The Digital Predistorter Goes Multi-dimensional: DPD for Concurrent Multi-band Envelope Tracking and Outphasing Power Amplifiers

Pere L. Gilabert, Gabriel Montoro, David Vegas, Nieves Ruiz and José Angel García

Pere L. Gilabert (plgilabert@tsc.upc.edu) and Gabriel Montoro (gabriel.montoro@upc.edu) are with the Dept. of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), c/ Esteve Terradas, 7, 08860, Castelldefels, Spain.

David Vegas (david.vegass@unican.es), Nieves Ruiz (mariadelasnieves.ruiz@unican.es) and José Angel García (JoseAngel.Garcia@unican.es) are with the Dept. of Communications Eng., Universidad de Cantabria (UC), Plaza de la Ciencia s/n, 39005 Santander, Spain.

I. Introduction

Over at least the last two decades, digital predistortion (DPD) has become the most common and widespread solution to cope with the power amplifier (PA) inherent linearity versus efficiency trade-off. In comparison to other linearization techniques such as Cartesian feedback (CF) or Feedforward (FF), DPD has proven to be able to adapt to the always growing demands of technology: wider bandwidths, stringent spectrum masks, re-configurability, etc. The principles of predistortion linearization (in its analog or digital form) are pretty straightforward and consists in preceding the PA with a subsystem (nonlinear function in a digital signal processor in the case of DPD or nonlinear device in the case of APD) that counteracts the nonlinear characteristic of the PA. Some great overviews on DPD can be found in [1]-[4]. Let us now have a look at the challenges that DPD linearization has been facing and will have to face in the near future with 5G New Radio (5G-NR).

In mobile communications, at some point after the evolution from 2G (GSM-> GPRS -> EDGE) to 3G (WCDMA-> HSPA -> HSPA+) and with the irruption of smartphones, the mobile data consumption began its unstoppable rise. Consequently, a more efficient usage of the precious commodity that is the radio electric spectrum was required. For example, in 2G GSM-EDGE the maximum link spectral efficiency was 1.92 bits/s/Hz, while only a few years later, with 3G-HSPA already achieved 4.22 bits/s/Hz [5]. However, the new waveforms were optimized from the spectral efficiency perspective, not power efficiency. From the power efficiency perspective, new challenges aroused, since we had to go from an efficient amplification of constant envelope modulation waveforms in 2G (i.e., 200 kHz bandwidth single carrier GMSK signals with 0 dB of PAPR) operating close or in saturation, to amplitude and phase modulated spread spectrum signals in 3G (i.e., 5 MHz bandwidth WCDMA signal with 6 dB of PAPR). High-fidelity amplification of non-constant envelope modulated signals with high PAPR requires avoiding envelope clipping at saturation, consequently, operating at significant power back-off levels where the PA efficiency figures are far below from the maximum achievable values.

With 4G (LTE, LTE-Advanced), to satisfy the need for higher data rates and users, spectral efficient multicarrier waveforms (i.e., 20 MHz of OFDM-based signals) with high-density modulations (e.g., 64-QAM) were introduced. Additionally, new technologies emerged in the evolution of 4G LTE, such as carrier aggregation (CA) and multiple input multiple output (MIMO), extending the maximum data rate for example up to 3 Gbps in LTE-Advanced when considering aggregated bandwidths up to 100 MHz and 4x2 MIMO. These technologies continue to develop to the massive speed and scale in 5G-NR.

In 5G-NR [6], the same network infrastructure will be able to efficiently serve different types of traffic with a very wide range of requirements such as a huge number of users for the Internet of Things, ultra-low latency and high reliability for mission critical systems or enhanced transmission rates for broadband mobile communications. 5G-NR intends to provide very high data rates everywhere. To achieve this goal, bandwidths up to GHz will be allocated at mm-wave bands, while at sub 6 GHz bandwidth of hundreds of MHz will be required.

Achieving these new capabilities requires coping with multiple demanding challenges which, particularizing for the design of radio transceivers, are related to: (i) ensuring the linearity of signals having bandwidths of several hundreds of MHz and peak factors exceeding 10 dB in order to ensure high transmission rates; (ii) improving energy and computational efficiency, as more dense deployments of base stations is expected to scale down the need for transmitted power; (iii) transmitting architectures with multiple antennas (massive MIMO in millimeter bands) and multiple power amplifiers to apply beamforming techniques that allow increasing the capacity and decreasing the radiated power; and (iv) simultaneous transmission and reception (full-duplex FDD in sub 6 GHz bands).
As mentioned before, DPD can overcome or at least mitigate the efficiency versus linearity problem in PAs. However, the resulting power efficiency achieved with linearization techniques applied to PAs operating as controlled current sources (e.g. class A, B, AB) is limited. To avoid wasting excessive power resources when handling high PAPR signals, either the operating conditions of a current source mode PA could be forced to follow its envelope, or switched-mode amplifying classes could be properly introduced. Among the set of techniques aimed to dynamic bias or load adaptation, envelope tracking (ET) [7]-[9], Doherty [10],[11] and LINC or outphasing PAs [12],[13] are the most widely proposed in literature.

In addition, significant efforts have been dedicated in recent past years to design wireless communication systems capable to handle multi-standard or multi-band signals at the same time. The advantage of having one single PA able to process signals in multiple bands simultaneously is the reduction of the number of components and cost of the RF subsystem [14]. When considering bands separated several hundreds of MHz, the implementation of a wideband DPD is unfeasible, especially in real-time platforms. Fortunately, DPD systems for multi-band signals can be significantly simplified assuming that the nonlinear distortions of concern are those that arise close to the band of interest, while the rest could be removed by filtering. When combining concurrent multi-band transmissions in PAs with dynamic load or dynamic supply modulation, the DPD needs to go multi-dimensional. That is, multiple-input single-output (MISO) DPD behavioral models are necessary to compensate for all the unwanted distortion effects that appear at the PA output.

Similarly, in multi-antenna systems where each transmit path has its own PA and antenna element, in order to avoid increasing the system complexity and cost, bulky components like isolators (placed between the PA and the antenna) are removed. As a consequence, these integrated multi-antenna transmitters typically suffer from nonlinear distortion due to the mixing of the antenna crosstalk and mismatch with the PA output, in addition to the nonlinear distortion caused by the PAs [15]. This is another example in which the DPD needs to go multi-dimensional in order to compensate for the multiple sources of unwanted distortion effects.

To sum up, Fig. 1 shows a conceptual map or compact overview on the aforementioned applications where multi-dimensional DPD is used and some references, research groups and companies that have published in these topics.

![Fig. 1. Applications for multi-dimensional DPD models (MISO DPD)](image)

In the following this paper will focus in the multi-dimensional DPD required to compensate for concurrent multi-band transmission when using dynamic load or dynamic supply modulated PAs. ET and outphasing PA are unlikely
to be deployed for ultra wideband applications in 5G mm-wave bands due to the bandwidth and power/cost budget limitations of these high efficient amplification architectures that require DPD linearization. However, in 5G sub-6 GHz macro base-stations, by considering envelope bandwidth or slew-rate reduction techniques [25] properly combined with dimensionality reduction techniques [30] to meet the low complexity requirements of the DPD implementation, the use of dynamic supply (e.g., ET or class-G PAs) or dynamic load modulation techniques (e.g., outphasing or load modulated balanced PAs) can still be considered as interesting solutions for high efficient amplification.

II. DPD for Power Amplifiers with Dynamic Supply or Dynamic Load Modulation in Concurrent Multi-Band Transmissions

A. Dynamic supply vs. dynamic load modulation.

In envelope tracking PAs, the supply voltage of the RF PA is adjusted according to the envelope of the RF carrier. Thanks to the dynamic supply, the RF PA (linear current-source mode, Doherty or even LINC PAs) can be forced to operate close to saturation which increases the power efficiency at power back-off. Several strategies can be designed (through the so-called shaping function) to shape the supply voltage signal in order to achieve better linearity, efficiency or to meet the slew-rate and bandwidth restrictions of the drain modulator [7]-[9].

In outphasing PAs [12],[13], the idea introduced by H. Chireix was to employ phase control of two constituent branch PAs operated in saturation, by enabling constant output envelope signals that were summed at the output of the system to allow amplitude modulation. The concept was later reintroduced by D.C. Cox who generalized the linear amplification with nonlinear components (LINC) approach. The main difference is that, while the Chireix combiner is non-isolating (and therefore the output signal from each PA changes the load impedance seen by the other), which provides good efficiency but not so good linearity, the LINC combiner isolates the two PA outputs (allowing them to see a fixed load at all times) which favors having good combining linearity but suffers from high dissipation at high outphasing angles.

![Fig. 2. Gain and efficiency evolution versus output power: a) at VGS = -2.3 V and RL = 50 Ω for different VDD values (left); b) for different outphasing angles (right).](image)

![Fig. 3. Gain (left) and efficiency (right) evolution versus output power at VGS = -2.3 V and VDD = 28 V for different RL values.](image)
If the PAs are assumed to operate as ideal voltage sources, with their outputs connected to a floating load R (usually provided by a balun), as the relative phase between the two sources is varied between 180 and 0 degrees, the effective loading on both PAs is therefore varied between R/2 and infinity. The output power from each branch PA is decreased, thereby reducing the dissipated power losses [12]. This is the principle of the active load modulation concept proposed by Chireix, who also added two fixed compensating reactances in order to provide some control over the efficiency versus output power profile.

The dynamic supply or active load modulation concepts can be appreciated in Fig. 2, showing the gain and efficiency evolution versus output power for different supply voltages (Fig. 2-left) and different outphasing angles (Fig. 2-right) when characterizing a single or two reactively combined continuous class-J mode GaN HEMT PAs, respectively [38]. For ET, the Nujira company (acquired by Qualcomm) introduced the concept of isogain shaping [39], where the instantaneous supply voltage is designed to achieve a particular constant PA gain. In this case, the ET PA system achieves low AM/AM distortion at the price of having less efficiency than when it is optimized for it. Analogously, the isogain approach can be applied in outphasing systems [38] or alternative dynamic load modulation techniques as shown in Fig. 3, where the gain (left) and efficiency (right) evolution versus output power of a single PA is shown for different resistive loading values. The isogain trajectory is the one providing constant gain at the price of losing certain efficiency in comparison to the optimum efficiency trajectory.

In the following subsection we will particularize the shaping of the supply voltage in ET PAs or the outphasing angle in outphasing PAs when considering concurrent multi-band transmissions.

B. Slow envelope generation in Multi-band Tx.

The amplification of concurrent multi-band or even carrier aggregated transmissions with high efficient topologies based on dynamic supply modulation (e.g., envelope tracking) or dynamic load modulation (e.g., outphasing) faces several challenges. On the one hand, to guarantee the desired linearity levels, the DPD has to be designed taking into account the difficulty of running the DPD at around 5x the signal instantaneous bandwidth (due to the bandwidth expansion occurring in the DPD process). When considering bands separated several hundreds of MHz, the implementation of a wideband DPD is unfeasible, especially in real-time platforms. Fortunately, DPD systems for multi-band signals can be significantly simplified assuming that the nonlinear distortions of concern are those that arise close to the band of interest, while the rest could be removed by filtering.

On the other hand, both dynamic supply and dynamic load modulation techniques are capable of achieving high power efficiency figures even when operated with amplitude and phase modulated signals presenting significant PAPR (good efficiency profiles with back-off operation). However, the efficiency decays with the signal’s bandwidth. Both dynamic techniques show a trade-off between the mean power efficiency that can be achieved and the instantaneous signal bandwidth handled by the PA.

A Chireix outphasing PA is intrinsically narrowband. The frequency dependent characteristic of the non-isolating combiner, including the compensating reactances, and the reactive elements in the device model usually impose strong limitations to its bandwidth. As a way of illustration, if designed with class-E switched-mode PAs (a highly attractive operating class in terms of efficiency performance), the efficiency contours rotate counterclockwise with the increasing frequency, while the mutual load modulation trajectories offered by a passive (Foster) combiner rotate in the expected clockwise sense [40]. In a pure outphasing operation, the constant-envelope phase-modulated signals to be handled by the constitutive branches may have a bandwidth several times wider than the original signal to be reproduced. For wideband or multi-band signals, a broadband input matching network would be also required to avoid undesired PM-to-AM conversion. In ET PAs, one of the main challenges regards the design of efficient envelope modulators capable of supplying the power required by the transistor at the same speed of the signal’s envelope. In concurrent multi-band transmission, the envelope of the resulting RF signal can present bandwidths that are several times (according to the rule of thumb around 3x) the instantaneous bandwidth. For example, with carriers separated some hundreds of MHz, the envelope modulator should present a slew-rate capable to efficiently amplify supply signals of GHz of bandwidth. Currently, commercial envelope drivers for base stations can efficiently handle signal bandwidths up to 40 MHz with power efficiencies greater than 70% [41].

Because the efficiency of the envelope modulator drops at high frequencies, to avoid dealing with this high-speed envelope variations in multi-band signals, some methods to reduce the bandwidth [24] or slew-rate [25] of the signal’s envelope have been proposed. In addition, in the particular case of dual-band signals, two main approaches to deal with the instantaneous dual-band envelope of the transmitted signal were proposed: i) perform the sum of
the modulus of the baseband signals (i.e., the peak of the instantaneous dual-band envelope, \( p=1 \) in (1)), as proposed in [26]; or ii) perform the square root of the sum of the squared modulus of the baseband signals (i.e., the average amplitude of the instantaneous dual-band envelope, \( p=2 \) in (1)), as proposed in [28]. These two approaches are particular solutions of the following method that can be generalized for multi-band transmissions. Therefore, to generate a slower envelope version \( E_s[n] \) of the original instantaneous envelope \( u[n] \), the generalized mean or power mean with exponent \( p \) can be extended for \( m \) bands:

\[
E_s[n] = \left( \frac{1}{m} \sum_{i=1}^{m} |u_i[n]|^p \right)^{\frac{1}{p}}
\]

(1)

Where the multi-band complex base-band signal is defined as

\[
u[n] = \sum_{i=1}^{m} u_i[n] e^{j\Omega_i n}
\]

(2)

where \( \Omega_i = \frac{2\pi f_i}{f_s} \), and with \( f_i \) and \( f_s \) being the center frequency and the sampling frequency, respectively.

Fig. 4 show the time domain waveforms and spectra of the original RF signal’s envelope and two examples of slow envelope generation (peak, \( p=1 \), and average, \( p=2 \), approaches) when considering a dual-band \( (m=2) \) and a multi-band transmission \( (m=5) \). The generated slow envelope then goes through a shaping function (a comparison of different shaping functions can be found in [7]), which in the case of multi-band transmissions basically consists in a detroughing function. Detroughing prevents the supply signal \( (V_{DD}(t)) \) to drop down to zero volts and thus avoiding sharp amplitude nulls in the time domain that may increase the bandwidth requirements. Moreover, a supply is typically more efficient for higher voltage levels and a limited voltage swing, while at low power levels the efficiency is not as important. Fig. 5 depicts two different shaping functions, the hard and soft detroughing. The soft detroughing can be defined with a function as described in [42],

\[
E[n] = \left( (E_{S_{TH}}) + (E_s[n])^6 \right)^{\frac{1}{6}}
\]

(3)

where \( E_{S_{TH}} \) determines the shaping curve and the minimum clipping level.
C. Multidimensional DPD

Dealing with signals presenting instantaneous bandwidth of several hundreds of MHz is a challenge not only for ADC/DACs that need to operate at sampling rates at least 5x the signal’s instantaneous bandwidth (despite there are undersampling techniques for the ADC [1], at least the DAC has to operate at full rate), but also because the FPGA has to operate at clock rates that, being optimistic, may be at the limit of current technology (e.g., current commercial available IPs for DPD are limited to signal bandwidths up to 200 MHz). As mentioned before, in multiband signals we can relax the clock rate constraints assuming that the nonlinear distortions of concern are those that arise close to the band of interest, and thus individual DPDs will take care of each one of the multi-band transmissions.

As explained in the introduction, several multiple input single output (MISO) DPD models have been published to compensate not only for the intra-band intermodulation distortion (both in-band and out-of-band distortion), but also for the cross-band intermodulation distortion (see Fig. 1). In addition, using a significantly slower envelope than the original instantaneous multi-band envelope for dynamically supplying or dynamically load modulating the PA will introduce additional nonlinear distortion. Consequently, to compensate for the slow-envelope dependent distortion effects, we need a concurrent multi-band DPD model that includes the information of the slow envelope,

$$x_i[n] = f_{DPD} \left( u_1[n], u_2[n], \ldots, u_K[n], E[n], n \right)$$

where $x_i[n]$ with $i=1,2,\ldots,K$ is the $i^{th}$ output of the MISO DPD, $u_1[n], u_2[n], \ldots, u_K[n]$ are multiple inputs corresponding to each one of the concurrent bands, and $E[n]$ is the generated slow envelope. Thus, the baseband signal processing for generating the predistorted I-Q signals is the same for both ET and outphasing PAs (which are considered as black-box systems) and consists in the MISO system described in (4), with as many input signals as concurrent bands and the slow-envelope signal used for both dynamically supply the PA or to code the outphasing angle.

Fig. 6. Block diagram of a multi-dimensional DPD for both envelope tracking and outphasing power amplifiers.
Fig. 6 shows a block diagram of a multi-dimensional DPD for compensating for the intra-band, cross-band and slow envelope dependent distortion effects in a concurrent multi-band system when considering either dynamic load modulation (i.e., outphasing PA) or dynamic supply modulation (i.e., envelope tracking PA). In the case of the outphasing PA, the outphasing angle coding and the calibration of the gain and phase mismatch of the two signal paths is carried out in the outphasing transformation block. Additionally, the time alignment between the RF signal and the supply modulated signal is carried out in the slow-envelope generation block. Finally, the time alignment and amplitude normalization between the input-output I-Q signals is done in the DPD adaptation block.

D. Experimental results: envelope tracking and outphasing PA in a concurrent dual-band transmission.

Let us now particularize the general block diagram in Fig. 6 for the specific case of a concurrent dual-band transmission. Fig. 7 shows the structure of the proposed 3-D distributed memory polynomial model (3-D DMP) presented in [26], that includes three branches to compensate for in-band, out-of-band and cross-band intermodulation distortion and the slow-envelope distortion, respectively.

This particular 3-D DPD model was used in a concurrent dual-band transmission in [26], [30], to compensate for the unwanted nonlinear distortion of an envelope tracking PA. For example, Fig. 8-left shows the output spectra of an ET PA used in [30] before and after 3-D DPD. Using the slow envelope to dynamically supply the PA is a suboptimal solution from the power efficiency point of view in comparison to using the original RF signal’s envelope. However, is the only feasible solution taking into account the drain modulator bandwidth limitations. Moreover, this approach is always more efficient than considering fixed supply. Fig. 9-top summarizes the results presented in [26] for a concurrent dual-band ET PA where, even using the slow envelope (both the peak of average approaches) to dynamically supply the PA, we can double the drain efficiency in comparison to fixed supply. In addition, the ACLR threshold of -45 dBc is guaranteed applying 3-D DPD linearization with less than 100 coefficients [26].

Let us go back to the general block diagram in Fig. 6 to now consider a concurrent dual-band outphasing amplification. In particular, we will consider a mixed or hybrid-mode outphasing amplification [43], since we will allow some amplitude variation in the phase-modulated outphasing components. This hybrid approach helps to mitigate the bandwidth expansion that occurs when amplifying pure constant amplitude phase modulated components. This way, the hybrid outphasing PA becomes more linearizable, even if is at the price of slightly degrading the power efficiency. In fact, at deep power back-off there may be no benefits in terms of power added efficiency when operating in a pure outphasing mode. Therefore, both hybrid outphasing signal components can be defined at baseband as

\[
x_{\text{outph}}^{-1}[n] = A[n] e^{j(\alpha[n] + \phi[n])} = x_{\text{outph}}^{-1}[n] + jx_{\text{outph}}^{-2}[n] \\
x_{\text{outph}}^{-2}[n] = A[n] e^{j(\phi[n] - \alpha[n])} = x_{\text{outph}}^{-2}[n] + jx_{\text{outph}}^{-2}[n]
\]  

Where half of the outphasing angle, \(\alpha[n]\), is defined as a function of the slow envelope \(\alpha[n] = f(E[n])\), \(\phi[n]\) is the original instantaneous phase of the dual-band signal, and where \(A[n]\) is the amplitude of the hybrid or mixed-
mode outphasing components where, unlike the pure outphasing components, some amplitude variation is allowed:

\[ A[n] = \left| \frac{x[n]}{E[n]} \right| \beta \quad \text{with} \quad 0 < \beta < 1. \]

Fig. 8-right shows the dual-band output spectra of a hybrid outphasing PA before and after 3-D DPD linearization where again, dimensionality reduction techniques are applied to keep the number of coefficients below 100. On the one hand, to select the most relevant basis of the DPD function in the forward path, a greedy algorithm such as for example the orthogonal matching pursuit (OMP) [30],[44] can be used; on the other hand, feature extraction techniques such as the principal component analysis (PCA) [45] or the partial least squares (PLS) [30] can be used to reduce the number of basis in the adaptation subsystem. Fig. 9-bottom shows that, even using the slow envelope, the efficiency is doubled in comparison to not applying any kind of dynamic load modulation. By coding the slow envelope into the outphasing angle, an efficient concurrent dual-band transmission could be feasible using a dual-band Chireix topology [46],[47], with a much less demanding implementation than the wideband counterpart that would be required if coding the original envelope.

To validate the closed-loop DPD algorithms, a common solution as a previous step to a future real-time implementation, is the use of a hardware in the loop architecture as the ones shown in Fig. 10. In a PC controlled system, the digital signal processing is carried out in Matlab. Then, instrumentation or commercial off-the-shelf boards are used to both generate (DAC) and upconvert the RF signal to be amplified and after going through the PA, downconvert and/or digitize (ADC) the feedback signal.

---

Fig. 8. Unlinearized and 3-D DPD linearized output spectra of: a) an envelope tracking PA [29] (left), b) a hybrid outphasing PA (right).

Fig. 9. 3-D DPD linearization of concurrent dual-band transmission for both ET [26] and outphasing PAs.
Fig. 10. Hardware in the loop general architecture.

III. Conclusion

DPD is the linearization technique par excellence to deal with the inherent PA linearity versus power efficiency trade-off. In the last decade the DPD has gone multi-dimensional in order to address new linearity challenges that have arisen due to, in the one hand, the inclusion of techniques oriented at maximizing spectral efficiency such as concurrent multi-band or carrier aggregated transmissions and MIMO; on the other hand, the adoption of high power efficient architectures based on dynamic supply or dynamic load modulation of the PA. Particularizing for multi-band transmission in ET or outphasing PAs, we have shown that one strategy to cope with the bandwidth limitations of both amplification architectures consists in using a slower version of the original RF signal’s envelope to both dynamically supply or load modulate the PA. Finally, some results considering a concurrent dual-band transmission have shown that power efficiency values can be significantly enhanced in ET and outphasing PAs using slow envelopes and 3-D DPD to compensate for all the PA unwanted linear and nonlinear effects.

Acknowledgements

This work has been supported in part by the Spanish Government and FEDER under MICINN projects TEC2017-83343-C4-1-R and TEC2017-83343-C4-2-R and by the Generalitat de Catalunya under Grant 2017 SGR 813.

References


