses the geometry parameters that could result in the required inductance (and Q-factor) value, calculates inductance, parasitics and Q-factor given the guessed parameters, thereafter repeating this process until one is satisfied with the performance of the inductor. A flow chart of such approach is shown in Figure 2.33.

Reference [81] isolates some general guidelines that can assist in designing a high quality inductor, irrespective of the geometry of the inductor and its model:

1. Where possible, one should use the highest resistivity substrate available. This will reduce the eddy losses that decrease the Q-factor.
2. Placement of inductors should take place on the highest possible metal layers. In this way, substrate parasitics will have a less prominent role because the inductor will be further away from the silicon.
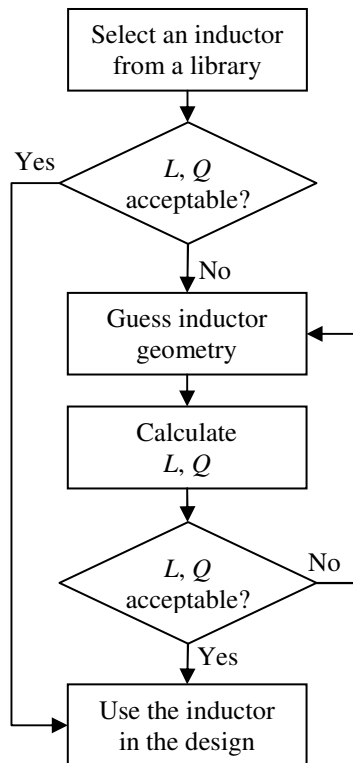
Figure 2.33. A flow chart of conventional spiral inductor design procedure.

3. If necessary, parallel metal layers for the body of the inductor can be used to reduce the sheet resistance.

4. Unconnected metal should be placed at least five turn widths away from inductors. This is another technique that helps reducing eddy current losses.

5. Excessively wide or narrow turn widths should be avoided. Narrow turns have high resistances, and wide turns are vulnerable to current crowding.

6. The narrowest possible spacing between the turns should be maintained. Narrow spacing enhances magnetic coupling between the turns, resulting in higher inductance and Q-factor values.

7. Filling the entire inductor with turns should be avoided. Inner turns are prone to the magnetic field, again resulting in eddy current losses.

8. Placing of unrelated metal plates above or under inductors should be avoided. Ungrounded metal plates will also aid the eddy currents to build up.

9. Placing of junctions beneath the inductor should be avoided. Presence of a junction close to the inductor can produce unwanted coupling of AC signals.

10. Short and narrow inductor leads should be used. The leads will inevitably produce parasitics of their own.

2.4   TECHNOLOGIES FOR POWER AMPLIFIER INTEGRATION

Traditionally, several device technologies have been used for integrating PAs [5, 52, 53]. Initially, gated structures such as MOSFETs, metal-semiconductor field effect transistors (MESFETs), high-electron-mobility-transistors (HEMTs) and pseudomorphic HEMTs (pHEMTs) had found a widespread use. After the introduction of bipolar transistor with a wide-gap emitter, or HBT [82], bipolar transistors emerged as a preferred choice, because of their higher gain and current densities at RF. In the late twentieth century, most of the on chip PA implementations were implemented with HBTs [52]. This resulted in transmitter systems that included at least two ICs in their implementation: a silicon CMOS based front end and a PA fabricated in another technology, which made them bulky and expensive. In the first decade of this century, BiCMOS devices are emerging as an alternative to the two-chip solution because they are able to bridge this integration gap, which as a result reduces the cost of transmitter manufacturing.

Technologies such as gallium-nitride (GaN), aluminium-gallium-nitride (AlGaN), silicon-carbide (SiC), and indium-phosphate (InP), have been struggling for survival in the battle between Si, SiGe and GaAs technologies. The AlGaN/GaN is suitable for both high frequency and high power designs, but together with the SiC technology it has been sidetracked on the road to microwave applications by technical difficulties and high fabrication costs [83]. The InP technology itself does not provide any substantial benefits for the linear PA design [84].

Table 2.5 shows some typical process parameters for the Si BJT, SiGe and GaAs processes [84]. After the inspection of parameters such as cut-off frequency f_T, breakdown voltages and parasitic capacitances, it becomes clear that GaAs presents the best technology for the PA integration. Silicon technology, as expected, offers the worst process parameters. Similar observation is possible by analyzing the PAE, power gain and scattering parameters (*S*-parameters) obtained from the performance measurement study detailed in [84], as summarised in Table 2.6. Measurements were obtained at $f = 1.88$ GHz with power supply of 3.4 V with quiescent current of about 90 mA for 28 dBm output.

Table 2.5. Process parameters for Si BJT, SiGe HBT and GaAs HBT technologies [84].

	Si BJT	SiGe HBT	GaAs HBT
f_T (GHz)	27	44	46
Forward gain, β	100	200	120
Base-emitter voltage, V_{be} (V)	0.8	0.8	1.33
Early voltage, V_A (V)	36	100	1223
Collector-emitter breakdown voltage, V_{ce} (V)	6.2	6	14.3
Collector-base breakdown voltage, V_{cb} (V)	20	12	26
Emitter-base breakdown voltage, V_{eb} (V)	2	5	6.9
Power density (mW/μm$^{2)}$)	2	2	0.9
Thermal conductivity (W/cm-°C)	1.5	1.5	0.49
Base-emitter capacitance, C_{be} (fF)	11	10	2.4
Base-collector capacitance, C_{bc} (fF)	3.6	3.3	1
Possibility of NMOS and PMOS integration	Yes	Yes	No

Table 2.6. Various performance parameters for PAs fabricated in three different technologies (f = 1.88 GHz, V_{CC} =3.4 V and P_{OUT} = 28 dBm) [84].

	Si BJT	SiGe HBT	GaAs HBT
PAE (%)	33.1	35	39.3
Gain (dB)	22.1	21.8	27.1
S_{11} (dB)	-13.8	-16.9	-12.6
S_{12} (dB)	20.2	20.5	27.2

However, it should be noted that the comparison between Si and SiGe technologies is performed from BiCMOS perspective, that is using a BJT in case of pure Si technology, which is not available in the widely used CMOS technologies. Having that in mind and after noting that some SiGe process parameters (e.g. f_T) are comparable to the ones of GaAs technologies, it can be identified that the SiGe BiCMOS performance lies comfortably between the Si CMOS and GaAs HBT technologies and that it is suitable for the PA integration. As such, the SiGe BiCMOS can be considered a CMOS technology with a HBT available for use, allowing for traditional CMOS stages to be incorporated on the same chip, which is, as stated before, considered the main advantage of SiGe. Some additional advantages over the GaAs can also be identified [52]:

- Chip robustness is facilitated by heat-conductive silicon substrate (also seen from Table 2.5),
- Circuit design issues from thermal runaway (Section 2.2.7) are minimized, and
- Reliability at high current densities enables reduction in device size.

Although suitable for experimentation in this thesis, many challenges of high power design on SiGe remain. Nevertheless, new technology trends and ever increasing PA requirements

will always lead to a careful re-evaluation of this powerful technology for commercial use [85].

2.5   RAPID POWER AMPLIFIER DESIGN AND AUTOMATION

The term "rapid power amplifier design" is used to describe a process where many PAs with similar specifications must be designed simultaneously. This section first identifies the need for rapid PA design, followed by the description of several ways to automate this tedious task.

2.5.1   The need for rapid PA design

When wireless transmitters are designed, a set of specifications will normally exist depending on their application. Some of these specifications will directly or indirectly be applicable to the PA itself. These will most certainly include the output power and efficiency described in Sections 2.2.1 and 2.2.3. Additionally, bandwidth and carrier frequency are also of importance and they are dictated by the frequency band in which the transmitter must operate.

Devices with wireless capabilities, such as cordless phones, cellular phones, laptops, personal data assistants (PDAs), as well as many computer peripherals will typically operate in one of the unlicensed frequency bands. Three different ISM bands are available for utilisation in most countries in the world, with transmission channels centred close to 915 MHz, 2.4 GHz and 5 GHz [86]. Two groups of standards are greatly popular for WLANs: Bluetooth and other IEEE 802.11 standards.

Depending on the standard and the frequency band, there will be a number of channels available in each band for use by a transmitting device. For example, for the IEEE 802.11b standard operating at around 2.4 GHz, there are 52 different channels available for transmission, each with a different carrier frequency. If a transmitter is to operate over more than one channel in this band, which will typically arise in situations where two or more of the same devices are present in the same wireless area, then it will either have to contain more than one PA, each optimised for operation in a different channel, or a lone PA will have to be tuneable in some way. In either case, the complexity of the PA system will greatly increase, resulting in a tedious design task.

2.5.2   Automating the rapid PA design

This problem can be resolved by introducing a set of algorithms that will automate a PA design for a given set of specifications. These algorithms should be able to cover the design of several monolithic PA high-efficiency output stages, such as those of the Class E or F (Section 2.2.5). It should also be able to handle the design of an appropriate integrated inductor, as described in Section 2.3. Even in the age of fast personal computers, the speed of execution of the design software still remains important.

There are a number of simple tools addressing some of the given aspects. One of them is a parasitic aware program [22]. It can be used for a PA design optimised for maximum efficiency. However, it only tackles the two-stage Class-E amplifier, with inductor design integrated into the tool. A spiral inductor design tool is described in [26], which can be used to design adequate spiral inductors. These and other available tools can be used in conjunction with commercial synthesis tools, such as MATLAB, Tanner Tools [87] and/or Cadence Virtuoso (see Chapter 3 for description of some of these tools), but to create a really successful design process the designer must be acquainted with many pieces of software instead of only one, which inevitably decreases the speed of design process flow.

Creating a routine that offers comprehensive design coverage, as a part of methodology for design of PAs, therefore, presents a vast research opportunity for this thesis, as stated in the form of a research hypothesis in Section 1.2.

2.6   CONCLUSION

This chapter identified major concepts important for this thesis. After a thorough study of various PA output stages, the switching PA classes (Class E and Class F), were identified as the best choice for integration with wireless transmitters. Although many integrated inductor implementations are possible, the spiral inductor presents the best choice for low-cost integration, removing the need for modification of a standard process and/or outside-chip modelling. The SiGe BiCMOS technology was seen as a very good compromise between the widely popular and inexpensive CMOS on one side, and more expensive GaAs HBT technology, traditionally used in high power design, on the other. As such, it can be used for prototyping the PA devices experimented with in this thesis. Finally, a need for a rapid PA design has been identified, and a method to accomplish this,

incorporated in a software routine that aids PA and spiral inductor design, was proposed to fill in the gap in available knowledge. Chapter 3 follows with the description of the methodology used in this research.