Digital Predistortion for Dual-band Doherty Power Amplifiers with Mixed-mode Carrier Aggregation

Noel Kelly, Wenhui Cao, and Anding Zhu

University College Dublin, Ireland

Introduction

As new wireless communication standards seek to meet the myriad demands of today's end users, the traditional RF transmitter chain is faced with an evolving set of challenges. To maximise information throughput, signal bandwidths are continuously growing. In fourth generation long-term evolution-advanced (LTE-A), for example, to meet the stated goals of 1 Gb/s downlink and 500 Mb/s uplink speeds, the carrier aggregation (CA) technique was introduced, expanding the maximum transmit bandwidth up to 100 MHz [1]. With mobile traffic expected to increase by a factor of 1,000 over the next decade, future fifth generation (5G) systems will continue to place new demands on the transmit chain [2].

The resulting need for highly linear and power efficient signal amplification across wide bandwidths has led in recent years to advanced development of alternative power amplifier (PA) architectures such as envelope tracking, outphasing and Doherty [3]-[5]. Despite the success of these new PA architectures, there remains a fundamental trade-off between linearity and efficiency at the transmitter [6]. Power-efficient amplification cannot be achieved without some degradation in signal quality. In light of this, digital predistortion (DPD) continues to play an important role. By applying compensation in the digital domain prior to the signal entering the nonlinear transmitter chain, DPD improves the terms of the linearity/efficiency trade-off without adversely affecting overall complexity [7]. As with all components of the transmitter chain, DPD techniques have had to evolve to meet the challenges of new standards. The newest generation of DPD techniques address problems from enabling the low-complexity and power and cost efficient predistortion required by small-cell base stations to developing novel model structures for nontraditional PA architectures [8]-[9].

The CA technique gives rise to another interesting DPD scenario, which was the focus of the 2016 IEEE Microwave Theory and Techniques Society (MTT-S) International Microwave Symposium (IMS) DPD student design competition, sponsored by Technical Committees MTT-9 and MTT-11. A common case when using CA is the so-called dual-band operating mode, where there is a large frequency gap between two carriers. To reduce costs, it is common for the transmitter architecture to use a single PA chain to simultaneously amplify both carriers [10]. This presents a problem for traditional DPD as to capture the full transmit bandwidth needed to train a single wideband model would require an extremely wideband and costly feedback loop. On the other hand, treating each carrier as a separate signal ignores the distortion due to cross modulation between the two and degrades the linearization performance. The two-dimensional DPD (2D-DPD) technique [11] has been proposed to enable efficient, accurate predistortion for both carriers in this scenario and forms the basis for the solution presented here.

Competition Details
Competition Scenario

This year’s competition focused on the design of a digital predistortion solution to maximize the output power of a highly efficient gallium nitride (GaN) high-electron-mobility transistor Doherty power amplifier excited by a dual-band mixed-mode signal modulated onto a 2.6 GHz carrier frequency, within a specified set of linearity constraints. The test signal, shown in Figure 1, is a single 500 µs frame composed of two carriers: a 4-carrier Global System for Mobile Communications (GSM)-like signal and a single-carrier orthogonal frequency-division multiplexing (OFDM) signal. Table 1 outlines the signal specifics. The goal is to linearize the PA so that it can deliver maximum output power to a reference load under single or dual band operation. Figure 2 shows the AM/AM and AM/PM characteristics of the complete transmitter chain excited by a dual-band signal with input power -21 dBm. As can be seen, in addition to the nonlinear behavior, the wide bandwidth also leads to substantial memory effects.
Figure 1. The characteristics of the dual band test signal: (a) frequency-domain normalized power spectral density and (b) time domain normalized amplitude.

<table>
<thead>
<tr>
<th></th>
<th>Band 1</th>
<th>Band 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal Type</td>
<td>GSM</td>
<td>OFDM</td>
</tr>
<tr>
<td>Frequency Offset</td>
<td>-50 MHz</td>
<td>+50 MHz</td>
</tr>
<tr>
<td>No. Carriers</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Carrier Bandwidth</td>
<td>200 kHz</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Carrier Spacing</td>
<td>1.67 MHz</td>
<td>-</td>
</tr>
<tr>
<td>Total Bandwidth</td>
<td>5.2 MHz</td>
<td>10 MHz</td>
</tr>
</tbody>
</table>

Table 1. Test signal characteristics.
Scoring Metrics

The competing DPD solutions were evaluated using a scoring metric designed to account for three parameters of interest: output power on each active band, adjacent channel power on each active band, and time domain normalized mean square error (NMSE) on each active band. The total score formula is given by:

$$S = \sum_{i=1}^{3} \sum_{b=1}^{N_a} \frac{(P_{av,b} - 3) + (IMD_{i,b} - 45) + \min(0, -30 - NMSE_{i,b})}{N_a}$$  \hspace{1cm} (1)

where $N_a$ is the number of active bands, $b$ selects the band, and $i$ selects the measurement. The score is improved when the output signal average power is greater than 3 dBW, the intermodulation ratio (IMR) or adjacent channel power ratio (ACPR) value is greater than 45 dB, and the NMSE value is less than -30 dB. Conversely, the score is reduced if the signal measures worse than any of the above limits.

Figure 2. Measured AM/AM and AM/PM characteristics.
Measurement Platform

The competition measurement platform used a Rohde & Schwarz SMW200A vector signal generator (VSG) and a FSW8 vector signal analyser (VSA) controlled via a MATLAB server. The VSG and VSA operate at a sampling rate of 200 MS/s. The baseband test signal is generated in MATLAB and uploaded to the MATLAB server by the remote user. The baseband signal is then loaded into the VSG where it is up-converted to the RF frequency and fed to a Wolfspeed (previously Cree) GaN-based Doherty PA. The PA is designed using the CGHV27015S and CGHV27030S transistors. The output signal is attenuated by approximately 40 dB before being sent to the VSA receiver where it is down-converted and demodulated back to baseband. The baseband data are then captured and downloaded from the VSA by the MATLAB server and relayed to the remote user. Control of the relevant transmit and receive parameters for the instruments is also handled by the MATLAB server. The complete remote measurement platform structure is outlined in Figure 3.

Predistortion Solution

Signal-adaptive DPD

According to the competition specification, the DPD solution should be capable of linearizing the power amplifier across a range of signal formats. There are three possible transmit scenarios for the prescribed dual band signal: both carriers turned on, only the OFDM carrier turned on, and only the GSM carrier turned on. Together these scenarios represent a very wide range over which to apply a single DPD model. The parameters of such a model would need to be chosen to give the best “average” performance across the three cases, leading to sub-optimal performance for each individual case.

The need for separate models and coefficient sets to maximise performance, despite the physical power amplifier remaining unchanged, stems from an important nuance of the digital predistortion technique: the optimum DPD model coefficients extracted for one signal type can be very different
from those of a different signal even when the power levels and device under test remain unchanged.

A universal optimum set of coefficients do not exist for a number of reasons. First, the models used for the PA inverse are not ideal, meaning there is no perfect coefficient set for which the error will be completely removed. Thus, for a given coefficient extraction operation the aim is not to find the universally optimum coefficient set but instead the set that minimizes the error across the training signal distribution. This means that for each new distribution there exists a different set of coefficients that represent the optimum solution. As each signal format has a different distribution, it follows that a new coefficient set is required for each signal type.

In addition to the modelling side, the PA itself is affected by the signal format. Different signals can lead the device to settle into different “steady-state” operating modes. The PA temperature can vary with these different modes, leading to variations in the gain. As the PA characteristics change with its gain, a coefficient set is only valid for a specific gain value. Different signal formats therefore correspond to different operating modes and, ultimately, different optimum coefficient sets.

Overall performance can therefore be improved by tailoring different DPD models and their corresponding coefficient sets to suit the needs of each of the test scenarios. To do this, the competition DPD solution employs a simple power measurement in the two bands of interest to determine the type of input signal and select the appropriate model. In the competition, two different model types are used: a modified 2-D model when both bands are turned on and a conventional single-band model for cases where only one band is active.

Single-Band Operation

For single-band predistortion the decomposed vector rotation (DVR) model has been chosen [9]. The model structure is given as:

\[
\hat{y}(n) = \sum_{i=0}^{M} a_i \tilde{x}(n-i) + \sum_{k=1}^{K} \sum_{i=0}^{M} c_{k,i,1} \|\tilde{x}(n-i)\| - \beta_k |e^{j\theta(n-i)}| + \sum_{k=1}^{K} \sum_{i=0}^{M} c_{k,i,21} \|\tilde{x}(n-i)\| - \beta_k |e^{j\theta(n-i)}| \tilde{x}(n) + \ldots
\]

where \(|\cdot|\) represents the absolute value operator and \(a_i\) and \(c_{k,i}\) are the model coefficients. The memory length is given by \(M\), and \(K\) sets the number of DVR hyperplanes. The DVR model replaces the conventional polynomial type basis functions of the Volterra series with a vector decomposition and phase rotation operation. This is both easily implemented in hardware and better suited to modelling the more unusual nonlinear characteristics of devices like Doherty PAs compared to traditional Volterra series-based models [9].

Dual-Band Operation
Despite the large separation between the two carriers, the high-end receive chain used in the test platform enables the full band of interest to be captured in one step. This eliminates one of the most commonly cited problems of dual-band operation and theoretically enables the use of conventional single-input single-output (SISO) DPD models for linearization. To begin, we first investigated this simple approach, evaluating a number of popular models (memory polynomial, generalized memory polynomial, dynamic deviation reduction-based Volterra series [12]-[14]) in a conventional DPD architecture with a common model applied across the full captured bandwidth. Although each model succeeded to some degree in improving the transmit chain linearity, the performance was relatively poor, especially in terms of out-of-band cancellation for each individual band. This can be attributed primarily to the sampling rate limitations. Although the sampling rate is large enough to capture the full dual-band signal, it is not sufficient to represent the intermodulation products between the two bands. The result is that aliasing occurs and the intermodulation components are “folded-back” on top of the predistortion signal. The presence of these intermodulation components in the predistorted signal degrades the linearization accuracy.

The 2-D DPD architecture mentioned above and outlined in [11] offers an effective solution to this problem. Unlike in typical 2D-DPD applications where the primary motivation is reduced hardware complexity, here we apply the 2D-DPD architecture to allow more accurate modelling of each individual band. The 2D-DPD technique applies predistortion to both signal bands separately but accounts for distortion caused by cross modulation between the bands by including relevant terms from both bands. Distortion outside of the bands of interest is not considered, the objective is to suppress the spectral distortion located around each band. By operating on each individual band separately, the dual-band technique allows the full predistorted signal for each band to be fully represented within the bandwidth permitted by the sampling rate. This avoids the intermodulation product aliasing problems encountered by the conventional SISO models above.

Much of the conventional 2D-DPD architecture, as proposed in [11], is not relevant to this application as the competition hardware already provides the full dual-band spectrum. Figure 4 outlines the structure as used in the competition solution. The input signal is separated into its individual bands before separate predistortion models are applied to each branch. Information from both bands is used in each DPD block to allow compensation of cross modulation distortion.
products. After DPD is applied each band is shifted back to its original position before re-combination and transmission to the PA.

A modified dual-band version of the DVR model is implemented in the DPD 1 and DPD 2 blocks in Figure 4. To reach the maximum achievable linearization performance, the model combines terms from the approaches in [15] and [16], giving the following final structure:

$$
\tilde{y}_i(n) = \sum_{m=0}^{M} a_{\text{w},m} \tilde{x}_i(n-m) + \sum_{k=1}^{K} \sum_{m=0}^{M} c_{\text{lin},k,m} \|\tilde{x}_1(n-m) - \beta_{k,\text{lin},m}\| \tilde{x}_1(n-m) + \sum_{k=1}^{K} \sum_{m=0}^{M} c_{\text{cross},k,m} \|\tilde{x}_1(n-m) - \beta_{k,\text{cross},m}\| \tilde{x}_1(n-m) + \sum_{k=1}^{K} \sum_{m=0}^{M} c_{\text{cross},k,m} \|\tilde{x}_1(n-m) + \tilde{x}_2(n-m) - \beta_{k,\text{cross},m}\| \tilde{x}_1(n-m) + \sum_{k=1}^{K} \sum_{m=0}^{M} c_{\text{cross},k,m} \|\sqrt{\tilde{x}_1(n-m)^2 + \tilde{x}_2(n-m)^2} - \beta_{k,\text{cross},m}\| \tilde{x}_1(n-m)
$$

where $\tilde{x}_1(n)$ denotes a signal sample from the target band and $\tilde{x}_2(n)$ is a sample from the adjacent band. As in the conventional DVR model above, $M$ represents the memory length, $K$ is the number of hyperplanes and the absolute value operations serve as nonlinear basis functions, including both in-band and cross band terms. The DPD model for the alternate band (DPD 2 in Figure 4) can be obtained by swapping $\tilde{x}_1(n)$ and $\tilde{x}_2(n)$ in (3).

**Coefficient Extraction**

The DPD competition rules provide a 15-minute training time in which to calculate the model coefficients before testing the performance. The decision to apply different models for different signal types as described above means that a minimum of three coefficient sets must be extracted during the training period. As in practical applications, these criteria lead to a requirement for a robust, fast, and accurate model extraction routine.
Typically, DPD model extraction is performed using one of two architectures: direct learning (DLA) and indirect learning (IDLA) [17]-[18]. Indirect learning involves training a post-inverse PA model using the PA input and output signals and then transferring the coefficients to the DPD, or pre-inverse, model. The technique offers fast convergence – typically requiring only 2-3 iterations. However, due to the inherent difference between the post- and pre-inverse model input signals it is not possible to completely remove the initial deterministic modelling errors. This leads to a degradation in the extracted coefficients accuracy. By contrast, the direct learning architecture adjusts the coefficients to directly minimize the error between the observed PA output and the original input. In this way, the optimal DPD characteristic can be achieved for a specific DPD model, leaving only random noise and measurement errors. The cost of this improved accuracy is the greater number of iterations required to converge and the risk that the adaptation may be unstable if there is a big difference between the original PA input and observed output [19].

Based on the analysis of strengths and defects for each structure, a dual-loop strategy was proposed to overcome disadvantages in both IDLA and DLA, constructing a more accurate and stable model extraction [19]. Pictured in Figure 5, the dual-loop extraction procedure uses the indirect learning architecture for the initial coarse coefficient extraction and then switches to the direct learning architecture for a number of iterations to fine tune the coefficients. Combining the two in this way the dual loop approach achieves faster convergence to direct learning accuracy levels with robust protection against instability during the initial training phase. In practice the dual-loop technique requires a minor increase in implementation complexity but this is not a concern within the competition format.

**Stabilizing Performance**

As in a real implementation scenario, the exact configuration of the test signal is not known during the training phase in the competition. This makes it necessary for the model and its coefficients to be robust across a range of input signal characteristics. Controlling the predistorted signal peak to average power ratio (PAPR) is particularly important. To achieve the optimum score (or optimum efficiency in a real system) the power amplifier should be driven at the maximum possible output.
power. It is well known that the power amplifier linearity characteristics change at different drive levels. For DPD, this leads that a different set of coefficients are required for each different drive level. Given the limited training time available in competition, training multiple coefficient sets for each of the three possible signal formats is not considered feasible. Thus, a single drive level must be chosen for each signal format during the training stage and maintained throughout the DPD run.

As shown in Figure 6, all three signal formats exhibit some variance in PAPR across different sets of randomly generated test signals. In particular, the dual-band and OFDM-only cases can vary by over 3 dB between sets. To protect the competition test platform, a hard limit of -10 dBm is enforced on the maximum peak input power to the power amplifier. During the coefficient extraction phase a training signal is chosen with PAPR in the region of maximum likelihood within the range of possible values. This signal is used to determine the optimum input power and extract the DPD coefficients.

With a fixed average input power, accurate control of the predistorted signal PAPR is important to ensure the peak input power limit is not exceeded. To control the PAPR and ensure stable performance across the range of possible input signals we employ a two stage PAPR reduction solution based on the scaled peak cancellation method (SPC) outlined in [20]. Figure 7 illustrates the basic operation of SPC, the signal $p_n$ is composed of all components of the input $x_n$ that are greater than the chosen clipping threshold. It is passed through a filter to reduce the out of band components and scaled before being subtracted from the original input signal. The result is that the peaks in the original signal are reduced according to the level of scaling applied to the correction signal, $p_f_n$. If necessary, a number of iterations can be used to achieve the desired output PAPR level, the technique is then referred to as scaled repeated peak cancellation (SRPC).

In the competition solution, we employ two SRPC crest factor reduction (CFR) blocks, as shown in Figure 8. The first block is applied to the input signal prior to its entering the DPD model. The peak threshold is chosen to match that of the training signal used to extract the DPD coefficients. The DPD model itself typically increases the signal PAPR in order to compensate for the gain compression at signal peaks caused by PA saturation. To account for this, after the DPD, a second CFR block is applied to safeguard against excessive PAPR growth and ensure the peak power limitations are met.

![Figure 6. The PAPR probability distributions for each signal type.](image)
before the signal is sent to the power amplifier. Note that this second PAPR model is applied only when the training is complete and the DPD is in “run” mode. Although some techniques, such as [21], have been proposed to apply a single CFR block after DPD without impacting training, for the competition scenario given the limitations on training time, a simpler, robust approach was deemed most appropriate.

Another important consideration to ensure stable performance across the range of possible input signals is the training sequence used to extract the DPD coefficients. For each of the three signal formats, the training sequence PAPR was selected to optimize PA output power without violating the peak power limitations. The final PAPR value represents a trade-off between the goal of limiting signal peaks to increase overall average output power and the requirement to maintain signal integrity to prevent a degradation in overall score in terms of NMSE or out-of-band power which are both impacted by intense clipping.

**Final Predistortion Solution**

Figure 8 shows the final predistortion solution as used in the competition. In both single and dual-band cases the structure is as outlined in Figure 3 with the bands of interest being first shifted to centre frequency before applying the predistortion model. The only difference between single-band and dual-band mode is that the cross terms are used only in the dual band case.

The signal identification block determines the signal type and selects the appropriate DPD model. The signal type is also used to determine the appropriate clipping threshold for the CFR block, the next step in the chain. After the signal PAPR has been controlled the predistorted signal is generated and passed to the final CFR block. Table 2 reports the DPD model parameters for each scenario as used in the competition.
Figure 8. The complete DPD block.

<table>
<thead>
<tr>
<th>Signal type</th>
<th>Linear memory</th>
<th>Nonlinear memory</th>
<th>Number of hyperplanes</th>
<th>Total no. of model terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Band</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>164</td>
</tr>
<tr>
<td>OFDM only</td>
<td>3</td>
<td>3</td>
<td>10</td>
<td>154</td>
</tr>
<tr>
<td>GSM only</td>
<td>3</td>
<td>3</td>
<td>10</td>
<td>154</td>
</tr>
</tbody>
</table>

Table 2. DPD model parameters for each signal format.
Figure 9. The linearized spectra for (a) dual-band, (b) GSM only, and (c) OFDM only signals.
Performance Review and Conclusions

In the competition, the linearization performance was evaluated using three blind test signals covering each possible format for the specified dual band signal: GSM and OFDM, GSM-only, and OFDM-only. Figure 9 reports the performance of the complete predistortion solution for each of these cases. Table 3 lists the performance improvement with DPD for each test case in terms of ACPR and NMSE; output power measured at the VSA (after 40dB attenuation) is also listed.

For the dual-band case illustrated in Figure 9(a), the competition score of 9.08 was slightly lower than expected due to lower output power from the competition device. For the OFDM only case in Figure 9(b), where the GSM carrier is turned off, the algorithm achieved a score of 12.10 in competition, approximately in agreement with our target score based on pre-competition testing. Finally, the GSM-only case shown in Figure 9(c) achieved a score of 15.75 in the competition test. This is lower than expected based on pre-competition testing, the likely cause being insufficient cancellation of out-of-band spectral re-growth at the PA output. This can be attributed to differences between the test signal and the training signal used to extract the DPD coefficients. Summing the above results gives a final score of 36.93 which, although slightly lower than the maximum score obtained in testing, represents strong linearization performance for each of the test scenarios.

This year’s competition proposed a challenging scenario where, in order to obtain a high score, the proposed DPD solution must operate accurately across a range of signal formats. To meet this challenge we used a selective DPD technique using different models for the range of possible signal configurations. To ensure training time limitations were met, training complexity was reduced by selecting a single power level for each signal configuration. A dual-loop coefficient extraction architecture was also employed to maximise extraction accuracy in the available training time. In competition at IMS 2016, the complete DPD solution performed strongly and successfully linearized the provided Doherty power amplifier across all three single configurations.

Acknowledgements

We would like to sincerely thank Mr. Filipe Esturrenho Barradas, Prof. José Carlos Pedro and Prof. Telmo Reis Cunha from the Universidade de Aveiro in Portugal for their work in organizing the competition, including hosting and maintaining the excellent Weblab platform. We also thank Prof. Koen Buisman, Prof. Christian Fagger, and Dr. Per Landin from the GigaHertz Centre at Chalmers University of Technology, Sweden. Finally, thanks to Rohde and Schwarz for their sponsorship of the platform measurement equipment and to the IEEE MTT-11 and MTT-9 Technical Committees for their support of the DPD design competition.

<table>
<thead>
<tr>
<th></th>
<th>Dual-band (Figure 9(a)) (Carrier 1/Carrier 2)</th>
<th>GSM-only (Figure 9(b))</th>
<th>OFDM-only (Figure 9(c))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No DPD</td>
<td>w/ DPD</td>
<td>No DPD</td>
</tr>
<tr>
<td>Output Power (dBW)</td>
<td>3.8/3.8</td>
<td>3.6/3.7</td>
<td>7.6</td>
</tr>
<tr>
<td>ACPR (dB)</td>
<td>25.6/34.4</td>
<td>52.7/52.5</td>
<td>24.3</td>
</tr>
<tr>
<td>NMSE (dB)</td>
<td>-21.6/-23.6</td>
<td>-31.0/-33.5</td>
<td>-20.2</td>
</tr>
</tbody>
</table>

Table 3. ACPR and NMSE improvement with DPD for each signal type.
This work has been supported by research grants from Science Foundation Ireland (SFI) and co-funded under the European Regional Development Fund under Grant Number 13/RC/2077 and Grant Number 12/IA/1267.

REFERENCES


