RF Transmitter Architectures for Nomadic Multi-radio: 
A Review of the Evolution Towards Fully Digital Solutions
Antoine Diet, Fabien Robert, Geneviève Baudoin, Martine Villegas, Philippe Cathelin, Pascal Triaire, Ajib Bahi

To cite this version:

HAL Id: hal-00854740
https://hal-supelec.archives-ouvertes.fr/hal-00854740
Submitted on 28 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
RF Transmitter Architectures for Nomadic Multi-radio: A Review of the Evolution Towards Fully Digital Solutions

Antoine Diet*1, Fabien Robert2, Geneviève Baudoin3, Martine Villegas3, Philippe Cathelin4, Pascal Traïre3 and Ajib Bahi3

1L2S-DRÉ. UMR8506. Supélec, F-91192 Gif S/ Yvette, France; 2OMMIC SAS, Limeil-Brevannes, France; 3Université Paris-Est, ESYCOM, EA2552. ESIEE-Paris, Noisy-Le-Grand 93162, France; 4ST-Microelectronics. Crolles, France; 5ST-ERICSSON SAS. Grenoble, France

Received: July 08, 2013; Accepted: June 27, 2013; Revised: June 27, 2013

Abstract: In this paper, a review of the major evolution steps of mobile transmitter architectures is summarized. We propose a classification and discuss about improvements and evolutions towards the full-digital solutions. This discussion is mandatory while the cognitive-radio concept implies the design of new structures for multi-radio front-ends. Technological process improvements, for example CMOS for high frequency, opens new possible solutions thanks to digitally based buildings blocks and functions. The conception of fully digital architectures is discussed in that idea. Our work in this domain is presented as a contribution in the study of transmitter architecture for multi-radio. Two ways are explored: (1) improvements of analog and digitized architectures, and (2) digital architectures. We present, in these two fields, a state of the art of published works and patents [US20130051440A1], [US20130084816A1], [US20120320957A1]. The Multi-radio context drives us to redefine the important figures of merit for the architecture design, as seen in recent publications and patents such as [US20130049854A1], [US2013009710A1], [US8432219], [EP2541781A1]. The traditional efficiency/linearity trade-off has to be considered with a mandatory frequency flexibility and power control ability, defined here as “power scale-ability”. These constraints made the conception of fully digital architectures attractive, thanks to the flexibility of digital systems. The first part of this paper will present the challenges of designing transmitter architectures for cellular and wireless local/extended area networks (WLAN and WiMAX for example), under the hypothesis of nomadic communication system. The design method in that context is consequently different from classical architecture design. In the second part, a classification of the RF transmitter architectures and their evolution are proposed. Evolution from linearization to linear architectures and introduction of digitized functions are presented, in that order. Solutions are appreciated differently, considering the multiple tradeoffs on efficiency, linearity, flexibility and complexity, while taking into account integration and realization (process).

Keywords: RF architecture, linearisation, fully digital architecture, EER, ET, LINC.

1. THE MULTI-RADIO TRANSMITTER : CONTEXT AND ARCHITECTURE DESIGN CHALLENGES

Wireless communications mean mobility and connectivity for users. Nowadays, the interest in multimedia broadcasting and other high data rate transfer implies an evolution of wireless communication systems. Most of the Radio-Frequency (RF) standards for mobile communications and Wireless Local/Personal Area Networks (WL/PAN) are located in the frequency range from 700 MHz to 6 GHz, see Table 1. Currently, the user demand is to provide a high data rate transparent connection over these different standards. Also, the multi-radio transceiver evolution drives it to the ability to receive or to be compliant with (and not to interfere with) connectivity applications such as Bluetooth or wireless broadcasting systems (DVB-H, Digital Video Broadcasting - Handheld). The transmitter has to limit its spectral emissions in order to respect its standard requirements and to guarantee the receivers immunity (same equipment). To efficiently use these standards, RF transceivers have to adapt to the RF characteristics mentioned above, and to respect the network medium (air interface) access. This method can be Time or Frequency Division Duplexing (TDD/FDD), or CSMA/CA (Carrier Sense Medium Access / Collision Avoidance). This has a strong impact on the front-end components, for example duplexer selectivity, switch losses and isolation.

The use of the spectrum and the user’s habits is the starting point of the radio evolution. We can reasonably suppose that the spectrum sharing is subject to numerous access methods and that scenario should concern simultaneous connectivity and possible opportunistic radio access. For example, someone can be supposed to use GPS (Global Positioning System) or WLAN internet connection with a cellular phone and, additionally, with a Bluetooth earphone. This example is extended to the possibility of keeping/switching the mobile phone services via different wireless links, maybe the WLAN connection. Following that idea, the architecture is also supposed to be able to duplex these communications. The challenging complexity of this

*Address correspondence to this author at the L2S-DRÉ. UMR8506. Supélec, F-91192 Gif S/ Yvette, France; Tel/Fax: +33 (1) 69 85 17 12 / 64; E-mail: antoine.diet@lss.supelec.fr
The question is an aspect of the Cognitive Radio (CR). CR is a concept where the communication system is able to sense the environment (electromagnetically), analyze it and determine a strategy for transmitting and receiving data, and reconfigure itself (upon the acceptance of network providers) dynamically with respect to the PHY (Physical Link) and MAC (Medium Access Control) protocols of the different possible networks. This is not a multi-standard approach in the sense where the RF characteristics of the transmission can evolve and be combined, following a strategy of resources sharing. Under the term “multi-radio” we will consider a system where the RF part (RF front-end) is supposed to be dynamically parameterized. The RF transmitter architecture design, for multi-radio, has to deal with a mandatory problem: how is it possible to provide such “RF flexibility” for a wireless transceiver? We focus, herein, on the design of mobile transmitters, implying the battery lifetime and efficiency to be carefully considered without lowering the other figures of merit. Such a transmitter should also be aware of standards coexistence.

Key functions of a conventional RF transmitter are: (i) baseband processing (modulation scheme, coding) and digital to analog conversion (DAC), (ii) high frequency transposition (needing a frequency synthesizer), (iii) power amplification and (iv) Tx/Rx multiplexing (duplexer for FDD / filter and switch for TDD) before the radiation by the antenna, and an emission filter). All these functions are subject to many sources of imperfections [1-6]. DAC resolution adds some noise and distortion on the signal and any baseband or RF bloc causes bandwidth limitation. The purity of the frequency synthesizer is challenging because its phase noise is transposed on the emitted signal and cannot be filtered. One of the major imperfections is the non-linear response of the transmitter, dominated by the Power Amplifier (PA) response, which can cause damaging spectrum and signal degradation [7-11] in the case of high Peak to Average Power Ratio (PAPR) signals. The effects of these imperfections can be quantified by different well-known criteria such as EVM (Error Vector Magnitude) and ACPR (Adjacent Channel Power Ratio). EVM expresses the signal quality/understandability of the message and ACPR represents the amount of interference caused by a system (emitter with imperfections, mainly non-linearities, NL). This is to be compared with the maximum allowed out-of band spectral re-growths (coexistence and channel interference) and is often completed by the spectrum emission mask, specified by the standards. Other technological criteria are to be considered for designing integrated transceiver: consumption, size, system complexity, integrability [12-14]. In order to design RF transmitter architectures for nomadic multi-radio applications, performance must be preserved regarding the different criteria, for the different “RF configurations”. The main qualities needed for the design of such architectures are discussed here: efficiency, linearity, power scale-ability, complexity, flexibility. All of these criteria are to be optimized, whatever the output power, carrier frequency, modulation schemes and channel bandwidth are.

“Power scale-ability”: Connectivity and mobile wireless networks have to manage their resources for performance optimization. Phenomena such as near/far effect, multi-paths, fading or multi-user interferences drive it to power allocation strategies for each transmitter. In order to improve the effective network capacity, the average RMS power of the signal is dynamically set at different level, following an optimized law of variation. Power control implies a power scale-ability of the front-end. The emission at maximum power level is far from being the best solution. Consequently, it is very important to be able to vary the RMS average power of the transmitted signal. Some strategies need to dynamically characterize the channel statistical behavior for determining the mean power of the signal. This implies that the architecture figures of merit have to be maximized for max, medium and low levels. In Fig. (1), we illustrate this by shifting of the different signal input power distribution in two different PA cases, optimized for maximum power (profile n°1) for average RMS efficiency (profile n°2). We compare three signals cases: low dynamic (case A), high PAPR with Gaus-
sian like signals (case B) and high PAPR with Rayleigh like signals (case C).

**Efficiency:** The total consumption of a transceiver is difficult to evaluate because it depends on the baseband part current consumption, influenced by the modulation scheme and symbol frequency. For wideband signal (WB) such as WCDMA (Wideband Code Division Multiple Access) and high data rate OFDM (Orthogonal Frequency Division Multiple Access) signals, the RF power amplification part of the

![Graph showing efficiency and power distribution](image)

**Fig. (1).** Resulting peak and average efficiencies for two types of PA characteristics and in three cases of signals. (Blue) Colors indicate increasing power of the input signal.
transceiver highly contributes in the total current consumption but has to deal with the additional consumption due to an important baseband processing (wideband). This is more crippling than GSM transceivers (NB) although emitted powers are lower. However, if we evaluate each element separately, the PA has the highest consumption and its optimization is critical for the battery lifetime because it directly impacts the overall consumption at a fixed output power. The efficiency is often evaluated at peak power emission (drain/collector efficiency or Power Added Efficiency, PAE) because it represents a maximum current consumption. In fact, the statistical distribution of the modulated signal strongly impacts the PA behavior (large signal), and so its efficiency over the time. That is why we emphasize to evaluate the average efficiency over the signal statistic rather than the peak efficiency. This is true for drain and power added efficiency. The distribution of the signal is partly characterized by the PAPR (Peak to Average Power Ratio). This needs an analysis of the architecture/PA efficiency for the different radio-communication signals considered. Fig. (1) shows that the distribution of the signal (ex. Rayleigh for case C, and Gaussian for cases A and B) results in a different average efficiency for two types of PA. These two types illustrate different PA optimization. 70% is an arbitrary value that represents the state of the art performances of switched PA at these frequencies. Fig. (1) reveals that the signal power distribution implies an average efficiency that can be maximized if an accurate choice is made between type n=1 or type n=2 PA. The variation of the efficiency with the power, as explained for power scale-ability criteria, is a penalty in our multi-radio approach. Following this point of view, PA type n=2 is preferable for high PAPR signals, corresponding to high data rate signals.

**Linearity:** The quality of the transmitted and received signal is measured basically by the EVM (and ACPR for the Tx), under the hypothesis of ideal receiver without imperfections and noise. Modulation schemes such as OFDM and W-CDMA have different statistical variations (power distribution, as shown in Fig. (1)) and Peak to Average Power Ratio. Due to the PAPR, the architecture is subject to a large power variation, close to its physical limit, and its response presents unavoidable non-linearities. Also, memory effects are increased by this high power variation close to the component physical limits [2, 15]. The influence of the PAPR on the non-linearity can be reduced by modifying the average power of the signal, or by the possible clipping on the signal: minimum or maximum limitation. By clipping, the PAPR is reduced but the EVM increases (non-linear operation). This modification is useful but distorts the information. Some better techniques can be used for PAPR compression, but at the expense of complexity and sometimes frequency bandwidth. Linearity requirement for above-mentioned standards is quantified in terms of EVM values. In fact, EVM only reveals the noise and unwanted signals in-band added on the modulation scheme symbols. As linear imperfections can be reduced by equalization, NL effects are mainly responsible for the EVM values. The higher the EVM, the higher the possibility to mistook symbols (causing errors). High data rates of communication employ M-QAM modulation scheme which is very sensitive to this noise (in function of M). The EVM limit is in the range of 3% (OFDM, 64-QAM for WiMAX) to 17% (QPSK, for UMTS). These low EVM values imply to carefully limit some imperfections like synthesizer phase noise, LO leakage and IQ modulators imbalances [16], and also to linearize the whole Tx line up (power amplification part included). A useful calibration of the transceiver contributes to reduce the imperfection effects. Generally, PAPR compression and linearization techniques are used jointly to optimize both efficiency and linearity. Other approaches consist of building an architecture adapted to the signal characteristic, i.e. a linear architecture such as EER (Envelope Elimination and Restoration [17]), ET (Envelope Tracking) or LINC (Linearization with Non-linear Components [18]). Until now, classical solutions were following two ways: designing efficient transmitter with linearization technique or designing linear transmitter with efficiency improvement technique.

**Complexity:** RF multi-radio transmitters are designed for integrated technologies. Increasing the number of components implies new imperfections, consumption and size increase, and EMC (Electro-Magnetic Compatibility) potential problems. Linearization techniques and additional functions or filtering blocs are often necessary for guarantying the integrity of the signal and spectrum emission requirements. The added complexity is a challenge for the RF architect because it drives him to a more difficult design (process and packaging). In some correction techniques or linear architectures, the signal bandwidth is widened by non-linear functions (root, modulus, phase...) and/or a feedback loop that may introduce a stability-bandswidth tradeoff.

**Flexibility:** Multi-radio targets the RF flexibility of transmitters. The ability to operate at different frequencies means that linearity, efficiency, power scale-ability and complexity have to be satisfied at each frequency band of operation. There are mainly three approaches: (i) Designing a very wideband front-end with such qualities, which is currently quite-impossible, (ii) Parallelizing different blocks, which represents a very high increase in complexity, size, consumption and losses due to additional connections, and (iii) Parameterizing an efficient linearized front-end.

Whatever approach chosen, the flexibility can be attained at the expense of sub-optimal performance, making the architecture suitable for multi-radio applications [19]. While efficiency and linearity are always the first highlighted tradeoff in a transceiver design, the criterion of frequency flexibility is now considered much important. We discussed the key point of such criterion evolution in [22].

Another important quality that must be added, but has to be considered separately: all of the communication systems are evolving toward miniaturization and adaptability. This can be facilitated thanks to electronic integration and digitization of the different functions. Transceiver architectures will be more flexible if we can digitize a part of it. The integrability depends on the number of components, the size, complexity and sensitivity to the process of each functional bloc. It is very attractive to simplify the architecture design process thanks to the integration of the different functions. These qualities are summarized in Fig. (2). Next part will present a classification of RF transmitters’ architectures from basics towards linearized, linear based, digitalized and digital ones. We define these criteria in order to define a methodol-
ogy for the design of architecture. Once the different requirements are clarified, the choice of the architecture structure can be made. The classification presented in the following part is done by identifying the different improvements brought by the different architecture types.

2. ARCHITECTURE CURRENT CLASSIFICATION

RF transmitters architecture are historically based on analog narrow band transmission [2-5]. The evolution towards IQ digital modulation schemes and wideband signals favored news structures based on well-known homodyne, heterodyne and/or modulated Phase Lock Loop (PLL) [2, 5]. The need of linearization implied by wideband and high PAPR signals drives it to linearized structures or linear based ones. Linear architectures are named that way when the information, and so the structure of the transmitter, are decomposed and recombined in order to cause the lowest NL effects (at the expense of complexity). We consider that RF architectures are classified in three major types: (i) analog linearized or linear architectures, (ii) digitization of some functions in linearized or linear architectures and (iii) fully digital based architectures, employing some new “digital–RF” functions. Moreover, the complex envelope signal, representing the information, can be treated in its polar or cartesian form.

Fig. (3) summarizes the possible classification of such architectures. As it can be seen in this figure, linearization and linear based architecture can be completed by special optimization techniques, often dedicated to efficiency improvement. This section will define and describe the different linearization techniques (Feed-Back, Feed-Forward, Predistortion) and linear architectures (EER, LINC, Envelope Tracking…) before comparison and discussion about efficiency improvement and digitization. Fully digital architectures will be described in a dedicated sub-section.

A) Basic architectures, linearization and linear architectures. Basic RF architectures were traditionally dedicated to narrowband FM transceivers. Block functions realized the transposition to high frequency in one step (homodyne structure) or two steps (heterodyne), involving an intermediate and selective frequency part (IF). The advantage of the homodyne structure was its simplicity and the possibility to use only one frequency synthesizer as a local oscillator signal (LO). By comparison, heterodyne structures suffer from added phase noise due to the presence of two synthesizers. However, the problem of RF coupling between Antenna, PA and LO signal is sometimes crippling in homodyne transmitter due to un-efficient shielding, particularly true in GMSK where PA is saturated. Many improvements were made in that way, and the advantages offered by integrated electronics made the homodyne structure more popular for the future of RF transceivers, where the number of components to integrate is minimized.

In order to reduce the impact of phase noise, the possibility of modulating the PLL, used for frequency synthesis, was carefully investigated. Modulated loop architectures are the base of various proposed solutions [23, 24]. Depending on the bandwidth of the information, the PLL can be modulated by the reference signal, by modifying the correction loop or directly by modulating the VCO command signal (Voltage Controlled Oscillator). Bandwidth enhancement of modulated PLL is an interesting task for RF transceivers because it drives it to the potential use of modulated PLL for high data rate communications. One major challenge is to transmit...
amplitude variations in such structures because modulated PLL is not reproducing AM information. The polar “lite”, as seen in Fig. (4), was one of the solutions proposed for GSM/EDGE. This structure uses a modulated PLL and so no AM information is supposed to be transmitted (see the cons of Fig. (4)). In the polar “lite” solution for GSM/EDGE, the envelope variation is reintroduced by a Voltage Gain controlled Amplifier, at the RF frequency [24] as modified in Fig. (4).

Using these basic architectures for wideband communications, which are often synonyms of higher PAPR signals (OFDM, W-CDMA...), the non-linearities (NL) introduced by the PA made the linearization of the transmitter mandatory [25]. There are mainly three solutions for linearizing the transmitters: (1) Avoiding the NL effects by backing off the power. This solution is simple but degrades the PA effi-

---

**Fig. (4).** Transmitters’ architecture classification.

**Fig. (4).** Modulated PLL architecture, and Polar Lite principle.
ciency and corresponds to an over-sizing (in power, and consequently in size) of the transmitter, compromising for integrated circuits. (2) Correcting the NL effects before or after their occurrence on the signal. It mainly exists Feed-Back (FB, polar or cartesian loop reactions), Feed-Forward (FF) and Pre-Distortion (PD). Transmitter’s architectures where the NL effects are corrected are called linearized architectures. (3) Modifying the NL effects and the architecture by using (vectorial) decomposition and recombination processes. This implies an important modification of the architecture following the decomposition function. The idea is to treat the signal on different forms (and number) in order to amplify it equivalently, with lower NL generated at the output. These structures are called linear architectures.

The common point in linearized architectures is to modify the modulated signal in function of the PA behavior. Sometimes, a selected NL model of the PA is useful (Volterra series, Wiener or Saleh model…) but increases the complexity of the correction. Adaptivity to the signal (amplitude and/or phase) can be introduced in order to compensate for the model’s lack of accuracy and the PA memory effects (a temperature influence can be considered).

Each structure presents drawbacks and interests. The Feed-Back can be performed on the amplitude (Polar FB) or on I and Q quadrature components of the signal (Cartesian FB), and both are dedicated to narrowband signals (the bandwidth can be improved for dedicated structures). Moreover, FB reduces the gain of the amplification and introduces a bandwidth limitation due to the transfer function of the loop (stability and dynamic response). The principle of FF is to reproduce NL components (without the information) and subtract them from the NL amplified signal before the emission. Feed-Forward increases significantly the signal processing and RF blocks complexity in the transmitter. Also, the hypothesis of a precise synchronization and accuracy between the two paths (NLs and reconstructed transfer functions) are required. The improvement in linearity will be costly in terms of consumption and size. The advantages are stability and the possibility to process potentially wideband signals. The most interesting of the three techniques is digital pre-distortion because of its flexibility: the anticipation can be done in the digital part and, by doing so, can provide adaptability of the technique if using a feed-back path (with an additional ADC). The signal is widened in frequency because of the non-linear law of the pre-distorter (as for NL effects of compression on the spectrum), requiring baseband and RF parts to be wideband designed. Because of this increase of the bandwidth, the digital pre-distortion leads to an increase in the signal sampling frequency and, therefore, increases the working speed and power consumption of the digital signal processing (DSP). If a feedback path is used for adaptability, the power consumption of the necessary ADC in the feedback path may be non-negligible. Needed performance of this ADC drives it to a high cost increase. In order to reduce the computation load, the digital pre-distortion function is generally implemented by look-up tables [26, 15] or a DSP processing.

Linear architectures are based on a vectorial decomposition of the signal and the amplification of the created components at RF frequencies before their recombination [17]. These techniques are used when high efficiency is needed and if NL effects are so important that it is mandatory to preserve the linearity at the same time. Avoiding NL is often achieved by driving high efficiency switched mode RF PAs with constant envelope (constant power) signals at their input. We can consider the problem of linearization in the communication chain from the digital part to the antenna (front end). This drives one to completely modify the architecture and its elements, specifications in baseband, IF/RF and power RF. After the amplification of constant envelope parts of the signals, the challenge is to reintroduce the variable envelope information with lower NLs than in a direct amplification case, keeping high efficiency of the architecture. Basic examples of these techniques are the LINC and EER, and their recent optimizations by carefully designing the RF blocks such as the RF HPA (switched or not) [1, 4, 21, 22, 27-31].

The LINC principle relies on a decomposition of the modulated signal into two constant envelope signals, see Fig. (5). The decomposition can be computed by a Digital Signal Processor (DSP) or by combining two VCOs in quadrature PLL configuration: CALLUM. This latter configuration is an interesting architecture but presents the possibility of instability and additional manufacturing costs. The amplification of the two constant envelope signals implies the design of two identical High Power Amplifiers (HPAs) at RF frequency, and this often causes signal distortion due to imbalance mismatches. Also the HPA must be wideband because the signal decomposition is a non-linear process (widening of the spectrum), and the phase modulation index is increased. Whatever the decomposition technique is (LINC/CALLUM), the problem is that the efficiency is directly determined by the recombination process: a sum of the powers. It is very difficult to avoid losses at RF while designing a broadband RF power combiner.

Another decomposition technique was proposed by Kahn in 1952 [17] and is basically an amplitude and phase separation technique (polar, see Fig. (5)): Envelope Elimination and Restoration (EER). This method was first proposed for AM signals. The advantage of EER is that it drives the RF PA with a constant envelope modulated signal (carrying the phase information), enabling the use of a switched (SW) high efficiency amplifier [7, 10, 32, 33, 34]. The difficulty is to reintroduce the amplitude information linearly using the variations of the PA voltage supply. This implies a power amplification of the envelope signal at a frequency of the order of the symbol rate (lower than RF). The recombination can be done with a SW class PA because the output voltage is linearly dependent on the supply voltage for this PA mode. Two difficulties are to be considered in such a linearization technique: (i) synchronization between the phase and the amplitude information [4, 16] particularly for larger BW (like LTE 20 MHz) and (ii) linear and efficient amplification of the amplitude before the recombination (directly impacting the overall efficiency). Envelope Tracking (ET) is different from EER in the sense that the PA driving signal is not only the phase information but the complete RF modulated signal. The envelope is carried by the driving signal and tracked by the supply voltage at the same time. This theoretically reduces the sensibility on time mismatch and the bandwidth at the input of the PA. Additionally, the calibration in
ET is less complicated than in EER. The low level distortions are minimized compared to EER, although there are numerous points of comparison between EER and ET. The PA should be able to amplify the AM information at its input, and it is supposed that the consumption is following the supply voltage value. For high PAPR signals, a clipping can be necessary to increase the mean power level driving the PA, requiring pre-shaping of the envelope signal. Both EER and ET will not work for weak signals (need to control PA drain voltage – or whatever – with a “close to zero” voltage linearly). Such transmitters must also be able to switch to a more traditional mode of operation below a certain output power, which is, by far, easier with ET. Table 2 summarizes some major points of comparison concerning above mentioned linearization techniques and linear architectures. The improvement in efficiency is often high in linear architecture but relies on non-linear functions that widen the signal bandwidth and increase the complexity of the system, once a difficult synchronization is done, requiring calibrations. If the complexity can be partly transferred in the digital domain, this will improve the flexibility of the transmitter and enable adaptability (PD can be adaptive Digital PD). Another important aspect for today’s application is the frequency flexibility, especially for multi-radio transmitter. Often, this quality depends on the structure/topology and class of the PA.

In polar architectures, another challenge concerning these signals is the amplification of the amplitude information. In short, the amplitude can be restored by supply modulation of the RF PA or being coded and restored, in a constant power

![Fig. (5). EER and LINC principles.](image)

Table 2. Comparison of Linearization Techniques and Linear Architecture

<table>
<thead>
<tr>
<th></th>
<th>Complexity of the Architecture</th>
<th>Efficiency Improvement Key Factors</th>
<th>Main Limitation and Imperfections Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB</td>
<td>Feedback path (polar or cartesian)</td>
<td>-</td>
<td>- BW limitation due to the loop</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Stability criterion</td>
</tr>
<tr>
<td>FF</td>
<td>Addition of components</td>
<td>-</td>
<td>Consumption of additional components</td>
</tr>
<tr>
<td>(D-)PD</td>
<td>Depends on the algorithm implantation</td>
<td>-</td>
<td>Complexity of the algorithm</td>
</tr>
<tr>
<td>LINC (CALL.)</td>
<td>high</td>
<td>- PA classes</td>
<td>- Synchronization of the paths</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Combiner (losses)</td>
<td>- Gain and phase imbalance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Combiner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- BW enhancement due to NL mod.</td>
</tr>
<tr>
<td>EER</td>
<td>high</td>
<td>- PA classes VS PAPR</td>
<td>- Synchronization (envelope restoration)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Signal statistical variation</td>
<td>- BW enhancement due to NL mod.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Supply mod. or envelope coding efficiency</td>
</tr>
<tr>
<td>ET</td>
<td>high</td>
<td>- PA classes VS PAPR</td>
<td>- BW enhancement due to NL mod. (less than EER)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Signal statistical variation</td>
<td>- AM driving of the PA (classes of operation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Supply mod. or envelope coding efficiency</td>
</tr>
</tbody>
</table>
form (pulsed PWM or Sigma-Delta, $\Sigma\Delta$), before driving the PA. In the first case, the amplification efficiency before supply modulation directly impacts the RF PA efficiency. In the second case, a mandatory filtering has to be done (thanks to the PA if possible) because it lowers the useful power at the PA outputs. This is due to the fact that PWM or Sigma Delta coding generates very important spectral re-growth. However, an amplified pulse coded signal cannot be emitted due to the standards spectrum mask requirements.

A lot of work has been done on the EER and ET based architectures, [35, 36]. These two types are often named as “polar” ones [1, 21, 23, 27, 31, 34, 37, 38] but this is abusive for ET due to the fact that the modulated signal is driving the PA. In fact, the term “polar” is used for characterizing the AM information modulating the RF signal. The generation of the amplitude and phase components can be expected to be done digitally thanks to the power of DSPs (or FPGA, ASIC,…). The bandwidths of the envelope and phase signals are widened due to NL processing and make it necessary to design the circuit for three to four times the symbol rate or even more (as for LINC or any other NL decomposition) [16]. Fortunately, a clipping in frequency and on the envelope amplitude is possible, increasing the EVM and ACPR under acceptable levels. These polar architectures are suited for new high data rate standards where efficiency of the transmitter and linearization is mandatory. Also, the multi-standard and multi-radio concepts have helped polar based architectures to evolve in multiple ways by asking for the need of transmitting the AM information. For example, the recombination is possible at the input of the PA thanks to an envelope coded signal which is constant envelope (PWM or $\Sigma\Delta$). This coded signal can modulate the phase signal (RF) and the AM information can be restored by the band-pass filtering function of the following cascaded blocks: PA, transmission filter and antenna. The emitted spectrum is the criterion of quality to be considered carefully, because the PWM or $\Sigma\Delta$ envelope coding is the source of useless and crippling spectral re-growth [38]. The efficiency is also penalized by the power amplification of such frequency components. Moreover, this is counter-balanced by the advantages of high flexibility of this architecture.

Current needs in terms of transmitter linearity and efficiency for high data rate applications have caused the RF designer to consider RF architectures based on combined techniques. For example, digital pre-distortion is generally used in polar implementation to mitigate the impacts (and constraints) of recombination circuitry. This review of linearization techniques reveals that the different parts of an efficient and linear transmitter cannot be designed separately: baseband, frequency transposition and power amplification (and antenna for wideband systems). The modification of the architecture for global performance improvements must be done, considering the impact of each block, digital or analog, and their imperfections. To conclude this sub-section, linearized architectures are mandatory for the major part of actual and future systems, which are wideband and high data rate. The highest performance can be reached if a combination of different techniques is exploited: pre-distortion and EER or ET seems to be the most popular.

Fig. (6) presents an evolution of EER/ET architectures for today’s Multi-Radio transmitter based on the above consideration. The structure in Fig. (6) is based on a polar architecture where the phase signal can be transposed by a modulated loop (this is drawn but can be realized by homodyne structure) and drive a switched high efficiency parameterized PA. The driving signal can be modulated in order to generate different types of pulses thanks to Pulse Width Modulator (PWM) or Sigma Delta ($\Sigma\Delta$) modulators. These modulators can be designed digitally and combined with the DAC for a possible envelope restoration. The envelope signal can also be present on the collector/drain current, driving to a (partial) amplitude restoration directly on the RF output signal. The amplitude information and the mean power of the signal are combined dynamically and restored by different recombination types such as supply modulation or multiplication of the driving signal (coded or not). In order to respect the multi-radio criterion of flexibility, the switch PA has to be modi-

![Fig. (6). Possible Multi-Radio EER/ET transmitter.](image-url)
fied for frequency flexibility (for example with switched capacitor…). As it is pointed out in Fig. (6), the frequency flexibility is an important property for the functional bloc composed by the PA, the filter/duplexer and the antenna. If the signal is coded, a pass-band function has to be realized (pass-band profile) for the different carrier frequencies. The idea of a parameterized PA is based on the simulations of [19] where the potentials flexibility and amplitude scale-ability (power) are investigated. The main figures of merit are the “pass-band profile” of the functional bloc: “PA+Filter+Antenna” and the achieved efficiency at the different carrier frequencies, and for modulated signals (PAPR and statistical power distributions). As it can be seen in this figure, the management and coding of the envelope are the subject of different scenarii whether the envelope is (partially) coded or not and if the coding chosen is PWM, ΣΔ or another scheme. The Linearity is managed by the PAPR re-partition between the supply modulation and the input driving signal. This signal can be AM modulated and/or PWM/ΣΔ coded if the impacts on the architecture performance (mainly efficiency and spectral re-growths) are not damaging. As switched PA performance depends on the impedance seen by the transistor, dedicated antennas for such architecture are supposed to be multi-band and co-designed with the PA because they directly influence the wideband response of the loading network.

The architecture presented in Fig. (6) potentially fulfills the multiple tradeoffs involved in the design of reconfigurable front-end for multi-radio. Frequency flexibility is provided by a tuning ability of a highly efficient PA and the amplitude restoration (structure is polar based) process is subject to an optimal recombination following the modulated signal PAPR and statistical properties (Gaussian, Rayleigh… see Fig. (1)). Although these optimizations are currently studied, for example in [22], a preliminary pre-distortion is helpful, considering the standard requirements.

B) Improvement and digitization. There exist numerous improvements of analog based architecture (linearized and linear). The improvement goals/criteria can be the robustness, the NL reduction by coding and the efficiency maximization. Designing mobile transceivers using OFDM and multi-antenna MIMO and smart antenna techniques enables to achieve very high performance (throughput) in mobile channels with large delay spreads. Many new standards use OFDM and MIMO approaches [37]. Examples include: IEEE 802.11n for wireless local area networks, IEEE 802.16d and IEEE 802.16e fixed and mobile WIMAX or IEEE 802.20 for wireless metropolitan area networks, IEEE 802.22 for wireless regional networks, 3GPP LTE (Long Term Evolution) for beyond 3G cellular networks. OFDM and MIMO have also generated new challenges in terms of transmitter architectures with good efficiency and linearity, and in terms of integration of MIMO transceivers and antennas in mobile user terminals. MIMO technology uses multiple antennas at the transmitter and the receiver. The obtained diversity and spatial multiplexing allow for better BER (Bit Error Rate) and increased data rate or link range without increasing the bandwidth or the transmitted power. In comparison with a SISO system (Single antenna at the emitter and at the receiver) the maximal achievable increase of the data rate depends on the minimum number of antennas in the transmitter and the receiver due to the compromise between the diversity gain and the multiplexing gain. The MIMO approach can be associated with beam-forming to control the direction and shape of the radiation pattern. MIMO techniques allow for important improvements in throughput performances, but it is a challenge to integrate many antennas and transceivers in a small mobile user device (size, losses, EMC…). With a multiple antenna base station and a single antenna user device it is possible to achieve some non-neglectable improvement but a multiple antenna user device is necessary to really take advantage of MIMO technique.

Despite the linearization, it is attractive to lower the NL effects by reducing the PAPR of the signal. One of the multiple possibilities is to code the amplitude information (envelope) with a modulator such as PWM or ΣΔ. Such modulations provide the coding of the envelope information (or amplitude information) under the form of a two-state signal (square-like). It results in the representation of the amplitude by a square signal mean value of which is the envelope information. Such a signal is more easily treated, but very important noise (white and shaped by the information itself) and spectral components are added. The coding efficiency defines the power ratio between the amplitude information and the PWM/ΣΔ coded signal. This can be seen in Fig. (7) where the spectrum reveals the amount of power, considered as noise when emitting the signal. Unfortunately, this ratio is very low: in the range of 10% for cases studied in [21, 38], depending on the switching frequency and signal scaling. The added penalty is that this coding efficiency directly multiplies the overall efficiency of the front-end due to a necessary filtering. Those techniques are known as LIST (Linear amplification employing Sampling Techniques) and were proposed many years ago at lower frequencies than the GHz. The frequency improvement of such coding made LIST very popular for RF transmitters where the envelope (in polar architecture) or the IQ paths (in cartesian architecture) are coded and generate a constant power signal at the input of the PA, avoiding AM/AM and AM/PM [32, 39-42]. Therefore, if the PA designed is able to amplify such squared (wideband) signal, this improvement suffers for one major default that is the impact on the overall efficiency and the need of more selectivity of the emission filter [38]. This is due to the high amount of power added on the signal in out-of-band. These frequencies must be filtered before the antenna because the PA amplifies them, contributing in power efficiency lowering. Avoiding PA NLs by envelope coding is very attractive but the spectrum of the driving signal is highly widened and can cause (i) PA instability due to frequency characteristics, or/and (ii) amplification of the unwanted/unused frequency components that lower the useful output power, saturate the PA and imply a strong attenuation by the emission filter (co-existence). In the context of multi-radio, multi-band antennas for such transceivers are advantageous because they can relax the filtering requirement, which means a technological realization less challenging. Although the amplification of important spectral re-growths can saturate the PA or/and damage it by reflected power or/and lower the efficiency, architectures designed showed the reduction of NLs with such process and investigated the use of digitized block at GHz RF frequencies. Fig. (7) represents polar and cartesian architectures with envelope coding. The spec-
trum is drawn before and after the emission filter and shows the importance of the filter selectivity to fulfill spectral emission requirements. Several modifications of these architectures have been investigated. Some of them consider the supply modulation of the PA and others are based on the digitization of the input of the PA (digital PA, D-PA) [42, 43]. Moreover, the digitization of the RF part increases the flexibility of the transmitter. State of the art performance of DACs (at RF frequencies) and RF filters are limiting key factors that reveal some technological locks in a Front-End design.

In order to increase the efficiency at its maximum, PA classes of operation were carefully studied and modified. High efficiency classes based on switched behavior of the PA are more and more popular. However, they imply the challenge of reproducing the envelope information by supply modulation and/or envelope coding and filtering, see Fig. (6).

Doherty, presented in Fig. (8), is a combination of PA proposed in another way. The principle is to amplify separately an “average power” part (the main PA) of the signal and a “peak power” part (if necessary). As the two PAs are connected, there is a loading effect while turning “on” and “off” the PA dedicated to “peak power”, named “auxiliary PA” for that reason. This is known to modify the impedance (load line) seen by the PAs and so the matching and power transfer. By increasing the power at the PA inputs when the auxiliary PA is tuning “on”, the impedance seen decreases and the load lines are modified in the way that PA classes change from AB, B towards C. This potentially increases the efficiency of the amplification, at the expense of linearity. In order to produce the effect of impedance reducing, it is necessary to use a quarter-wavelength line of transmission at the output of the main PA, as it is shown in Fig. (8). Doherty can significantly improve the efficiency thanks to load-pull measurements, and a careful synchronization between the main and auxiliary PA paths. Using the quarter-wavelength line made it difficult for frequency flexibility due to the theoretical frequency tuning. Multi-modes complex structures with several PAs can be investigated in that way for multi-radio, but imply the design of additional PAs.

C) (Fully) Digital architectures. The advantage of digital architectures is to limit the imperfections due to variations in the process. In analog architectures, the multiple functions basic characteristics vary (with process, temperature, voltage …), and so the architecture performance varies as well. This leads to the use of calibration loops. Digitization provides size optimization and good stability of the circuit, because it is built with more simple operators such as current sources, switches, capacitor charge/discharge… Inspired by the digitization of some RF blocks and the growing performances of RF-Digital technologies, fully digital architectures are emerging [44, 45]. These architectures are separated from digitized ones by the fact that they can be entirely based on structures non-existent with analog blocks. This classification is extended to architectures where the majority of the functional blocks are cascaded/combined digital functions (and so the technology process) and drive it to cartesian-like and polar-like architectures. Each element of the transmitter must be designed with regard to several parameters: (high) frequency, dynamic power control and consumption (leading to high efficiency). Due to the use of Sigma Delta modulators, WB-CW standards such as WiMAX 3.5 and 5.8 GHz need an over-sampling frequency in the range of 15 GHz (more than four times the carrier frequency). Using CMOS 90/40 and 32 nm technology is a current solution to address this challenge, while reducing the number of analog blocks. The main interest of such digital architectures is their high flexibility for RF transmission reconfigurability needed in the multi-radio context. This flexibility of the transmitter implies the definition of wanted key functions in digital ar-
Recent Patents on Electrical & Electronic Engineering, 2013, Vol. 6, No. 2

Diet et al.

...flexibility and signal spectral purity. These functions can be achieved thanks to the use of the ADPLL (All Digital Phase Lock Loop), the RF-DAC (and D-PA) and the flexible RF filter (for co-existence requirement). We will present briefly three architectures representing the trends in fully digital transmitters design, see Figs (9 & 10).

The first architecture, presented in Fig. (9) (left), is based on a classic direct conversion architecture using only one frequency transposition. The “Direct Digital to RF Modulator” (DDRM) architecture [46] was developed as a basis toward further architecture digitization. In this architecture, the system which is closest possible to the amplifier (which is still an analog block of the transmitter) is digitalized. As the first stage is a digital over-sampling stage, there is no DAC conversion of an analog baseband signal before the DRFC (analog part is directly RF). In this architecture, mixing and D/A conversion are performed by a single block: “Digital to RF converter” (DRFC), a kind of RF-DAC. In a DRFC, the LO signal enables to convert the digital information at the RF frequency while the analogical conversion is performed. The principle is that the data is over-sampled at the carrier frequency. The amplitude signal is coded into a digital word, digitally driving Gilbert cells. The resulting current is so proportional to the code word. The linearity performance and the signal resolution increase with the number of parallel cells. Due to this parallelization, IQ imbalance is limited and can only result from the average cell imbalances. Power control can be achieved by a bias current variation, which reduces the output current from each unit cell. As it uses no filter, the choice of the converter frequency is determined on the emitted spectrum. Indeed, the only filtering applied is the zero-order hold. The frequency must be chosen so that baseband harmonics are cancelled thanks to the zeros of the filter response.

Thanks to CMOS evolutions, digitized architectures were optimized to address spectrum sharing issues under the name...
of “Sigma Delta – RF DAC” architecture [47], see Fig. (9) (right). This architecture is based on a Sigma Delta coding, which generates important spectral re-growths at high frequencies (noise shaping) and enables the reduction of needed bits to represent the in-band information. To suppress a high part of these spectral re-growths, a delayed path is added after the analogical conversion, at RF frequency. In this architecture, the gain control can take place throughout the transmitter. The first power control appears after the first up-sampling filter. It is then a low speed dynamic control, resulting from multiplying IQ signals with a binary word. This word size depends on the control resolution. Several Delta-Sigma modulators in MASH structures are then used and I and Q channels are duplicated (I and I’, Q and Q’). I and Q signals go through a modulator, and signals I’ and Q’ are first delayed before a second modulator. This will result in the recombination of mixed signals around the carrier, and the creation of notches in desired frequency bands (certainly used by other standards). An additional 6 dB power control can be done by by-passing I’ and Q’, and the corresponding RF-DAC, losing the advantage of notches as a result.

The third architecture presented, seen in Fig. (10) [48], is preferably considered as a hybrid structure. In fact, the amplitude information is evaluated by RF sampling at twice the carrier frequency, and the phase information is restored by means of a phase shifting of the sampling signal (corresponding to one of the four “quadrant’s” of the IQ symbol). There is no complete separation between phase and amplitude in the transmitted signal processing. IQ signals are over-sampled and filtered in order to place the spectral re-growths far enough (zero order hold) and setting notches at multiples of the over-sampling frequency. Two signals are created: “abs(I + Q)” and “abs(I-Q)”. Moreover, phase control signals (2 bits) are determined with the signs of I and Q (complex quadrant). Amplitude signals are alternately sent to a digital differential amplifier (Dif. D-PA) at a rate of twice the carrier frequency. This can be seen as an over sampling by four of the IQ symbols leading to a better estimation of the amplitude. This coded amplitude signal feeds a Digital Power Amplifier (D-PA). This block will mix the signal around a carrier which phase can be shifted dynamically, depending on the sign of I and Q. Knowing I and Q signs, gives us information on the phase (linked to the quadrant position). A part of the phase information can be taken from “abs(I+Q)” and “abs(I-Q)” signals. This amplitude-coded signal can then be modulated by an RF carrier with an estimated phase. As seen in Fig. (10), four carrier signals with four different phases (0° 90° 180° and 270°) feed the Dif. D-PA depending on I and Q signs (2-bit control signals). As we obtain a NZR (Non Zero Return) signal, the filtering condition after the Dif. D-PA will be less constraining than when using I and Q directly. One of the carrier frequency signals and the coded amplitude are fed to the Dif D-PA. The carrier frequency signal is logical (“0” or “1” values) and activates one or another of the two different pairs in the Dif. D-PA, in function with the sign signals. If the carrier is set to “0”, “1”, respectively, the output is alternately “–abs(I-Q)” and “abs(I-Q)”, or “abs(I+Q)” and “–abs(I+Q)”. The output network of the Dif. D-PA is a band-pass filter around the carrier frequency, to reconstruct the original signal. In [48], an example is given at 5.8 GHz for a 64QAM modulation scheme and 10 MHz bandwidth WiMAX signal, which is very stringent. The resulting spectrum re-growths are below -50 dBc/Hz.

Digital RF architectures for WB-CW are providing several advantages such as flexibility and size optimization. Their design is subject to the improvement in power consumption of RF blocks. This research topic is currently popular for mobile and connectivity high data rate standards. Challenges in the design of such architectures highly depend on the performances of ADPLL, RF-DAC and D-PA. However, the filtering of replicas due to digitization (sampling) is a major problem at the power amplification and before the
emission (spectral requirements to fulfill). As the transmitter is designed to be flexible, the filter has to adapt to that requirement. The design of the front-end not only has to face the NLs problem but also to manage the filtering of the signal before the emission. Filtering requirements for such filter are presented in [49] and point the mandatory design of flexible filter in the transmitter.

We investigated such a filter design and proposed a LC modified biquadratic structure, with active inductances for the reconfigurability. Active inductor and preliminary filter were realized at ST Microelectronics and the filter measurements are in good agreement with the simulation. In Fig. (11) the active inductor principle and the filter (tunable) characteristic are presented. Such a design proves the feasibility of the reconfigurable front-end with digital based structure. A key point in the RF architecture domain evolution is that: while evolving towards the fully digital domain, analog problems of realization will be the current technological locks. Additionally, the consumption will be very different to evaluate and budget for such new structures.
We presented the evolution of the RF transmitter architecture domain to justify the current technological choices. We studied the potentiality of polar-based reconfigurable architecture and present this solution, although some complementary studies are in progress concerning the signal power distribution. Finally, our recent works in the domain of fully digital architecture are presented, and confirm the growing interest for these new architectures solution types. Fully digital architecture will now be concerning HPA design.

CURRENT & FUTURE DEVELOPMENTS

As presented in the paper, the future development contexts are both fully digital architecture and digitized-linearized architecture. These solutions are concerned by the high efficiency PA design, in switched mode, in order to restore the amplitude information potentially by (i) supplying modulation, (ii) driving signal coded or (iii) digital coding and control of the PA RF signal. Following that idea, our works are focused on the design and optimization of PA in switching mode and their use as RF power block combined. Fig. (12) presents an optimized class E PA for the polar based EER/ET architectures. The evolution of this conception will be to focus on the RF flexibility of this PA.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

The authors acknowledge ST Microelectronics for the realization and measures of the active filter.

REFERENCES


Passif Semiconductors Corp. Amplitude control system and method for communication systems. US patent n°US8432219.


F. Robert et al. Study of a polar ΔΣ transmitter associated to a high efficiency switched mode amplifier for mobile WiMAX. 10th annual IEEE Wireless and Microwave Technology Conference, WA-MICON, april 2009, Clearwater, FL, USA.


A. Pozgay et al. A Fully Digital 65nm CMOS Transmitter for the 2.4-to-2.7GHz Wi-Fi/WiMAX Bands using 5.4GHz RF DACs. IEEE ISSCC 2008, pp 360-619.
