Flexible Power Amplifier Architectures for Spectrum Efficient Wireless Applications

Alessandro Cidronali, Iacopo Magrini and Gianfranco Manes
Department of Electronics and Telecommunications, University of Firenze,
Italy

1. Introduction

The wireless systems evolution known as "beyond the 3rd generation" (B3G) will make use of dynamic spectrum access techniques to provide wide bandwidth to mobile users via heterogeneous wireless networks. A consistent step toward this scenario is represented by the outcome of the last World Radiocommunication Conference [1] which established new primary frequency bands allocation spanning from the UHF band to low microwaves and thus reflecting the increasing demands for broadband mobile and cellular systems.

We have become used to the doubling of processing power of chips based on Moore's law, but the progress in radio interface technologies still poses significant challenges.

High spectrum efficiency performance becomes therefore another major requirement of the design, along with the more consolidated ones: energy efficiency, integration, cost and reliability.

While the IMT-advanced roadmap foresees a 100 Mbps data rate for mobile users and a peak of 1 Gbps is expected for nomadic users, the available spectrum for legacy wireless communications is fragmented and reaches the amount of 750 MHz in the S-C band. A radio technology that is expected to interact with a multi-services network should be able to change between different operating bands and adapt its features according with the different available standard and requirements. Most of the research efforts performed during the last years dealt with issues related to the physical layer of the communication stack [2]; however, despite the growing interest in multi-standard operation, less attention has been devoted to the radio-frequency front-end, which therefore remains one of the most challenging parts of a multi-band radio. One main reason for the delay in effectively implementing multi-standard transceivers can be attributed to the implementation of the RF transmit power amplifier (PA). Today, dedicated, single standard PAs achieve very good power added efficiency (PAE) and, in this way, long battery lifetime. Any multi-standard PA, needed for the support of different, not always predefined, communication systems, should compete with such dedicated solutions. A conceptual framework to this is provided by the so-called software-defined radio (SDR), i.e. a radio communication system, using software for the reconfiguration of the digital and analog parts in order to perform the modulation and demodulation of radio signals, [3]. In practice, however, due to the difficulty of implementing the fast signal processing implied in the SDR approach, most of

the systems on the market, based on more traditional approaches, are still supporting only a very limited number of standards (e.g. 4 GSM frequencies, UMTS and, possibly, Bluetooth). In the near future, further standards will have to be supported, and more could have to be added during the handset lifetime, hopefully without hardware reconfiguration. This will determine the need of multiband PAs capable to transmit efficiently more than one service with variable radio access schemes.

Example of realizations in different technologies are provided in this Chapter as demonstrators of the discussed multiband design methods. The flexibility of the operative frequency is thus introduced by analyzing new PA architectures and design methodologies which consider the inclusion of tunable and switching components to enable a change in the operative frequency. A review of the most promising circuit topologies suitable to design reconfigurable matching networks is given in this Chapter. Varactor diodes based and MOS switched based topologies are compared, highlighting their point of strength and weakness. It is shown as a concurrent dual-band PA implemented by the proper combination of frequency-dedicated PAs, each of them optimized to work in a given bandwidth would be an easy approach, it becomes unsuited due to the complexity of the power combiners. For this reason the true concurrent dual band PA presented in this chapter is to be considered as an enabler components for high efficiency multiband systems.

Spectrum efficiency is just one of the challenges a wireless system designer faces, further come from linearity and energy efficiency resulting from the use of multicarrier and complex envelope modulation schemes. As the spectrum efficiency increases a more demanding requirement in term of PA linearity faces to the designers. Energy efficiency and linearity are conventionally traded-off considering that increasing the power back-off increases the linearity at the expenses of lower energy efficiency. To maintain signal integrity, the resulting waveforms in turn require linear transmission paths for their successful deployment. A way to match signal integrity and energy efficiency consists in the use of digital predistortion algorithm applied at base-band and implemented in the digital section of the transmitter. In spite of their large development in frequency dedicated PA architectures, the development of a technique suitable for multi-band applications is not yet completely available. In this Chapter a comprehensive treatment a novel technique for Dual Band Digital Predistortion (DB-DP) is discussed. The DB-PD is based on the simultaneous predistortion of both channels at intermediate frequency (IF), it uses a single band memory polynomial DP for linearization, while the feedback path is based on a subsampling receiver. The memory polynomial DB-DP system is presented by simulation with Matlab-Simulink® for a deep understanding of performance.

2. A possible applicative scenario for multi-band transmitters

Extending the scenario to already experienced 3G voice/data systems, users may be moving while simultaneously operating in a broadband data access or multimedia streaming session. The need to support low latency and low packet loss handovers of data streams as users transition from one access point to another may require the concurrent use of more than one frequency band at the time. For full-mobile data services, no user interaction will be required to adapt their service expectations because of environmental limitations that are technically challenging but not directly relevant to the user (such as being stationary or moving). The enabling front-end of future mobile unit thus will accommodate more than

one system in a effective and efficient way to make possible the connectivity capabilities depicted in Fig. 1.
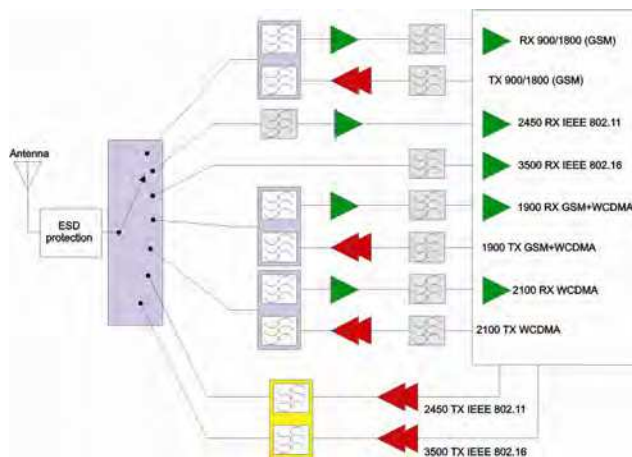

Fig. 1. Concept of a multi-band transmitter

The Wireless Local Area Network (WLAN) industry has become one of the fastest growing segments of the communications industry. This growth is due, in large part, to the introduction of standards-based WLAN products, regulated by the IEEE 802.11. The expectation of the WLAN's continuing growth stems from the promise of new standardized WLAN technologies, from improved cost/performance of WLAN systems, and from the growing availability of WLAN solutions that consolidate voice, data, and mobility functions. This, combined with market forecasts reporting that WLAN will experience a continuous growth in the next years, show that WLAN technologies will play a significant role in the future and will have a significant impact on our business and personal life styles. The WiMAX is an alternative and complementing standard for high data rate transmission, which will transform the world of mobile broadband by enabling the cost-effective deployment of metropolitan area networks based on the IEEE 802.16 standard to support notebook PC and mobile users on move. There are many advantages of systems based on 802.16, e.g. the ability to provide service even in areas that are difficult for wired infrastructure to reach and the ability to overcome the physical limitations of traditional wired infrastructure. The standard will offer wireless connectivity of up to 30 miles. The major capabilities of the standard are its widespread reach, which can be used to set up a metropolitan area network, and its data capacity of 75 Mbps. This high-speed wireless broadband technology promises to open new, economically viable market opportunities for operators, wireless Internet service providers and equipment manufacturers. The flexibility of wireless technology, combined with high throughput, scalability and long-range features of the IEEE 802.16 standard helps to fill the broadband coverage gaps and reach millions of new residential and business customers worldwide.

With WLAN 802.11 and now WiMAX 802.16, there has been a growing interest in technologies that allow delivery of higher data rates over large geographical areas. The IEEE 802.16 family of standards (802.16-2004 and 802.16e) are intended to provide high

bandwidth wireless voice and data for residential and enterprise use. The modulation used to achieve these high data rates is orthogonal frequency-division multiplexing (OFDM). WiMAX OFDM features a minimum of 256 subcarriers up to 2048 subcarriers, each modulated with either BPSK, QPSK, 16 QAM or 64 QAM modulation. Having these carriers orthogonal to each other minimizes self-interference. This standard also supports different signal bandwidths, from 1.25 MHz to 20 MHz to facilitate transmission over longer ranges and to accommodate different multipath environments. This represents a significant increase in system profile complexity as compared to the 802.11 standard, mostly to guarantee a wider, more efficient, more robust network. More subcarriers and variable-length guard intervals contribute to this enhancement.

The ability to develop and manufacture a single reconfigurable terminal, which can be configured at the final stage of manufacture to tailor it to a particular market, clearly presents immense benefits to equipment manufacturers. With the design, components used, and hardware manufacturing processes all being identical for all terminals worldwide, the economy of scale would be huge. This has the potential to offset the additional hardware costs which would be inevitable in the realisation of such a generic device.

Based on this, the scenario adopted reflects in the request for transceiver architectures capable to support cellular phone, WLAN and WiMAX in an 'always and everywhere connected' solution. The transceiver performance in this multi-standard operation, however, comes at the expense of RF specifications that are more difficult to achieve. Furthermore, the choice and definition of the proper transceiver architecture becomes a difficult task, since several parameters - as now imposed by two standards - must be taken into account.

3. Suitable architectures for multiband-multimode transmitters

The concept of a multiband or general coverage terminal is, strictly speaking, an extension of the basic SDR concept into that of a broadband flexible architecture radio, since the basic reconfigurability and adaptability aspects of operation do not depend upon multiband coverage. It would be possible, for example, to construct a useful SDR which operated in the 800-900 MHz area of spectrum and which could adapt between AMPS, GSM, DAMPS, PDC, and CDMA. It is now normal, however, for a handset to have multi-frequency operation and hence the extension of this principle to a SDR is a natural one. The international business traveler market is still seen as both large and lucrative, particularly in terms of call charges, hence making this type of handset attractive to both manufacturers and network providers. An ideal SDR is shown in Fig. 2; note that the A/D converter is assumed to have a built-in anti-alias filter and that the D/A is assumed to have a built-in reconstruction filter.

The ideal software defined radio has the following features [4]:
- The radio access scheme (i.e. modulation scheme, channelization, coding) and equalization for transmitter and receiver are all determined in software within the digital processing subsystem. This is shown containing a DSP in Fig. 2
- The ideal circulator is used to separate the transmit and receive path signals, without the usual frequency restrictions placed upon this function when using filter-based solutions (e.g., a conventional diplexer). This component relies on ideal matching between itself and the antenna and power amplifier impedances and so is unrealistic in practice over a broad frequency band. Since the primary

alternative, a diplexer, is very much a frequency-dedicated component, its elimination is a key element in a multiband or even multimode transceiver.

- The linear, or linearised, PA ensures an ideal transfer of the RF modulation from the DAC to a high-power signal suitable for transmission, with ideally no adjacent channel emissions. Note that this function could also be provided by an RF synthesis technique, in which case the DAC and power amplifier functions would effectively be combined into a single high-power RF synthesis block.

- Anti-alias and reconstruction filtering is clearly required in this architecture (not shown in Fig. 2

- . It should, however, be relatively straightforward to implement, assuming that the ADC and DAC have sampling rates of many gigahertz. Current transmit, receive, and duplex filtering can achieve excellent roll-off rates in both handportable and (especially) base-station designs. The main change would be in transforming them from bandpass (where relevant) to lowpass designs.
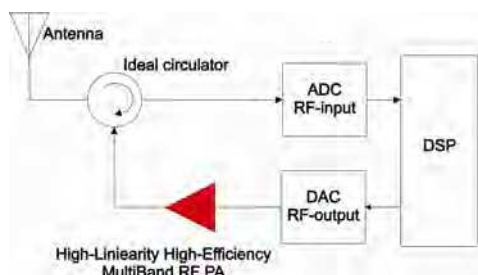

Fig. 2. Ideal software defined radio architecture

Possibly the most important element of any SDR system, whether in a base station or handset, is the linear or linearised multiband transmitter. Receiver systems have always required a high degree of linearity, as they must possess a good signal handling capability, in addition to good low-noise performance. In the case of transmitters, however, a high degree of linearity is a relatively recent requirement, arising predominantly from the widespread adoption of multi symbols envelope-varying digital modulations.

This follows from the fact that most modern modulation formats incorporate some degree of envelope variation, the only significant exception at present being GSM and its derivatives (DCS and PCS). The basic architecture of a SDR transmitter revolves around the creation of a baseband version of the desired RF spectrum, followed by a linear path translating that spectrum to a high-power RF signal.

Nevertheless the implementation of a true SDR poses a further very critical issues, i.e. the power consumption of the analogue-digital converter. Let's consider for instance the use of a flash converter, largely available in the market with a maximum number of bit about 18 preceded by a sample and hold circuit. Carrying out a simplified calculation, given the converter dynamic range, Dc, the power consumption of this systems is:

$$P_{dc} \frac{kT}{t_s} 10^{Dc/10}, \qquad (1)$$

where the k is the Boltzmann's constant = 1.38x10^{-23} J/K, T is the device temperature and t_s the sampling time. Furthermore the dynamic range of the converter is given by:

$$D_c = 6.02N + 1.76 - PAR + 10\log_{10}\big[2OSR\big], \tag{2}$$

where number of bit, N, a peak-average ratio for the signal, PAR, and an oversampling ratio, OSR. From this easy calculation we can straightforwardly estimate the AD power consumption Pdc in a significant scenario for SDR. Assuming to digitize a frequency band from 800 MHz to 5.5 GHz with a 11GS/s ADC and assuming that the receiver dynamic range is from -20 dBm to -120 dBm, with a SNR of 12 dB at minimum sensitivity, the average PAR of 4, the required N is 20; it results that a such ADC consumes hundred of watt, thus preventing the use of the ideal architecture in Fig. 2. in practical implementation.

4. Reconfigurable Matching Networks

The multiband-multimode demands of today's wireless market, is fulfilled by implementations based on parallel line-ups completed by antenna diplexers and switches to meet the specific requirements of each communication standard, (c.f. Fig. 1). Utilizing only one adaptive transmit path to replace the parallel path concept is conceptually simple, but practical design considerations place severe design constraints and technology. Major challenges consists in creating the tunable filters and PAs [5]. Addressing these challenges means to develop flexible PAs capable to maintain the power-added efficiency (PAE) and linearity while moving among different operating frequencies. In conventional PA implementations, the linearity requirement typically results in the use of class-AB operation for the output, which provides a workable compromise between linearity and efficiency. When considering linearity, the class-AB output stage must be dimensioned in such a way that it can provide its peak output power without saturation. As a result, for a given peak output power and battery voltage, the load impedance for a class-AB stage at the fundamental frequency is fixed to $R_L \approx 0.5 \cdot V_{cc}^2 / P_{Peak}$ .Unfortunately, class-AB operation provides its highest efficiency only under maximum drive conditions. When operated at the required back-off level, due to linearity reasons for a given communication standard, a rather dramatic loss in efficiency occurs. For these reasons improving amplifier efficiency, while maintaining linearity, is currently a major research topic in wireless communications. In linearity-focused researches, the circuit is designed so that the resulting overall linearity performance of the PA module is improved. In this way, the active device can be operated closer to its peak-power capabilities and still be able to meet the linearity requirements.
Techniques that address the efficiency in the back-off mode are dynamic biasing or regulation of the supply voltage of the output stage

[6]. Dynamic biasing provides only modest improvements in efficiency, and supply voltage regulation requires an efficient DC-to-DC conversion, increasing system cost and complexity and operative bandwidth. Nevertheless this techniques appear very promising for future transmitter architectures. An alternative for improved class-AB efficiency is load-line adjustment as a function of output power using an adaptive or reconfigurable output matching network.
An ideal Reconfigurable Matching Network has to provide:

- Low Loss
- High linearity
- High Tuning Speed
- Sufficient impedance coverage
- Low complexity
- Low area usage

Power handling of matching networks is a critical issue in PA applications. To reduce the losses in a matching network, the use of a limited number of reactive elements is mandatory, beside the choice of high Q tunable components. Typically, such a network is based on varactor diodes, PIN-diodes or FET switching of matching elements like inductors, transmission-lines or capacitors, also involving micro electromechanical systems to improve the power handling capability [7].

We can conclude that these integrated adaptive networks will play an important role for the realization of the next generation of adaptive transceivers and this paragraph is aimed to describe the ongoing basic researches on this subject.

4.1 Varactor based switching matching network

Varactor diodes, although characterized by a relatively low Q factor at microwave frequencies, can be a choice for enabling RF tuning. Unfortunately, because of their inherently non linear behavior, their use with modern communication standards (characterized by high peak-to-average power ratios), has to be carefully analyzed according to the specific case considered. In Fig. 3 are shown varactor diode based circuit topologies [5] suited to provide matching tuning overcoming the issue related to the linearity of the electron devices.
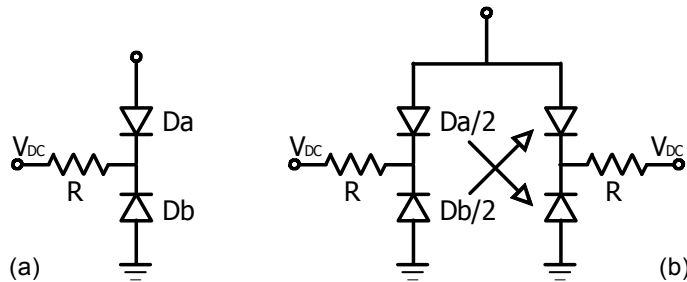

Fig. 3 – (a) Anti-series connection of varactor diodes to minimize third-order distortion, (b) Anti-series/anti-parallel connection of varactor diodes to minimize second and third-order distortion.

Basically, the capacitance of a single varactor diode can usually be expressed as:

$$C(V) = \frac{K}{(\varphi + V)^n} \tag{3}$$

where φ is the built-in potential of the diode, V is the applied voltage, n is the power law exponent of the diode capacitance, and K is the capacitance constant. The power law exponent can exhibit wide variation in different situations, from a value of n≈0.3 for an

implanted junction to n≈0.5 for a uniformly doped junction to n≈1.5/2 for a hyper-abrupt junction.

If the applied dc voltage is V_{DC}, then the incremental capacitance of a single varactor diode as a function of the incremental voltage v can be expressed as

$$C(v) = C_0 + C_1 v + C_2 v^2 + ...$$ (4)

where the term C_1 gives rise to second-order distortion and the C_2 term gives rise to third-order distortion. The diode configuration in Fig. 3a can be employed to realize a voltage variable capacitor with theoretically no distortion. Indeed, (referring to the circuit in Fig. 3a) evaluating the expressions for the linear and nonlinear terms of the capacitance, and setting to s the ratio of the diode areas Db/Da, it follows that:

$$C_0 = \frac{sK}{(1+s)(\varphi + V_{DC})^n}$$

$$C_1 = \frac{(1-s)nC_0}{(1+s)(\varphi + V_{DC})}$$ (5)

$$C_2 = \frac{C_0\left[(s^2+1)(n+1) - s(4n+1)\right]}{2(\varphi + V_{DC})^2(1+s)^2}$$

It can be noted that for n ≥ 0.5, C_2 can be made equal to zero, resulting in zero third-order distortion, by setting

$$s = \frac{4n+1+\sqrt{12n^2-3}}{2(n+1)}$$ (6)

It can be observed from eq. (6), that the particular case of constant doping profile in the diode (the abrupt junction case where n=0.5) results in a value of s=1. This case is particularly attractive because this set of conditions (n=0.5, s=1), sets both C_2 and C_1 equal to zero. A more elaborate analysis shows that all higher order distortion terms also vanish, yielding (in theory) a "distortion-free" operation for this unique case. When dealing with process technologies where n>0.5, eq. (6) provides a direct means of calculating the required diode area ratio to minimize C_2.

For example, in the case where n=1, the required area ratio is exactly two. In the case of n=2, which corresponds to the ideal hyper-abrupt junction, the required area ratio is 2.6.

Although this approach can minimize C_2, it is clear from eq. (5), that a value of s≠1 will result in a finite value of C_1. In this case, a relatively high third-order distortion product will unfortunately still arise, resulting from the secondary mixing of the fundamental with the second-order non-linearity C_1. Fortunately, this distortion contribution can be eliminated, by placing an identical varactor stack in anti-parallel configuration (see Fig. 3b).

The linear capacitance of the circuits of Fig. 3a and b are identical, but the circuit of Fig. 3b has C_1=C_2=0 when the proper area ratio is set. It must be underlined that, in this configuration (Fig. 3b), all the even-order coefficients are zero (C_1, C_3, C_5, …) in this topology, but the higher coefficients that create odd-order distortion (C_4, C_6, C_8, …) are not

zero, although the IM3 contributions due to the 5th and higher order nonlinearities are very small.

The implications of this analysis can be summarized as follows:

- The classical configuration of Fig. 3a provides theoretically a "distortion-free" varactor stack when n=0.5, corresponding to a uniform doping profile of the varactors.
- The more generalized configuration of Fig. 3b provides an ultra-low distortion varactor stack for any value n>0.5, by setting the proper ratio of the diode areas, which sets C_1 and C_2 to zero, providing more freedom for use in different process technologies.

It must be underlined that each of the circuits in Fig. 3 requires a very high tap impedance (R) for proper operation. A high tap impedance, limits the impact that forward biasing of one of the diodes by RF signal has on linearity.

An effective way to implement the high impedance while keeping the RC time limited for the control signal is the anti-parallel diode pair depicted in Fig. 4.
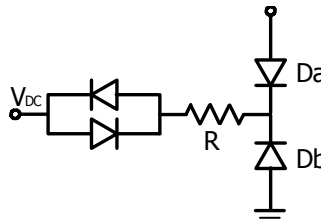

Fig. 4.– Anti-series connection of varactor diodes with modified center tap.

Analyzing, by a two tone test the described matching and comparing them with a single varactor diode we can observe the linearity of the different topologies. The IM2 and IM3 distortion of the circuit in Fig. 3a are comparably low, while the IM3 of the circuits in Fig. 3b is limited by fifth-order distortion due to complete cancellation of the third-order products. In this case indeed, a 1:5 slope dependence for the IM3 components can been found [8], confirming the elimination of the C_1 and C_2 contributions.
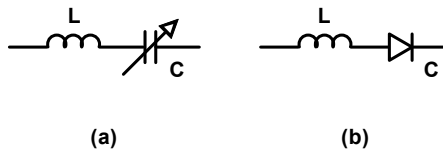

(a)   **(b)**

Fig. 5 – (a) Equivalent circuit of the "tunable inductor", (b) Simplified circuit of a varactor-based "tunable inductor".

A different option for the implementation of reconfigurable matching network is to act directly on the inductance value, rather than the value of the capacitors as discussed hereinabove. For this purpose varactor diodes can be exploited to create a "tunable inductor" [9]. By considering the circuits in Fig. 5, the equivalent impedance Z of the LC circuits can be written as:

$$Z = \frac{1 - \left(\omega/\omega_c\right)^2}{j\omega C} \tag{7}$$

where $\omega_c = \dfrac{1}{\sqrt{LC}}$ is the resonant frequency of the circuit. When the condition $\omega > \omega_c$ is achieved, the impedance Z represents an equivalent inductor L_{eq}, that, for the simplified circuits in Fig. 5, is analytically expressed as:

$$L_{eq} = \frac{\left(\dfrac{\omega}{\omega_c}\right)^2 - 1}{\omega C} \qquad (8)$$

The resulting equivalent inductor is a "tunable inductor" whose value is related to the varactor capacitance C (and also ω_c). In Fig. 5b, the varactor parasitic elements have not been represented, however they have to be taken into account during the design of the "tunable inductor".

4.2 Switching matching networks

If the output power of the PA is so high to make the use of varactor diodes prohibitive, a possible solution could be to replace the varactor diodes with fixed value capacitors controlled by switching PIN diodes. By observing to the examples in Fig. 6, if the PIN diode connected to the capacitor is in the ON state, the capacitor adds its own capacity to the global circuit. On the other hand, if the PIN diode is in the OFF state, the capacitor does not affect the global circuit.
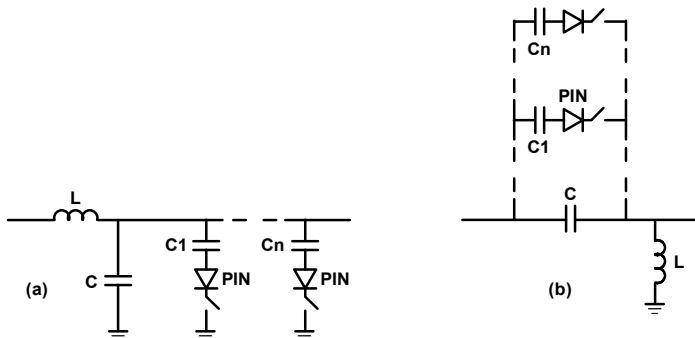

Fig. 6.– (a) Low-Pass Matching Network example, (b) High-Pass Matching Network Example

By using a set of digital signals that change the state of the PIN diodes to active or inverse condition, it is possible to generate $N = 2^n$ impedances (where n is the number of PIN diodes used). Unfortunately, such a technique permits only a discrete set of tuning possibilities, resulting inappropriate where fine tuning is requested, unless a considerable number of components is used, thus increasing losses, area usage and costs. A different option of switched matching network is based on MOS devices . The use of MOS devices rather than PIN diodes makes them more suitable for IC designs. The operative principle of the network is discussed by a following example.

Let's consider a Si-Ge HBT device, the optimum fundamental loads at one dB compression point are 15 Ω at 2.45 GHz and 5 Ω and 3.5 GHz respectively, the reactive parts are

negligible. These resistance values have been determinate accounting for the output power level to be supplied by the power amplifier in the two different bands. The aim of the output network consists in synthesizing these loads using a MOS switching network topology, schematically depicted in Fig. 7 and investigating the general features.
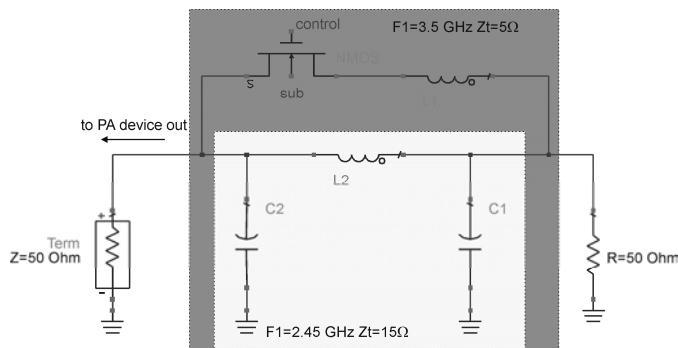

Fig. 7. MOS based switched matching network π topology

The basic simple network topology is based on a π-structure, with an additional branch composed by a NMOS device acting as a switch, with in series an inductor to change the network impedance when the switch is closed. The behavior of the network depends on the switch condition:

1. when the switch is ON (i.e. its R_{ds} value is low, zero in the ideal case) the network should present at its input an impedance value equal to 15 Ω that is the optimum value for the device at 2.45 GHz.

2. When the switch is OFF (i.e. its R_{ds} value is high, infinite in the ideal case) of course, the frequency behavior of the network change. In particular the equivalent value of the inductor between the two shunt capacitors becomes equal to the parallel of L_1 and L_2 and the input impedance of the network decreases to 5 Ω at 3.5 GHz.

Moving on the actual schematic, for the switch it has been necessary to introduce the biasing network in order to guarantee the right switching functionalities. The simulations have been performed biasing the NMOS device, with dc voltages on drain and source set to V_D=3V and V_S=3V respectively, to ensure that the NMOS device is properly biased in the origin of its output characteristics. The gate control voltage is raised to V_G=5V, when the NMOS switch has to realize a short circuit condition, and dropped to V_G=2V when the NMOS switch has to realize an open circuit condition. The feed lines for all three terminals have been realized using 4.8 KΩ resistor for each lines to guarantee the request isolation. Fig. **8** shows the small signal parameters S11 and S21 of the networks, as a function of frequency from 1 GHz to 4 GHz, respectively when the switch is OFF, e.g. when the network has to synthesize the load at 3.5 GHz, and when the switch is ON, e.g. when has to be synthesized the load at 2.45 GHz.

In particular in Fig. 8 left, the blue line in the right part of the figure represents the real part of the input impedance, which can be note is roughly equal to 15 ohm. Unfortunately, the input impedance shows a residual imaginary part due to the non ideal behavior of the inductor and capacitor elements present in the network. Similarly, Fig. 8 right reports the same features when the switch is OFF, e.g. when the network has to realize the load

required at 3.5GHz. Also in this case, the network exhibits the requested real part of the input impedance, with a residual imaginary part of the input impedance.
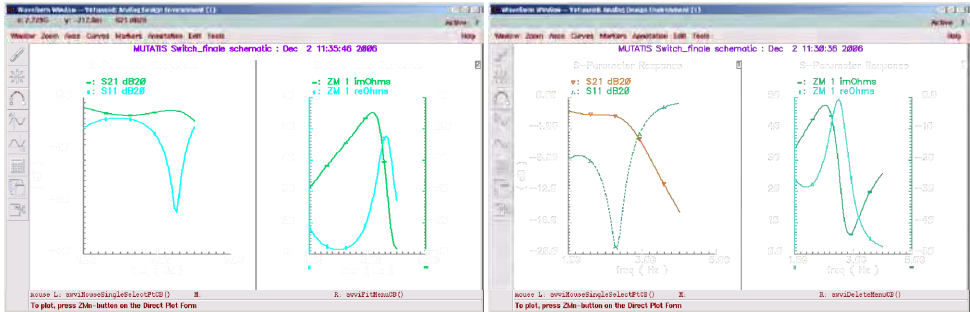

Fig. 8. S11 and S21 frequency behavior together with the real and imaginary parts of the input impedance when the switch is ON (left) and OFF (right).

5. Dual-Band reconfigurable SiGe HBT amplifier design

The above described matching network can been adopted to implement the two matching networks for a class-AB amplifier. In the following example of such design approach is presented. The design of the reconfigurable PA is based on a power device composed of 17x8 elementary SiGe HBT with and emitter area of 8.49 μm^2. The bias circuitry of the power section is designed to provide other than the required base bias current, a circuit-level linearization. Let's start introducing this latter part, referencing to the Fig. 9. The size of the devices used for the bias circuitry and the values of passive components scale accordingly, so that accurate (i.e. matched) current mirroring can take place. In addition, resistor R4 introduces base ballasting and its value is selected to be 350 Ω per emitter. It helps in reducing the risks due to thermal runaway. A larger value would have better effect, however, that could result in high DC voltage drop at high power drive, and thus early gain compression of T4, the power transistor, would take place.
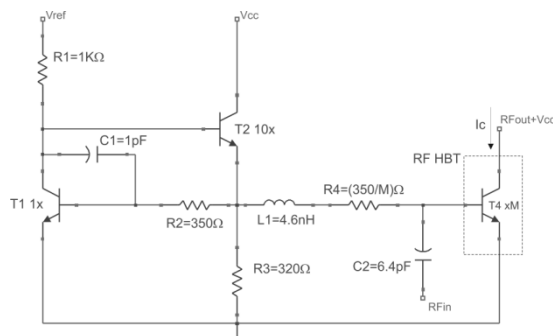

Fig. 9. PA core schematic including the power transistor and the bias network

The power device, which is shown in Fig. 9 within the dotted outline, is biased through the current mirror. By inspection of this figure the current through the reference transistor T1 is:

$$I_{REF} = \left(V_{REF} - 2V_{be}\right) \Big/ R1 \quad , \tag{9}$$

with Vbe ~ 0.8 V. If the ratio between the reference device and the RF device is M, in this case M = 17x8, then the current that will flow through the RF transistor is M*I$_{ref}$. in addition, the current mirror behaves like an ideal voltage source, since its output impedance is

$$R_{out} \approx \frac{1}{g_m A_{loop}} , \tag{10}$$

where g_m is the transconductance of the emitter follower device, and A$_{loop}$ is the loop gain of the loop formed by the reference device and the emitter follower (buffer) transistor. At high frequencies, R$_{out}$ becomes inductive and its value should be kept low throughout the frequency band of interest and could be further reduced if resistor R$_3$ was further reduced in value, at the expenses of an higher current trough this resistor.

At 1 dB compression point, the power transistor should be biased at a collector current which has a value near the optimum value that guarantees that the maximum transition frequency (f$_t$) of the transistor is achieved. Due to self bias effect, which forces the DC average collector current to increase with increasing input power, the value of the quiescent bias current is chosen to be much lower compared to the value that it reaches under full power drive conditions. So, when the transistor T4 is self biased it starts to draw more current. Since the current flowing through R$_2$ is constant, as the base-emitter voltage of T$_1$ is constant, this additional current is supplied by transistor T$_2$. The capacitor C$_1$ in the bias network helps to stabilize the loop formed by the reference transistor, T$_1$, and the emitter follower transistor, T$_2$. The stability analysis of the loop formed by T$_1$ and T$_2$ shown that for unity loop gain the phase margin is approximately 78 degrees, which guarantees the loop unconditional stability. Fig. 10 reports the DC current flowing in the collector of the power device, T4, as a function of the input power at the frequency of 2.45 GHz. From the curve is seen that the quiescent current increases according with the description given above and reaches the value related to the peak f$_T$, at the 108 mA for an input power of 6 dBm. The value of the bias current for low input power is defined according with the eq. (9), adjusting the value of either R$_1$ or V$_{ref}$, in this case to the R$_1$ = 1 kΩ, corresponds a V$_{ref}$ = 3.3 V.
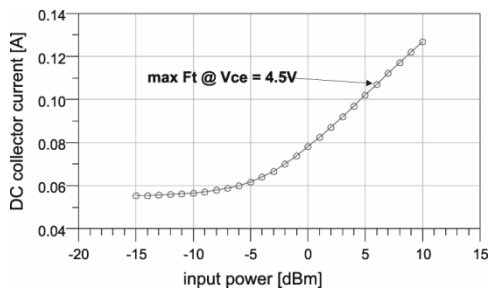

Fig. 10. Behavior of the collector current for the power transistor, T4 in Fig. 9, as a function of the input power at 2.45 GHz.

The layout of the reconfigurable PA is illustrated in Fig. 11 (left), where are clearly visible the inductors adopted for the matching and the area for the active part of circuit. The total size of the layout with the bonding pads is 1x1mm. While in operation, the voltage in the node between the inductors swings between positive and negative values. This causes the control MOS drain-bulk np junction to be forwarded biased, which degrades switch performance and may results in latchup. In order to overcome the specific drawback, we set an offset voltage (pin Vofs) which shifts the voltage in the abovementioned node in positive values.
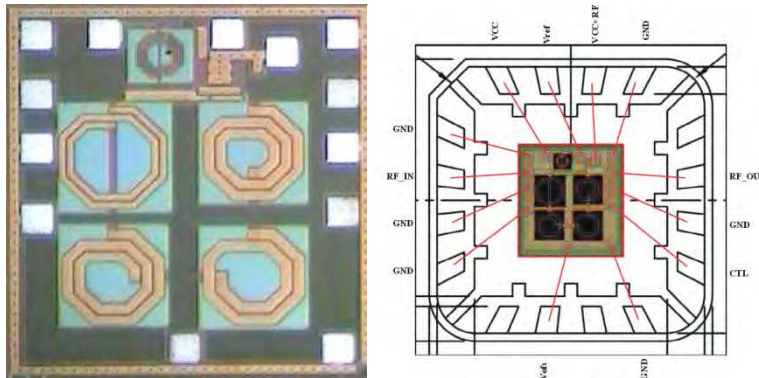

Fig. 11. Layout of the reconfigurable SiGe–PA prototype (left) and its packed wiring diagram (right)

For the design of the inductive part of the matching networks a planar electromagnetic simulator should be used, this permits to calculate any mutual inductances, between the inductors that are in proximity along with the losses in the silicon. Due to the large number of controls and dc-supply for the operation of the components a package is normally required and has to be taken into account during the design. The selected solution consists in the QFN package which allows up to 16 leads. The bonding diagram for this component is reported in the Fig. 11 (right), along with the pin description. At the time of this report editing, the packaged components were not yet available, for the reason the measurement results that will follow consider only the on-wafer device.
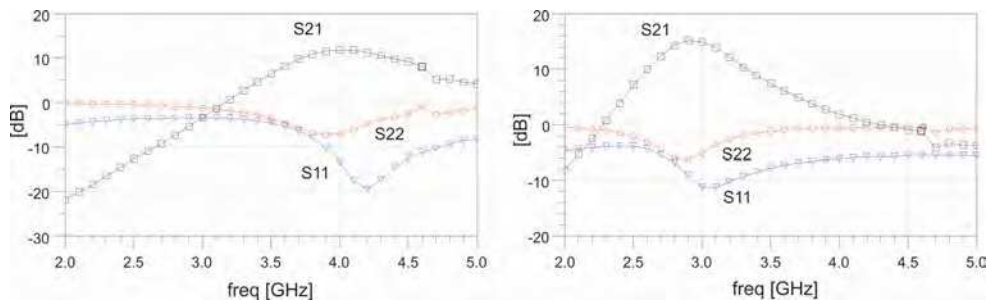

Fig. 12. S-parameters in the state corresponding to the lower and higher bands.

The first test of measured data considers s-parameter in the range 2 to 5 GHz. The S21-S11 and S22 in dB for the state corresponding to the lower and higher frequency band are reported in Fig. 12. From the figures is clearly observed the shift in frequency due to the above mentioned reasons, which is estimated in the range of 500 MHz. While the small-signal gain maximum is 2.5 dB lower than the simulated in the lower frequency band and 1 dB less in the higher band. This reduced matching have to be attributed to the matching which, being lower than estimated for the discussed reasons, introduces and matching loss consistent with the reduced gain observed during the characterization phase. We exclude problems related to the measurement set and calibration, although a problem related to measurements, the ringing in the 4.7GHz, was observed.

The characterization of the sample in the large-signal regime is reported in the below figures. It is carried out at the frequencies where the device exhibits the maximum gain and matching in the two states that are respectively at 2.9GHz and 3.9GHz, which correspond to 500MHz frequency shift from the design target as discussed above. The data related to the large signal gain up to the compression are reported in the Fig. 13, respectively for the state corresponding to the lower and higher frequency. In this figure the CW single tone signal is applied in the two states and the Pout-Pin curves are recorded. The value of the gain is consistent with the small-signal gain while the compression point is estimated to be about 16 dBm for the lower band and 15 dBm for the higher frequency. These figures are 1 dB lower that estimated during the simulations. All this data are consistent with the simulation and again supports the hypothesis that introducing the additional inductive parts due to the packaging the proper frequency behavior can be reached.

An intermodulation product characterization considering a two-tone signal with a center frequency at the selected frequencies of 2.9GHz and 3.9GHz and 1MHz offset, was applied to the device. The input power was swept from -10dBm to 13dBm. The results of the characterization, in terms of the higher IM3 are reported in the Fig. 13. It is observed that any consistent change in slope is observed in the traces. This allows concluding that the MOS involved in the switched matching networks doesn't introduce any additional nonlinearities. In fact, if the additional nonlinearity required changing the state of the device, was effectively excited we would have observed an addition component in the IM3. It is also worth to observe that the input power dealt by the device is below the threshold for which this effects become evident. This threshold from simulation is estimated in the range of 28-30 dBm.
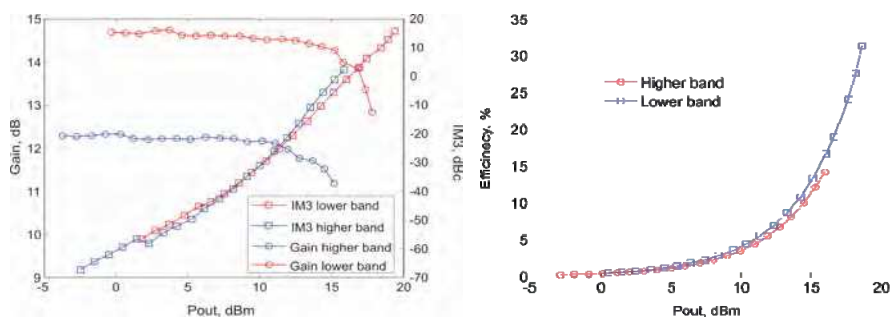

Fig. 13. single tone large-signal gain at 2.9GHz and 3.9GHz and two-tone large-signal inter-modulation offset frequency 1MHz, data reported for the higher IM3 (left), PAE

6. Dual-band power amplifier architectures

The main objective of this paragraph is about the consistent and quantitative evaluation of a two possible architectures of dual-band PA both suitable for their involvement in the concurrent dual-band systems, [12]. The first is based on two dedicated PAs combined by a frequency diplexer while the second is specifically designed to be operated in dual-band state. For the sake of the comparison the operative frequency are defined as 1.98 and 3.42 GHz respectively suitable for WCDMA and OFDM radio access technologies.

The two dual-band architectures considered in this paragraph are based on the schematic representations reported in Fig. 14. In the first (Fig. 14, left) the PA is implemented by making use of two dedicated PAs combined by a frequency diplexer. This latter device has to be designed to combine the two PAs introducing band-pass and band-stop behavior in each of the two branches. This is implement by the most innovative technique and technology and it still represents a very critical part of the entire PA structure. Indeed, this component must guarantee an almost lossless behavior in the two transmission paths and as much as possible isolation between them; without sacrificing the matching. In particular the transmission loss characteristic is required to preserve the combined efficiency of the entire structure, while the isolation is a required feature to avoid the cross-modulation between the two dedicated PAs. The constraints on the diplexer become more critical in the case of closer operative spectrum bands. During the two dedicated PAs design, the eventual combination with the diplexer implies a specific additional PA design consideration related to an accurate evaluation of the out-of band termination, which might degrade the output power and efficiency of the two units. The treatment of the harmonic termination of the due to the diplexer is out of the scope for the present treatment.
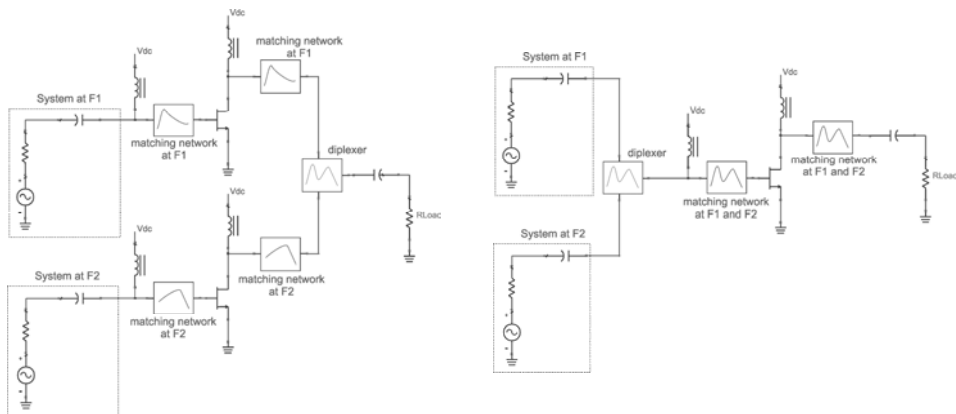

Fig. 14. Schematic of the concurrent dual-band PA implemented by two combined dedicated PAs (left), and by a dual-band PA (right).

Either the dedicated frequency PAs and concurrent PA topology consist of conventional class-AB PA designs, where the tuned matching networks can be synthesized by either passive and\or distributed elements properly dimensioned, without external tuning controls. For the concurrent PA (Fig. 14, right), the two signal sources are combined prior to

be applied to the PA input. The design method of concurrent dual-band PA based on multi-tuned networks composed of lumped elements is discussed in [10].

The investigation carried out in this chapter relies on prototypes designed and fabricated using low cost off the shelf active devices along with discrete SMD passive components assembled on FR4 0.8 mm thick evaluation printed circuit board designed with microstrip technology. For this specific test several electron device technology can be considered, eider bipolar or FET fabricated using several different material, spanning from Si to GaAs and possibly GaN. In the present case we choose as active device a GaAs FET, namely the ATF50189 from AVAGO Technologies, a medium power enhanced mode p-HEMT with a cut off frequency of 6 GHz and a 1-dB compression point of 29 dBm at 2 GHz. Optimum bias point for efficiency, linearity and gain can be fund either from manufacturer specifications or CAD simulations, on the basis of nonlinear model analysis. For the specific device possible bias point is found to be 4.5V drain supply voltage with a corresponding quiescent current of 200 mA. The chosen bias point drives the ATF50189 transistor in the AB class operation. The design was based on load and source pull simulations carried out at the two fundamental frequencies of 1.98 and 3.42 GHz, adopting a nonlinear device model which included the package parasitic. Simulations provided saturated output power of 28 dBm and 26 dBm respectively in the lower and higher frequency bands with a power added efficiency of approximately 40% and 35% at the 1-dB gain compression point. The resulting load and source constant power contours are shown respectively in Fig. 15.
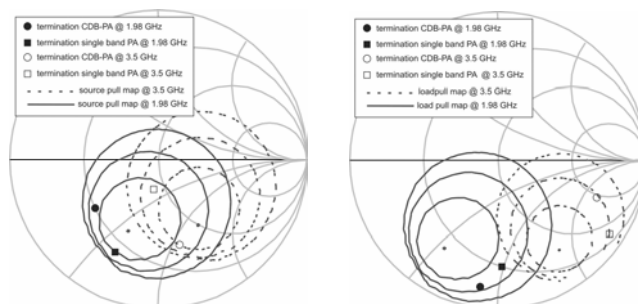

Fig. 15. Simulated source-pull (left) load-pull (right)contours at 1.98 GHz and 3.42GHz, 1 dB steps, and terminations at fundamentals for the single band and dual-band prototypes

The implementation of the source and load terminations defined by the source- and load-pull analysis was obtained by using lumped elements matching networks. This technique, by employing a different approach with respect to standard microstrip technology, enabled the achievement of highly compact prototypes. All the designed PAs adopt the same general topology for the input and output matching networks. Different nominal values and the absence of some of the components determine the difference between the prototypes. In addition, the input network accommodates a stabilizing network which has bee implemented by all three prototypes. The presence of shunt capacitors at both the gate and drain terminals, provide a short circuit to the second harmonic.

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- ➢ HTML (Free /Available to everyone)

- ➢ PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)

- ➢ Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below