



Departament de Teoria
del Senyal i Comunicacions



UNIVERSITAT POLITÈCNICA DE CATALUNYA

Ph.D. Thesis

MULTI LOOK-UP TABLE DIGITAL
PREDISTORTION FOR RF POWER
AMPLIFIER LINEARIZATION

Author: Pere Lluís Gilabert Pinal

Advisors: Dr. Eduard Bertran Albertí
Dr. Gabriel Montoro López

Control Monitoring and Communications Group
Department of Signal Theory and Communications
Universitat Politècnica de Catalunya

Barcelona, December 2007

Chapter 3

Overview of Linearization Techniques

3.1 Introduction

Linearity requirements in the transmitter chain, as it has already been introduced in Chapter 2, are a must for equipment that has to be certified according to a particular communications standard. In addition, in current mobile communications systems we can identify a triple compromise among: *transmission rate*, *batteries autonomy* and *coverage boundaries*. Therefore several issues have to be taken into consideration, such as current modulation schemes (spectral efficiency) presenting high *crest factors* (CF), the equipment power consumption (power efficiency) and the system linearity requirements, all of them contributing at the definition of this triple compromise.

Linearizing structures are capable to deal with the existing trade-off between linearity and power efficiency. However, none of the existing linearization techniques can be considered as a universal solution for itself, furthermore their optimality depends on many factors such as the specific requirements of a particular application. In general linearizers are aimed at:

- Canceling or reducing out-of-band distortion (e.g. measured in terms of ACPR or IMD reduction).
- Canceling or reducing in-band distortion (e.g. measured in terms of EVM).
- Prioritizing power efficiency by operating close to compression but maintaining linear amplification.
- Operating with spectrally efficient modulation schemes (despite their high PAPRs) in order to foster spectral efficiency.

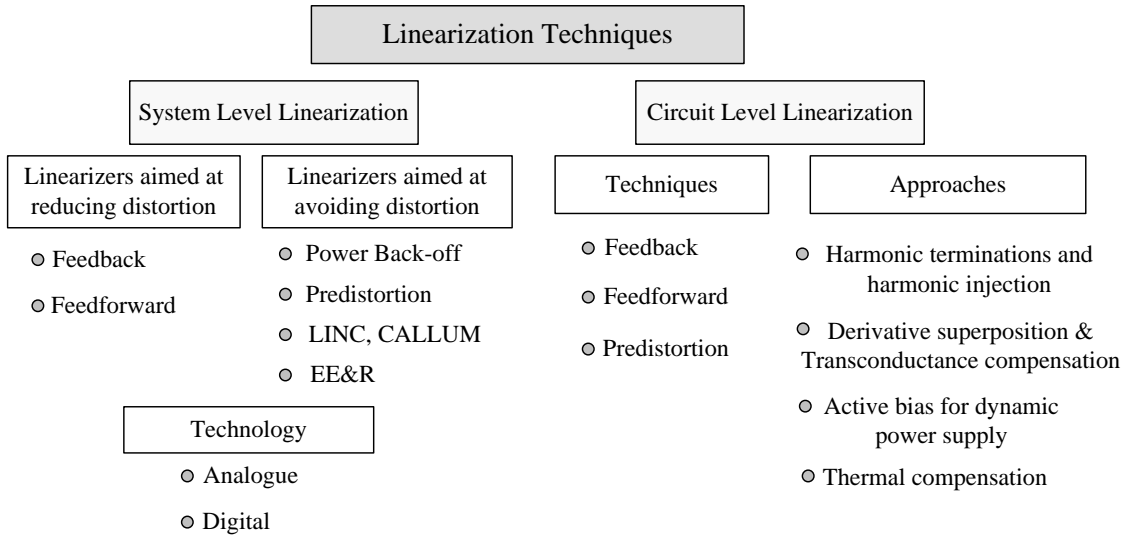


Figure 3.1: Classification of linearization techniques.

Figure 3.1 shows a scheme with a general classification of the most significant linearization techniques. Linearization can be carried out at two hierarchical levels: *circuit* and *system* level. *Circuit level linearization* consists in the application of linearization techniques at device (power transistor) level. At this linearization level, costs and size are more competitive in comparison to those at system level, so then results more suitable for consumer equipment (user terminal). However circuit level linearization does not perform significant IMD reduction improvement and becomes sensitive to future advances in transistors technologies becoming necessary detailed information on device models. On the other hand, *system level linearization*, is focused in the application of linearization techniques (both analogue and digital) at a higher level (system, sub-system). At this hierarchical level, higher IMD reduction can be achieved, but costs and size are also higher, thus results suitable for professional equipment (base stations).

This chapter will be focused in providing an overview of the most significant linearization techniques currently in use. To conclude this overview, a discussion on the suitability of employing one or another linearization technique for compensating PA nonlinear distortion in digital broadband systems will be presented at the end of this chapter.

3.2 Circuit Level Linearization

Linearization at circuit level is based in different approaches, depending on the main objective that they are aimed at solving [TAR06]. So then it is possible to distinguish among:

- *Harmonic terminations and harmonic injection approach*, aimed at reducing or canceling

nonlinearities.

- *Derivative Superposition and Transconductance compensation approach*, aimed at reducing distortion by linearizing the transconductance gain (g_m) in a common source FET amplifier.
- *Active bias for dynamic power supply approach*, aimed at reducing power consumption.
- *Thermal compensation approach*, aimed at compensating memory effects distortion derived from temperature variations, self-heating.

In addition, in order to carry out these linearization approaches, different techniques have appeared, either derived from system level solutions (e.g. feedback, feedforward, predistortion) or particular solutions for circuit level [Med06].

i) Harmonic terminations and harmonic injection.

At circuit level, an efficient strategy for power amplifier linearization is through a proper termination of the input and output ports at the harmonics of the carrier frequency. The basic idea behind this linearization technique is that a power amplifier always exhibits second-order nonlinearity, which is often even stronger than the third-order nonlinearity itself. The second-order nonlinearity can be exploited to reduce the effect of third-order nonlinearities in the generation of intermodulation products (IMP). By adding a second harmonic signal (centered at $2f_0$) at the input of the power amplifier, this signal will mix with the fundamental tone through the second-order nonlinearity, thereby producing an in-band contribution, that can be adjusted to compensate or even completely cancel the corresponding term arising from third-order nonlinearity. Such linearization techniques require first of all the capability to predict the correct amplitude and phase of the harmonic/baseband signals, resulting in effective intermodulation cancellation.

Following this idea, *harmonic injection* can be in general exploited to control linearity and can be applied in different ways:

Harmonic signals are *synthesized* and directly *injected* into the amplifier input by means of external sources. Some examples are provided in [Fan02, Ait01].

Harmonics are internally generated at the amplifier and fed back to the amplifier by means of proper *harmonic terminations*. In [Wat96, Mae95, Col00] some examples are given.

Finally, harmonics generated at the input and output ports can be exploited within the most standard linearization schemes:

- *Harmonic feedback*: part of the harmonic components at the output ports are fed back to

the input with a loop including an amplifier and a phase shifter to adjust the signal to be presented at the input port for linearization. Examples given in [Moa96, Ali98, Hu86].

- *Harmonic feedforward*: their use is limited since they need bulky and expensive control circuitry. Some examples implementing feedforward linearization at circuit level can be found in [Kan97, Hau01, Yan99].
- *Harmonic predistortion*: A particular class of linearization schemes is devoted to the linearization of two-stage amplifiers. In this case the second harmonic component of the interstage signal can be extracted, amplified and phase shifted to be exploited in the linearization of the final power stage. In [Kim99] an FET predistorter was introduced to generate a 3rd order IMD component that amplified by the main device, could cancel its own distortion.

ii) Derivative Superposition and Transconductance compensation.

Non-linearity in Common Source (CS) FET amplifiers mostly comes from transconductance (g_m) non-linearity. This non-linearity can be expressed from applying the Taylor expansion series for the drain-source current i_{ds} of such CS FET [TAR06]:

$$i_{ds} = I_{dc} + g_m V_{gs} + \left(\frac{g'_m}{2!}\right) V_{gs}^2 + \left(\frac{g''_m}{3!}\right) V_{gs}^3 \quad (3.1)$$

where V_{gs} is a small-signal gate-source voltage, and $g_m^{(n)}$ indicates the n-th derivative of g_m with respect to V_{gs} . The term $\left(\frac{g''_m}{3!}\right) V_{gs}^3$ is the responsible for producing in-band IMD3 products when a multi-tone signal is fed to the amplifier. The derivative superposition technique [Web96] enhances linearity in the main transistor by minimizing the $\left(\frac{g''_m}{3!}\right) V_{gs}^3$ term. This can be done by adding properly sized and biased transistors in parallel, because the $g''_m(V_{gs})$ characteristic can be either positive or negative depending on input level, biasing and/or threshold voltage. Some significant examples of this technique are reported in [Apa05, Kim03b].

iii) Active bias for dynamic power supply approach.

In this approach the use of active bias circuits to continuously optimize the performance of the circuit under varying bias conditions is an effective method for improving power efficiencies in active devices. This technique entails the use of active bias circuits which continuously optimize the performance of the circuit under varying bias levels, providing a high dynamic range when the receiver is near or in compression point. Moreover offers low power consumption when the receiver is in small-signal operation where a large dynamic range is not necessary. Some reported examples using this method are [Yan04, Hei02, Lar99, Kim03a].

iv) Thermal compensation approach.

The temperature variations caused by the dissipated power are determined by the thermal impedance, which describes the heat flow from the device [Vuo01]. Thermal impedance in the active device is not purely resistive, that is, some kind of dynamic forming a distributed low-pass filter is present. Temperature changes caused by the dissipated power do not occur instantaneously, but due to the mass of the semiconductor and package, frequency-dependent phase shifts always exist. Furthermore, since the temperature profile of silicon is quite sharp, it can be assumed that self-heating in the component is a more important cause of memory effects than heat generated by surrounding sources [Vuo01]. To overcome the difficulties associated with thermal memory effects a compensation module can be employed. Some interesting studies aimed at characterizing and compensating thermal memory effects are presented in [Alt01, Bou03, Ska02].

3.3 System Level Linearization

Linearization techniques at system level can be classified into five main groups. Nevertheless, it is possible to find in literature many alternatives and variations of the original ones [Sun95, Ken00]. According to the causality of the linearizer operation in relation to the distortion cancelation it is possible to split these five main families of linearization techniques into two groups [Zoz02]:

- *Linearizers aimed at reducing distortion.* First the nonlinear effects appear at the output of the PA and then, by taking some measurement of the present distortion, the objective is to reduce it. Two linearizer techniques are based in this principle of operation:
 - Feedback (and its variants: envelope, radiofrequency, Cartesian, polar)
 - Feedforward
- *Linearizers aimed at avoiding distortion.* These techniques are aimed at preventing from the nonlinear effects at the PA output. To obtain linear amplification at the output is then necessary to feed the PA with an input signal that has been previously processed (a signal presenting constant envelope, a predistorted input signal) or by simply adjusting the operating point (power back-off). The main linearizing techniques following this principle are:
 - Power back-off tuning (not strictly a linearizer, but the simplest technique)
 - Predistortion (analogue or digital; signal or data; base-band, IF or RF)
 - LINC /CALLUM
 - EE&R (Envelope Elimination and Restoration)

Since all kind of linearizers show as many relative advantages as limitations in practice, research in this area is still active, trying to find optimal solutions to specific problems. Moreover, current

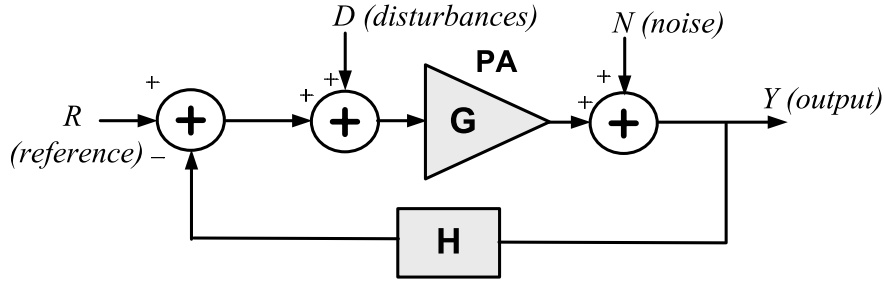


Figure 3.2: General block diagram of a negative feedback.

allowance of high speed digital signal processors (DSP's) for mandatory issues regarding signal processing (e.g. source coding, interleaving, IFFT for OFDM modulations) have revived classical analogue solutions and also have facilitated new approaches to the PA linearization problem.

In the following, these linearization techniques as well as some of their derivatives will be more deeply described. A final comparison presenting opportunities and weaknesses of each linearizer structure will be also provided.

3.3.1 Feedback Linearizers

Since Harold S. Black demonstrated the usefulness of *negative feedback* in 1927 [Bla34], feedback theory has been growing along time and an important number of feedback schemes have been developed [Fra02, Kuo95, Oga97, Nag82]. A general *feedback closed loop* block diagram including input and output additive disturbances is shown in Fig. 3.2.

Dynamics of the basic scheme in Fig. 3.2 can be described by studying the closed loop linear transfer function. Applying the superposition principle and denoting the system transfer function by G and the feedback loop transfer function by H (control theory notation), the system output is:

$$Y = \frac{G}{1 + GH}R + \frac{G}{1 + GH}D + \frac{1}{1 + GH}N \quad (3.2)$$

Where R is the reference signal (input signal), Y the output signal, D the input disturbance and N the output disturbance (noise). As it can be observed from (3.2), the bigger is H , the more insensitive is the feedback system to input disturbances. Moreover, the bigger is G or H , the more insensitive the output noise is. However, the price to reduce input disturbances is a loss in the overall system gain (amplification). Taking into account just the transfer function regarding the reference signal R , and considering H big enough, then

$$T = \frac{G}{1 + GH} \rightarrow H \uparrow \uparrow \rightarrow T \approx \frac{1}{H} \quad (3.3)$$

the system becomes insensible to variations in G (power amplifier) but at the price of, as mentioned before, decreasing the closed-loop (T) system gain.

In order to study feedback stability conditions, let us consider the frequency dependence of the system closed loop transfer function T :

$$T(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{G(\omega)}{1 + G(\omega)H(\omega)} \quad (3.4)$$

$T(\omega)$ will become unstable at the frequency (or frequencies) where:

$$G(\omega)H(\omega) = -1 \quad (3.5)$$

or equivalently, separating it into two conditions:

$$\begin{aligned} |G(\omega)H(\omega)| &= 1 \quad (0 \text{ dB}) \\ \arg(G(\omega)H(\omega)) &= \pm \pi \end{aligned} \quad (3.6)$$

The GH product is known as the 'open loop transfer function'. It is used in classical control for evaluating closed loop system dynamics taking computational advantage of this more simple 'transfer function'.

Fixing the amplitude condition, the second defines the phase margin, that is, the amount of phase that may be augmented before reaching the second instability condition. And in a similar way, fixing the phase condition, the first one defines the gain margin. These margins are easy to measure from Bode or Nyquist frequency response plots, and are the basis for designing the most simple loop controllers. Therefore, to preserve RF feedback stability it is necessary to keep enough gain and phase margins in the feedback path. These stability restrictions limit feedback technique to narrowband communications.

Among the classical feedback variations, some feedback structures have been proposed for communications systems. Such is the case of *Cartesian feedback*, *polar loop feedback* or *envelope feedback* linearizers. In addition, there are some advanced versions of the classical feedback based on the application of *control theory* but for communications problems, such as the *Hyperstable* linearizer based on passivity theory [Zoz04, ber01] or the *Robust Cartesian Feedback with Reference Model*, based in the H-infinity control theory [Mon05].

Cartesian Feedback

Figure 3.3 shows the Cartesian Feedback (CFB) architecture proposed in [Pet83], where the RF output signal is resolved into the In-phase (I) and Quadrature (Q) components. For simultaneously compensating both AM-AM and AM-PM distortion, the CFB linearizer uses two conventional I and Q feedback paths. The output signal ($y(t)$) is synchronously demodulated (Quadrature demodulator) and compared with the input Cartesian components ($x_I(t)$ and $x_Q(t)$)

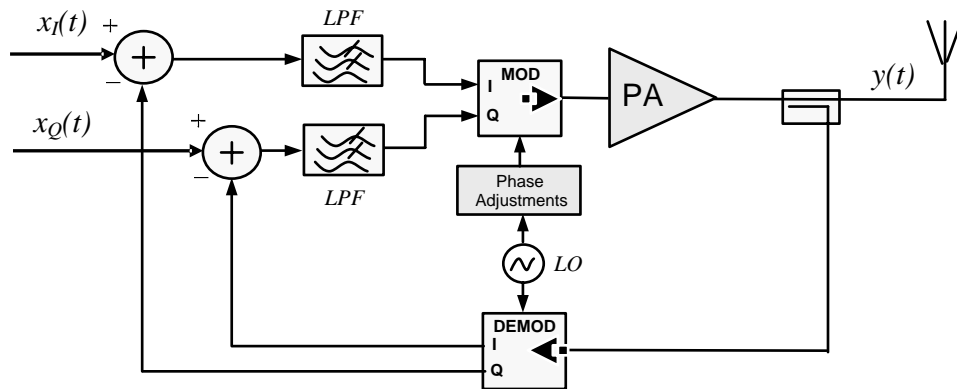


Figure 3.3: Block diagram of a Cartesian feedback transmitter.

to obtain the error signal. The error signal is fed to the loop filter followed by upconversion in a Quadrature modulator before it finally reaches the power amplifier.

The CFB linearizer can achieve good IMD reduction performance mainly considering narrowband signals [Cho02], despite the fact that CFB has been proven to work for wideband applications [Joh91]. Some of the issues related with the use of the CFB linearizer are related to its sensitivity with the stability margins of the linearizer loop, defined by its gain and phase margins. These stability margins are lower as higher is the frequency band. In addition, a phase shift in the loop when changing the carrier frequency, for example, can be another problem that degrades its performance.

Polar Loop Feedback

The block diagram of a polar loop transmitter is depicted in Fig. 3.4. The input signal is modulated at an intermediate frequency and it is split up in its polar components, amplitude and phase. Both input polar components are compared with their respective counterparts of the amplifier output signal. The resulting phase error signal controls a VCO that feeds the amplifier with a constant envelope but phase modulated signal. Similarly, the amplitude error signal modulates the collector voltage of the power amplifier. Thus a phase-locked loop is used to track the phase and a classical feedback circuitry to track the amplitude.

The required feedback bandwidths are different for the amplitude and phase components, thus limiting the available loop amplification for keeping positive gain or phase stability margins. Polar loop linearizers have shown good linearity performance for narrowband applications [Fer02, Sow04, Man00].

Envelope Feedback

Envelope feedback can be considered like a particular case of RF feedback and is an

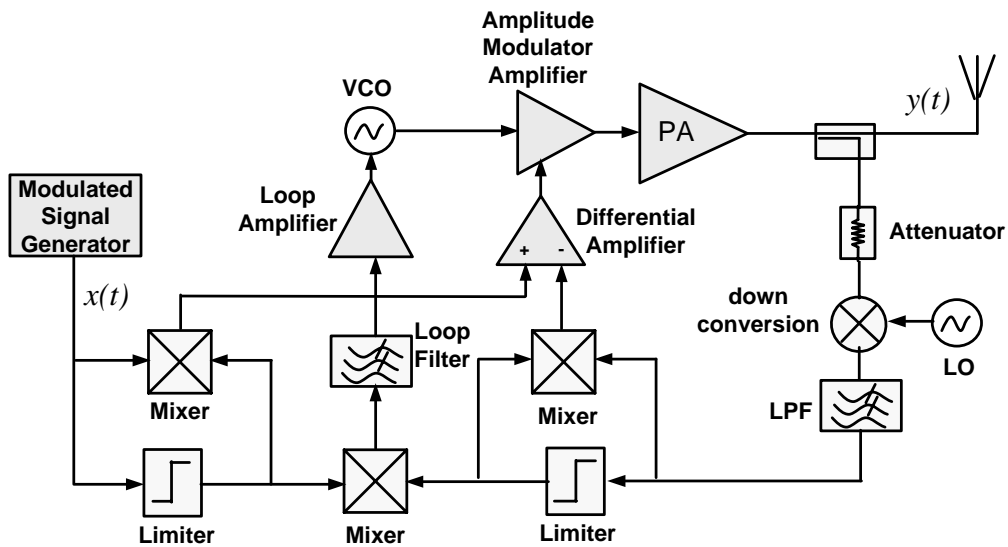


Figure 3.4: Basic configuration of polar-loop feedback.

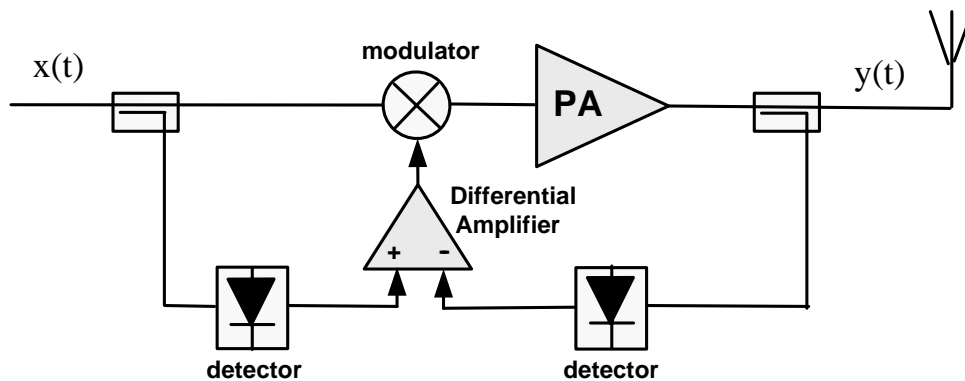


Figure 3.5: Basic configuration of an envelope feedback.

interesting linearization technique for such situations where device integrations is advisable [Par00, Gor02].

Figure 3.5 shows the block diagram of the envelope feedback linearizer. The error signal, obtained from the comparison between the envelopes of the input and output signals, is used as the modulating signal.

Hyperstable Linearizer

The basic configuration of an hyperstable linearizer is shown in Fig. 3.6. The structure of the hyperstable linearizer is based on an analogue implementation of the LMS algorithm that follows the reference model adaptive systems scheme [Lan85].

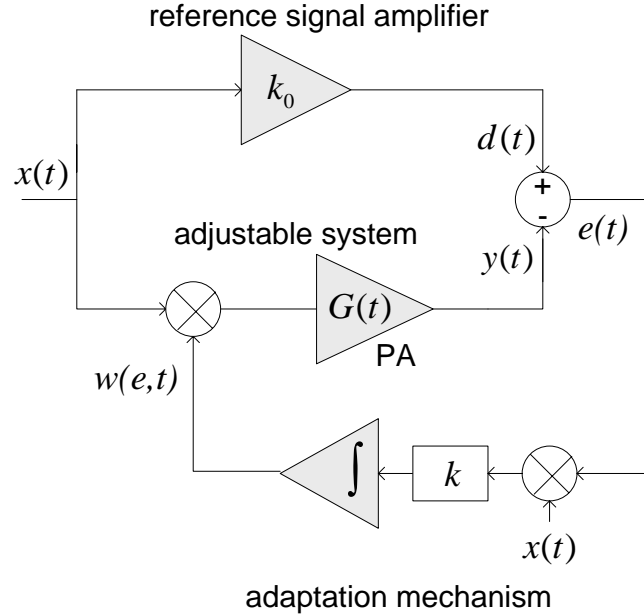


Figure 3.6: Basic configuration of an Hyperstable linearizer.

The reference model is simply a signal amplifier operating in the linear region. A sample of the reference amplifier output ($d(t)$) is compared with the output of the PA ($y(t)$), resulting then the error signal $e(t)$. The objective is to minimize or nullify the error signal by modifying the PA input according to the following expression:

$$w(t) = k \int x(t)e(t)dt \quad (3.7)$$

This expression can be seen as an analogue version of the LMS gradient algorithm [Ber03]. The resultant linearizer structure is an hyperstable system. Hence, according to *Control Theory*, and in particular from the *General Dissipative Systems* properties, $e(t) \rightarrow 0$ if $k > 0$. A detailed analysis from the Popov hyperstability criterion [Lan85] proves that this simple condition ($k > 0$) leads to $e(t) \rightarrow 0$.

Some examples of the hyperstable linearizer showing good IMD cancelation performance for narrowband signals can be found in [Zoz04, ber01, Ber03].

Robust Cartesian Feedback with Reference Model

The robust Cartesian feedback with reference model (CFB-RM) is a linearization technique that can be classified among those using the principles of distortion feedback and it is based on the application of *Robust Control Theories* over a Cartesian feedback structure.

A motivation for the original introduction of H-infinity methods by Zames [Zam81] was to incorporate system uncertainty into the controlled system. The H-infinity norm relates distur-

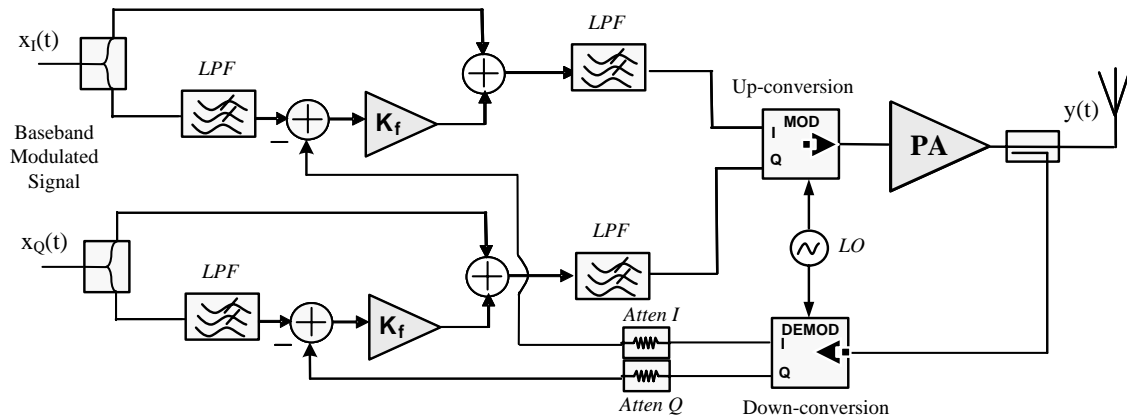


Figure 3.7: Block diagram of a Cartesian feedback with reference model.

bance inputs to error outputs in the controlled system and results appropriate for specifying the uncertainty level of the system.

Figure 3.7 shows the particular implementation of a CFB-RM linearizer. The main difference with respect to the classical CFB is the introduction of reference models (implemented with low pass filters) in both feedback paths. This solution permits adaptive linearization without having a priori any information regarding the nonlinear behavior of the PA. Further details on the design and final implementation can be found in [Mon05].

One of the advantages of the CFB-RM linearizer is its robustness. Despite the fact that exists a compromise (trade-off between linearity and efficiency) in the election of the K_f value (equivalent to the classical feedback loop gain), the penalization suffered by the CFB-RM linearizer is considerably lower than the one suffered by the basic CFB, that is, high linearity levels are achieved without a critical degradation of the overall system gain and power efficiency.

3.3.2 Feedforward Linearizers

The *feedforward linearizer* was first proposed by Harold S. Black [Bla77, bla37, bla29] when he was working for the Western Electric Company and was trying to improve the Bell System's new open-wire telephone system. Feedforward is recognized to be a suitable linearizer for operating with wideband signals. Fig. 3.8 shows the block diagram of a general Feedforward linearizer.

The feedforward functioning is very intuitive. Considering a two-tone test, the input signal is first equally split and fed to the upper and lower paths, respectively. The signal in the upper path is amplified by the PA. The output signal of the PA shows IMD and HD products. An attenuated sample of the distorted PA output signal is fed to the lower path and then subtracted with a delayed version of the input signal. Therefore, considering an ideal match between the

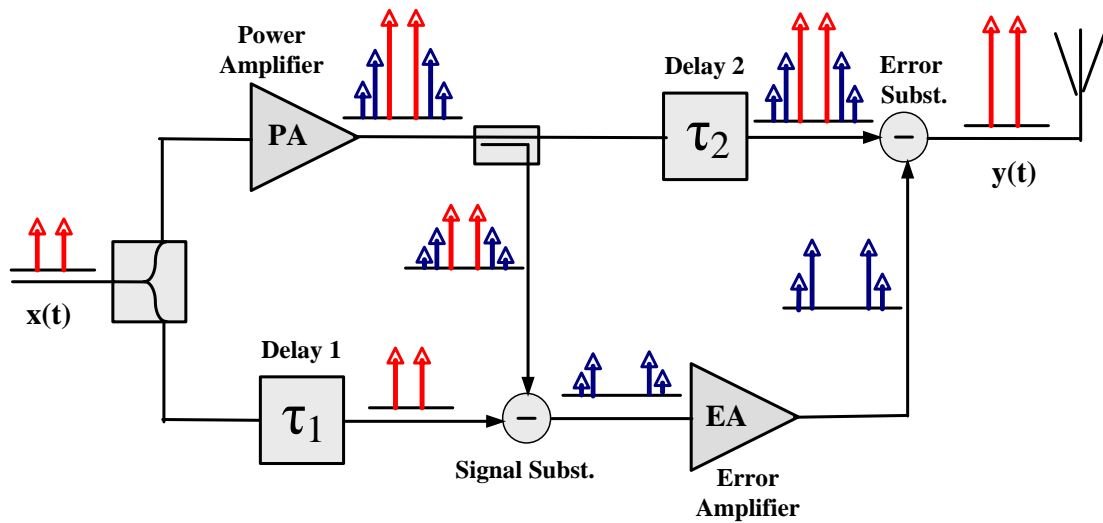


Figure 3.8: Feedforward simplified block diagram and principles of operation.

lower path delay (τ_1) and the delay introduced by the PA (in the upper path), the resulting error signal consists only of the IMD products caused by the PA. Then, the error signal is linearly amplified by the error amplifier and fed to the upper path (error subtractor). Contemporarily, the PA output signal in the upper path is delayed (τ_2) by an amount equal to the delay introduced by the error amplifier. So then, at the output of the error subtractor ideally only appears an amplified version of the desired two tone signal.

Feedforward is unconditionally stable, nevertheless it does not guarantee an optimal cancellation performance. Its open-loop nature makes it too sensitive to delay mismatches and tolerances of the components. These mismatches may produce imbalances among the different branches of the linearizer structure. If these imbalances are not compensated, the imperfect cancellation in both paths can take to a linearity performance degradation [Pot99]. Furthermore, it can also take to significant power efficiency degradation even when linearity levels are maintained [Gil04]. In addition, the same open-loop nature of the linearizer makes it too sensitive to nonlinearities and losses introduced by loop components (e.g. directional couplers).

Some patented improvements of the basic feedforward structure, such as the use of pilot tones to produce an arranged signal to control phase and amplitude imbalances [Yan00, Tat91, Mye86], or the use of compensation circuits for controlling the gain and phase components [Kha02, Cav95, Gad03], facilitate the monitoring and compensation of imbalances. However, the use of additional circuitry may penalize power efficiency, being already critical since the feedforward uses two amplifiers.

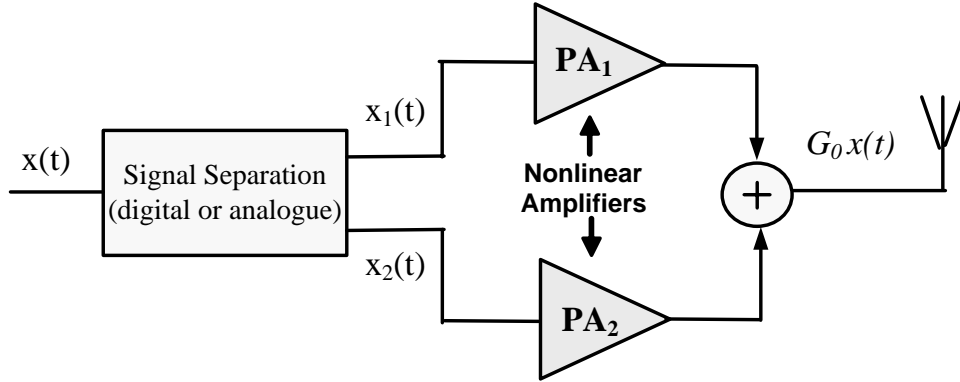


Figure 3.9: Block diagram of a LINC linearizer.

3.3.3 LINC and CALLUM Linearizers

The current allowance of high speed digital signal processors (DSPs) have revived classical analogue solutions involving signal processing, such as LINC or CALLUM. One of the main advantages of these linearization techniques is their applicability to high-efficiency PAs (including switched PAs). However, they are not exempt from practical problems in their realization.

LINC (Linear amplification using Nonlinear Components)

Linear amplification using nonlinear components -LINC- is the modern name given by Cox [Cox74] in 1974 when redesigned the Chireix *power outphasing technique* [Chi35] appeared in the mid 30's. The LINC linearizer technique consists of the separation of a variable envelope signal into two or more signal components of constant envelope. These constant envelope signal components are amplified separately by highly power efficient (class-D, class-E or class-F), but often strongly nonlinear power amplifiers, for a later combination of the component signals aiming to restore an amplified replica of the original signal.

Its operation is depicted in Fig. 3.9. It is based on a transformation of the original amplitude and phase modulated signal to be amplified. By considering a band-pass input signal,

$$x(t) = A(t) \cos(\omega t + \theta(t)) \quad (3.8)$$

The original input signal is split into two constant amplitude but phase modulated signals:

$$\left. \begin{aligned} x_1(t) &= \frac{A_{\max}}{2} \cos(\omega t + \theta(t) + \phi(t)) \\ x_2(t) &= \frac{A_{\max}}{2} \cos(\omega t + \theta(t) - \phi(t)) \end{aligned} \right\} x(t) = x_1(t) + x_2(t) \quad (3.9)$$

where A_{max} and $\phi(t)$ can be defined as follows:

$$\begin{aligned} A_{max} &= \max \{A(t)\} \\ \phi(t) &= \cos^{-1} \left(\frac{A(t)}{A_{max}} \right) \end{aligned} \quad (3.10)$$

These constant envelope signals are amplified separately in high efficient power amplifiers (see Fig. 3.9 and finally the combination of both signals again ($x_1(t) + x_2(t)$) will produce the desired output signal presenting linear amplification:

$$\begin{aligned} x(t) &= x_1(t) + x_2(t) = \\ &= \frac{A_{max}}{2} \cos(\omega t + \theta(t)) \cos(\phi(t)) - \frac{A_{max}}{2} \sin(\omega t + \theta(t)) \sin(\phi(t)) + \\ &+ \frac{A_{max}}{2} \cos(\omega t + \theta(t)) \cos(\phi(t)) + \frac{A_{max}}{2} \sin(\omega t + \theta(t)) \sin(\phi(t)) = \\ &= A(t) \cos(\omega t + \theta(t)) \end{aligned} \quad (3.11)$$

As it has been pointed out before, both input signal components are constant in amplitude, thus highly efficient power amplifiers presenting significant levels of nonlinear distortion can be employed. However, similarly to feedforward, LINC is extremely sensitive to amplitude and phase imbalances between both signal paths, as it has been reported in [Cas90, Cas93] for QAM modulations. Another issue that LINC linearizers have to handle is that the constant envelope signals will have a larger bandwidth than the output signal. The actual bandwidth expansion depends on the modulation depth of the transmitted signal. For EDGE modulation where modulation depth is moderate, a bandwidth expansion factor of about 3 has been calculated [Str02].

Finally, one of the biggest challenges for a LINC transmitter is power losses due to the combination of signal components [Sun94], since this could in some cases outweigh the benefits of employing highly power efficient PAs.

CALLUM (Combined Analog-Locked Loop Universal Modulator)

The *Combined Analog-Locked Loop Universal Modulator* -CALLUM- was proposed by Bateman [Bat92] and its principles are similar to the LINC linearizer. It combines two constant amplitude phasors to generate the output signal, but instead of having a signal component separator, the two constant amplitude phasors are generated by means of two Cartesian feedback loops. The block diagram of CALLUM transmitter is shown in Fig. 3.10.

The baseband equivalent to the transmitter output signal is obtained with a Quadrature demodulator (I-Q components) and compared with the corresponding Cartesian components of the input signal. The resulting error signal controls one VCO in each loop that in turn drives a power amplifier.

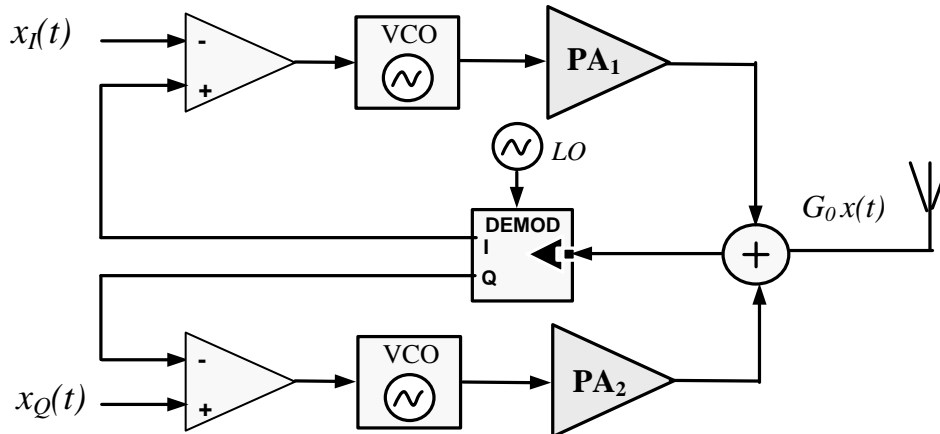


Figure 3.10: Bloc diagram of a CALLUM linearizer.

One of the main problems regarding CALLUM technique is related to its stability. The structure can become unstable if some I or Q components suffer a phase shift implying positive feedback. Due to potential stability problems, CALLUM is limited to narrowband applications.

3.3.4 Envelope Elimination and Restoration Linearizers (EE&R)

The *Envelope Elimination and Restoration* -EE&R- technique was first proposed by L.R. Kahn [Kah52] in 1952. This technique has been satisfactorily used in high-power transmitters due to its potential efficiency. Kahns technique has also been proposed for different communication systems such as orthogonal frequency-division multiplexing (OFDM) and digital video broadcasting (DVB). The use of modified EE&R in OFDM and QAM modulations for satellite applications in X-band has also been reported. Many papers discuss hybrid architectures and methods that can improve the overall system performance with respect to a combination of requirements on linearity, efficiency and bandwidth [Raa99, Rud03, Mil04, Die04].

The block diagram of an EE&R linearizer is shown in Fig. 3.11. The modulated input signal ($x(t)$) is split to recover separately envelope and phase information. For recovering phase information, the signal is hard-limited ($x_2(t)$), thus eliminating envelope variations and allowing a constant amplitude input to the nonlinear PA (low or null OBO operation). To restore the envelope information ($x_1(t)$), the PA is amplitude modulated swinging the DC bias (voltage).

Transmitters based upon the Kahn technique generally present good linearity levels because its linearity depends upon the modulator rather than the RF power transistors. The two most important factors affecting the linearity performance are the envelope bandwidth and the alignment of the envelope and phase modulations [TAR06].

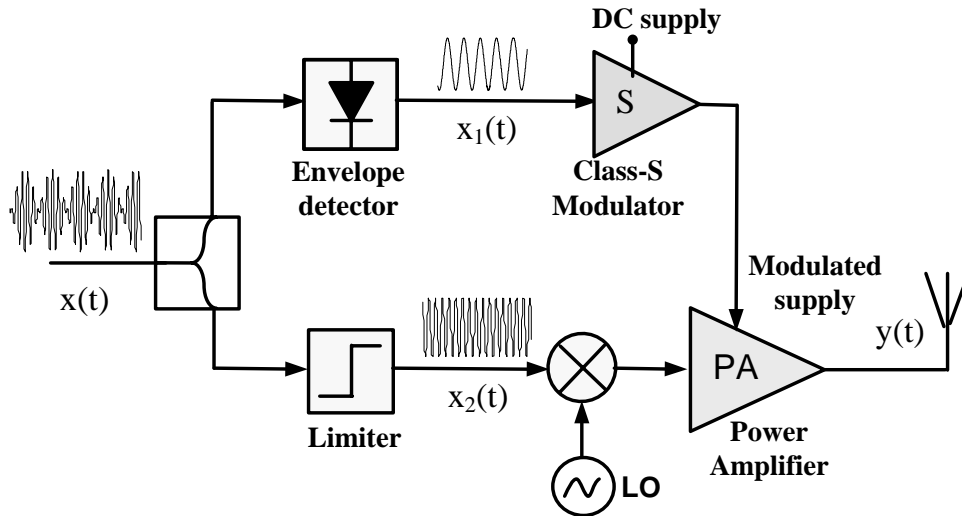


Figure 3.11: Basic configuration and principles of a EE&R linearizer .

In a complete EE&R *transmitter*, the use of a DSP permits performing the input signal processing, that is separating amplitude and phase of the input signal, already at baseband. This prior signal processing removes the need to modulate the carrier elsewhere in the transmitter architecture and also eliminates the requirements for a limiter and an amplitude detector to perform this signal components separation process at the input. The use of current high speed DSPs has permitted to overcome certain limitations and drawbacks derived from analogue circuitry, such as delays imbalances.

3.3.5 Predistortion Linearizers

Conceptually, *predistortion* (PD) is a linearization technique quite easy since it consists in preceding the PA with a device called predistorter in order to counteract the nonlinear characteristic of the PA. The objective of the predistorter is then to ideally reproduce the inverse PA nonlinear behavior and as a result having linear amplification at the output of the PA.

Figure 3.12 shows the basic principles of an open-loop predistorter. Several solutions have been developed to realize the predistorter, from digital baseband processing, which represents the main scope of this thesis, to processing the signal directly at RF using diodes as nonlinear devices. Moreover, most of current predistortion solutions already introduce some kind of feedback mechanism in order to allow *adaptive predistortion* what permits a more robust operation of the linearizer. Therefore all different solutions proposed for the realization of the predistorter can be somehow classified attending at some specific categories. An example of possible alternatives in the realization of the predistorter is depicted in Fig. 3.13.

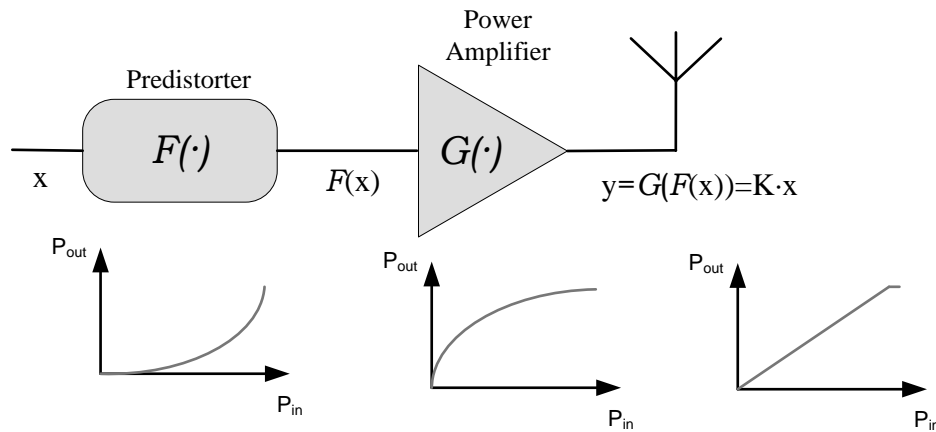


Figure 3.12: Fundamentals of predistortion linearization.

According to the position of the predistorter in the complete transmitter, predistortion can be carried out at *radiofrequency* (RF), *intermediate frequency* (IF) or *baseband* (BB). Predistortion at IF and BB have the advantage of being independent on the final frequency band of operation, as well as they are more robust in front of environmental parameters. Besides, the cost of ADC (analog to digital converters) and DAC (digital to analog converters) decreases at low frequency operation. One drawback regarding the predistortion compensation at IF or BB is the increasing linearity requirements since the up-converters can introduce additional distortion. However, up-converters (or I&Q modulators) can be avoided operating at IF and using software radio techniques, such as the so called *IF sampling*. Another significant issue is nonlinear distortion in the feedback loop (e.g. introduced by down-converters) when considering adaptive predistortion. This nonlinear distortion does not have to be compensated but can mask the open loop linearity performance and produce unwanted nonlinear compensations.

One example of *RF predistortion* using *analogue devices* is the so called *Cubic Predistorter* shown in Fig. 3.14. The cubic RF predistorter is aimed at canceling third-order IMD by adding a properly corrected in amplitude and phase cubic component to the input signal. In the case of a band-pass system only the third-order IMD products are usually reduced.. When a higher linearity improvement is required, predistortion is usually combined with feedforward architectures.

Figure 3.14 shows the block diagram of a typical cubic predistorter. The input signal is split into two paths. In order to ensure a perfect match in the recombination process of both split signals, a time-delay element is added in the upper path. The lower path is formed by a cubic nonlinearity performing the nonlinear predistortion, a gain (variable attenuator) and phase (variable phase-shifter) controllers that ensure a correct match at the combiner and finally, a post-distortion amplifier that buffers and amplifies the resulting signal. The amplifier in this

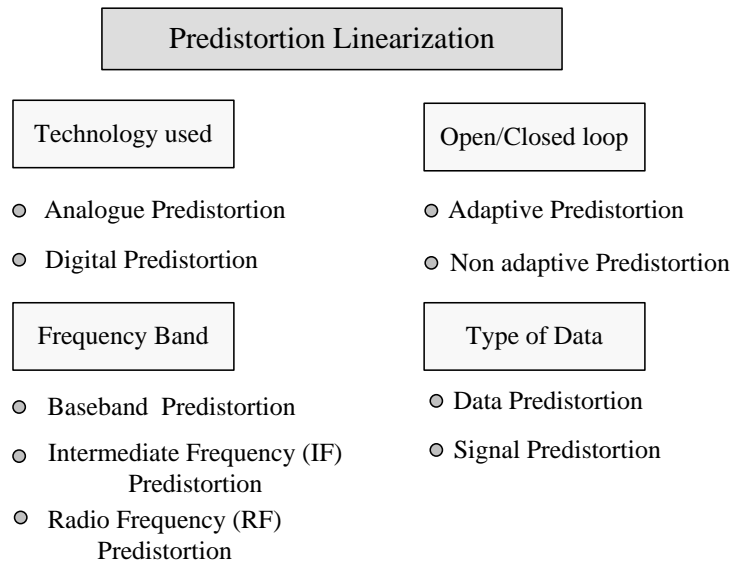


Figure 3.13: Classification of predistortion linearizers.

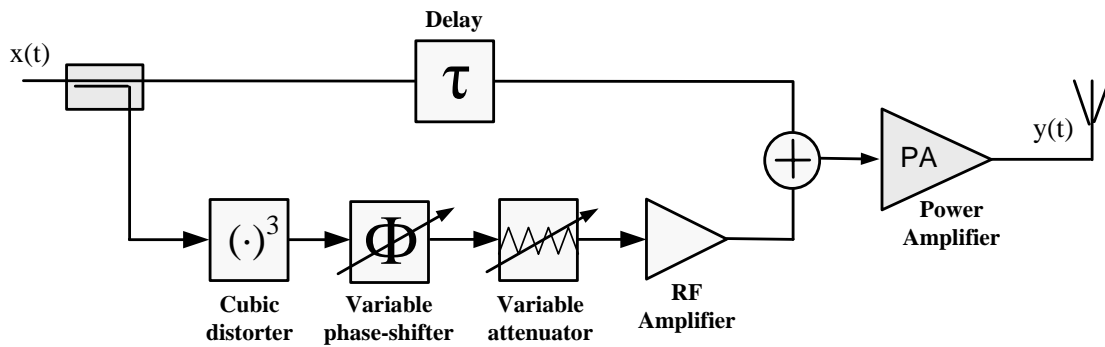


Figure 3.14: Block diagram of a RF cubic predistorter.

path is considered to be a small-signal device that does not contribute to the overall nonlinear distortion. Finally the predistorted signal in the lower path is subtracted from the signal in the upper one and fed to the PA.

The nonlinearity in the predistorter is usually created making use of the nonlinear characteristics of diodes: single diode, anti-parallel diode, varactor-diode [Yam96, Yu99] and also FET transistors [Nie02].

On the other hand, digital predistortion makes use of digital processing devices such as digital signal processors (DSPs) or Field Programmable Gate Arrays (FPGAs). Fig. 3.15 shows the simplified block diagram of a transmitter containing an adaptive digital predistortion module at BB. Two main approaches can be found in digital predistortion which are *signal* and *data predistortion*. The main difference between signal and data predistortion regards the position in

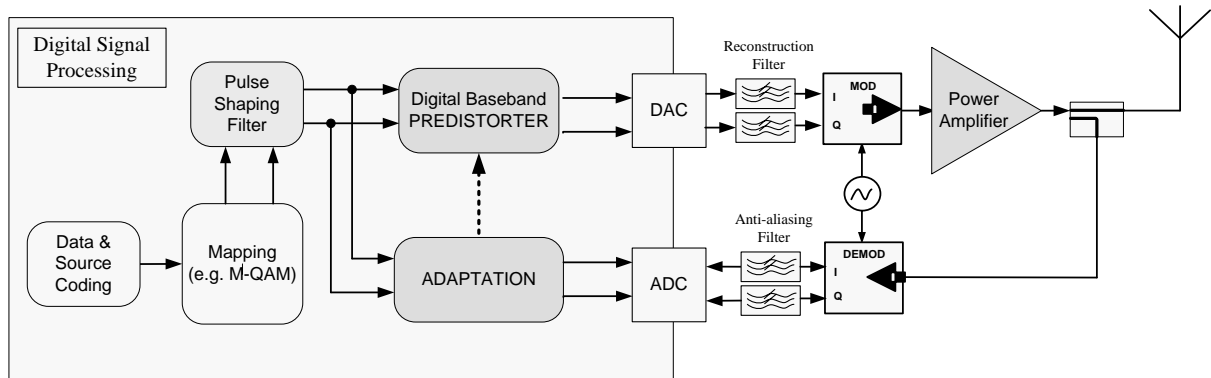


Figure 3.15: Block diagram of a transmitter with digital baseband predistortion.

the transmitter chain where the predistortion is carried out [And97]. The predistorter module in *data predistortion* is situated before the pulse-shaping filter, while in *signal predistortion* the predistortion is carried out after the pulse-shaping filter. Thus, Fig. 3.15 shows an example of signal predistortion.

The digital *data predistortion* technique is custom tailored to specific digital modulation formats, so then, the predistortion function of data predistorters aims at compensating the data vector space (constellation). Therefore, the predistorter coefficients are optimized by means of minimizing the Error Vector Magnitude (EVM), that is, compensating the in-band distortion introduced by the PA. But digital data predistortion does not compensate the out-band distortion. In addition, data predistortion is not transparent to all different modulation formats, so then, it results suitable for specific applications but unsuitable from a more versatile point of view. On the other hand, digital *signal predistortion* is aimed at canceling both in-band and out-band distortion, as long as the saturation level of the PA permits it. Moreover, digital signal predistortion is totally independent from the PA type, band of operation, class (A, AB, C), power technology (Solid State Power Amplifier (SSPA), Traveling Wave Tube Amplifiers (TWTAs)) and the most important, it is theoretically independent from the modulation format.

In Chapter 5 of this thesis the principles of digital baseband predistortion will be described and analyzed more in depth. In addition a specific contribution on this topic will be presented from both theoretical and experimental points of view.

3.4 Summary: Strengths and Challenges of System Level Linearization Techniques

Since a global optimal linearization technique does not exist, each linearization technique can be considered optimal for some particular applications, thus the decision for using one of another

linearization technique relies in several factors such:

- Linearity requirements (required levels of IMD cancelation).
- Complexity that will add the system level linearizer (trade-off between accuracy and complexity).
- The signal bandwidth that has to support the linearizer (high signal bandwidths make reconsider memory effects in PAs).
- Technology used, that is, linearization based on analogue circuitry or digital processing devices.
- Power efficiency requirements, trade-off between linearity and efficiency.
- Robustness in front of possible aging or unexpected changes

Feedback family linearizers (Envelope, Cartesian, Polar) have been extensively used in the past, but despite they are still used in some specific applications (circuit level linearization solutions), current demands on signal bandwidth make them unsuitable due to their stability problems in feedback loops. Moreover the use of feedback linearizers imply a loss in the overall system gain what can affect the overall system efficiency. However, feedback linearizers are insensitive to possible disturbances or changes in the transmitter subsystem due to its feedback adaptive nature and good IMD cancelation can be achieved in systems dealing with low or moderate signal bandwidths.

On the contrary to what happens with feedback linearizers, the EE&R and the outphasing techniques (LINC, CALLUM) have revived thanks to the use of current high speed digital signal processors. The use of DSPs has permitted to overcome certain limitations and drawbacks derived from analogue circuitry. Both EE&R and the outphasing techniques are being used to linearize high efficient switched power amplifiers (class-D, class-E, class-F power amplifiers) which are highly nonlinear. One of the biggest challenges of these linearization techniques is related to power efficiency, since they have to be very careful reducing power losses (e.g. power losses associated with the combination of the signal components) since this could in some cases outweigh the benefits of employing highly efficient PAs. On the other hand, current limitation with signals bandwidths are imposed by the processing speed of DSPs, thus since digital processing devices are always improving processing speed, these linearization techniques represent an interesting research field for engineers.

Feedforward together with digital predistortion linearization are the most suitable linearization techniques to cope with wideband signals. Feedforward linearization is unconditionally stable, can achieve effective cancelation of IMD (up to 30 dB in an 8-tone continuous random phase

test) and it is immune to PA memory effects. The main issue regarding feedforward linearization is related to its open-loop nature what makes it too sensible to possible unbalances in cancelation loops, thus degrading the overall linearity performance due to an imperfect cancelation. This can be compensated by means of a proper loop control, what introduces additional circuitry to the original feedforward structure. In addition the Error Amplifier and the directional couplers losses (mainly the output combiner) are critical for the IMD cancelation performance and power efficiency of the overall system.

Analogue predistortion is a more simple solution than digital predistortion (no need of DSPs) that can cope with high bandwidths despite provides moderate IMD suppression. However digital predistortion can achieve better correction capabilities and can be adaptive permitting corrections of possible unexpected unbalances. The main drawbacks are related to power consumption associated to the DSP device and bandwidth limitations due to the speed of the DSP processors. Since DSP are currently present in most of the communications transmitters for codification and digital modulation purposes, the allowance of efficient digital predistortion algorithms are an interesting field of research, moreover when memory effects are of significant importance in the IMD cancelation performance.

