2016 Student Design Competition Winners IMS

Power Amplifier Linearization Through Digital Predistortion Linearization

Preparing Linearity and Efficiency for 5G

Noel Kelly, Wenhui Cao, and Anding Zhu

IMAGE LICENSED BY INGRAM PUBLISHING

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 maximize information throughput, signal bandwidths are continuously growing. In fourth-generation long-term evolution-advanced (also called 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

Noel Kelly (noel.kelly.1@ucdconnect.ie), Wenhui Cao (wenhui.cao@ucdconnect.ie), and Anding Zhu (anding.zhu@ucd.ie) are with University College Dublin, Ireland.

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 tradeoff 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 tradeoff 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 addresses problems ranging 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, in which a large frequency gap occurs 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. On the one hand, to capture the full transmit bandwidth needed to train a single wideband model would require an extremely wide-band and costly feedback loop. On the other hand, treating each carrier as a separate signal ignores the distortion created 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 DPD solution to maximize the output power of a highly efficient gallium nitride (GaN) high-electron-mobility transistor Doherty PA 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 four-carrier Global System for Mobile Communications (GSM)-like signal and a single-carrier orthogonal frequency-division
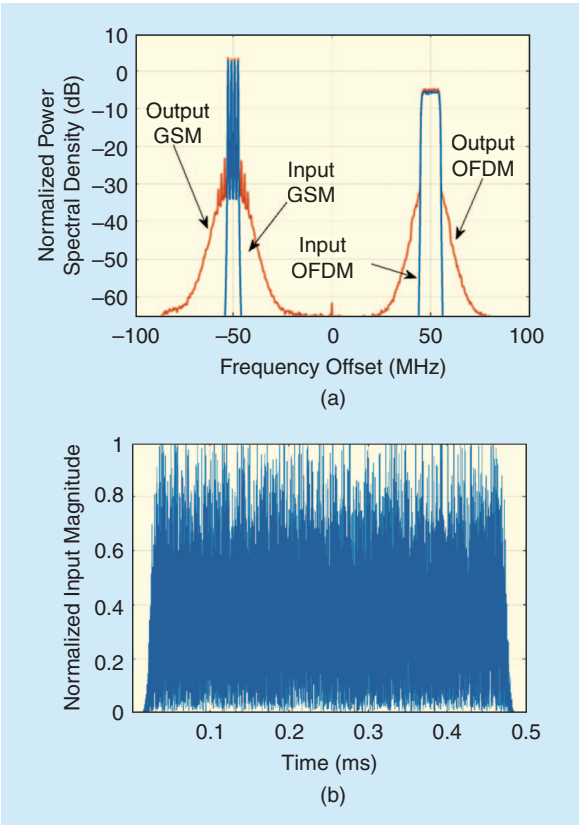

Figure 1. *The characteristics of the dual-band test signal: (a) frequency-domain normalized power spectral density and (b) time-domain normalized amplitude.*

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 amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) characteristics of the complete transmitter chain excited by a dual-band signal with input power of –21 dBm. As can be seen, in addition to the nonlinear behavior, the wide bandwidth also leads to substantial memory effects.

Scoring Metrics
The competing DPD solutions were evaluated using a scoring metric designed to account for three parameters

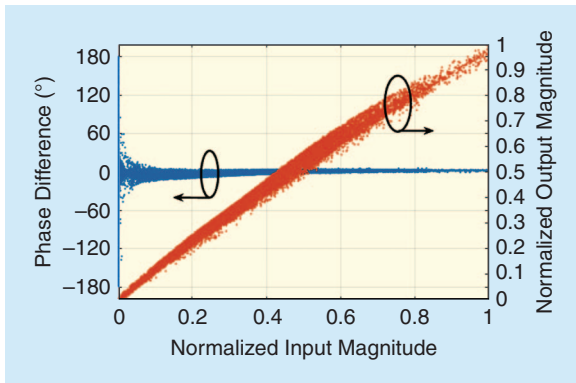| TABLE 1. The test signal characteristics.		
	Band 1	**Band 2**
Signal Type	**GSM**	**OFDM**
Frequency offset	–50 MHz	+50 MHz
Number of carriers	4	1
Carrier bandwidth	200 kHz	10 MHz
Carrier spacing	1.67 MHz	—
Total bandwidth	5.2 MHz	10 MHz

Figure 2. *The measured AM/AM and AM/PM characteristics.*

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

$$S = \sum_{i=1}^{3} \sum_{b=1}^{N_a}$$
$$\times \frac{(P_{av,i,b} - 3) + (IMD_{i,b} - 45) + \min(0, -30 - NMSE_{i,b})}{N_a},$$
(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 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 downgraded if the signal measures worse than any of the above limits.

Measurement Platform
The competition measurement platform used a Rohde & Schwarz (R&S) SMW200A vector-signal generator

(VSG) and FSW8 vector-signal analyzer (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 upconverted 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 downconverted 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 PA 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 suboptimal performance for each individual case.

The need for separate models and coefficient sets to maximize performance, despite the physical PA remaining unchanged, stems from an important nuance of the DPD technique: the optimum DPD
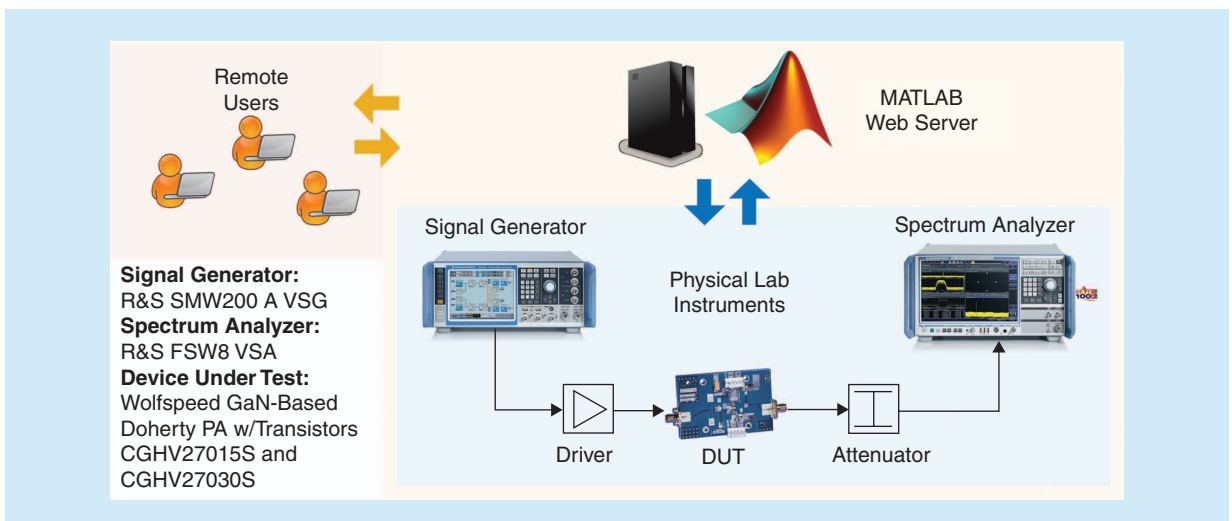

Signal Generator:
R&S SMW200 A VSG
Spectrum Analyzer:
R&S FSW8 VSA
Device Under Test:
Wolfspeed GaN-Based
Doherty PA w/Transistors
CGHV27015S and
CGHV27030S

Figure 3. *The remote measurement platform setup.*

model coefficients extracted for one signal type can be very different from those of a different signal type, even when the power levels and device under test (DUT) remain unchanged.

A universal optimum set of coefficients does 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 modeling 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's 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.

Thus, overall performance can be improved by tailoring different DPD models and their corresponding coefficient sets to suit the needs of each test scenario. 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

$$
\begin{aligned}
\tilde{y}(n) = & \sum_{i=0}^{M} a_i \tilde{x}(n-i) \\
& + \sum_{k=1}^{K} \sum_{i=0}^{M} c_{k,i,1} \big| |\tilde{x}(n-i)| - \beta_k \big| e^{j\theta(n-i)} \\
& + \sum_{k=1}^{K} \sum_{i=0}^{M} c_{k,i,21} \big| |\tilde{x}(n-i)| - \beta_k \big| e^{j\theta(n-i)} |\tilde{x}(n)| \\
& + \cdots,
\end{aligned}
\tag{2}
$$

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.

Unlike typical 2D-DPD applications where the primary motivation is to reduce hardware complexity, here we apply a new 2D-DPD architecture to allow more accurate modeling of each individual band.

This is both easily implemented in hardware and better suited to modeling 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, and 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 previously and outlined in [11] offers an effective solution to this problem. Unlike typical 2D-DPD applications where the primary motivation is to reduce hardware complexity, here we apply a new 2D-DPD architecture to allow more accurate modeling 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 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
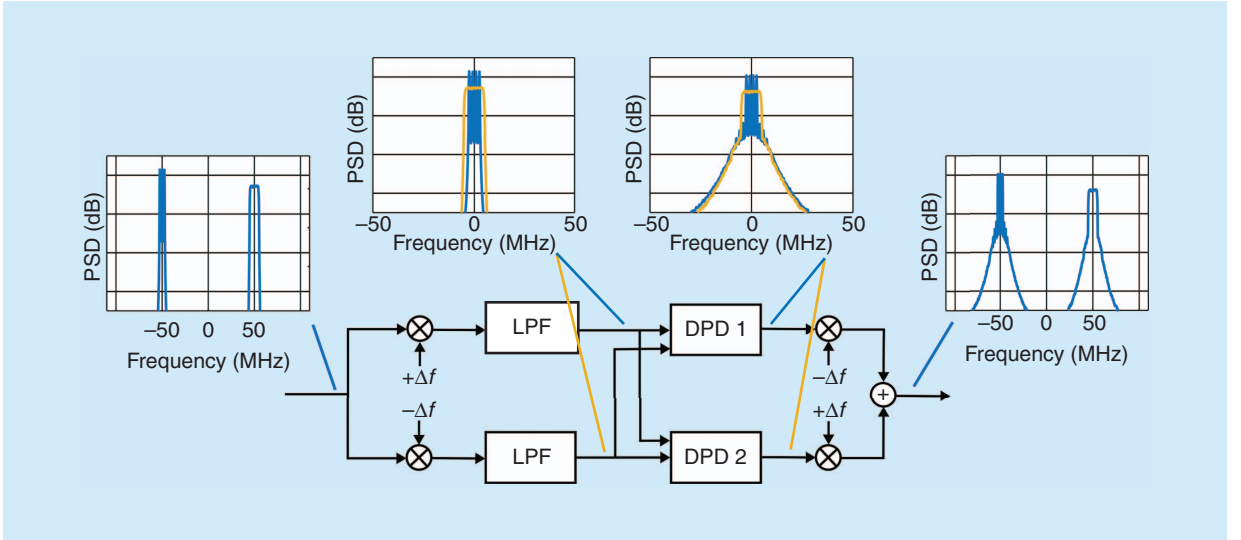
Figure 4. *The dual-band DPD structure. LPF: low-pass filter.*

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 mentioned previously.

Much of the conventional 2D-DPD architecture, as proposed in [11], is not relevant to this application because 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 recombination 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 in (3), 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 described previously, 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-min 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 previously, 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

$$
\begin{aligned}
\tilde{y}_1(n) = &\sum_{m=0}^{M} a_{1,m}\tilde{x}_1(n-m) \\
&+ \sum_{k=1}^{K}\sum_{m=0}^{M} c_{1,km,\mathrm{in}}\left|\left|\tilde{x}_1(n-m)\right|-\beta_{1,k,\mathrm{in}}\right|\tilde{x}_1(n-m) \\
&+ \sum_{k=1}^{K}\sum_{m=0}^{M} c_{1,km,\mathrm{cross},1}\left|\left|\tilde{x}_2(n-m)\right|-\beta_{2,k,\mathrm{cross},1}\right|\tilde{x}_1(n-m) \\
&+ \sum_{k=1}^{K}\sum_{m=0}^{M} c_{1,km,\mathrm{cross},2}\left|\left|\tilde{x}_1(n-m)\right|+\left|\tilde{x}_2(n-m)\right|-\beta_{2,k,\mathrm{cross},2}\right|\tilde{x}_1(n-m) \\
&+ \sum_{k=1}^{K}\sum_{m=0}^{M} c_{1,km,\mathrm{cross},3}\left|\left|\tilde{x}_1(n-m)\right|-\left|\tilde{x}_2(n-m)\right|-\beta_{2,k,\mathrm{cross},3}\right|\tilde{x}_1(n-m) \\
&+ \sum_{k=1}^{K}\sum_{m=0}^{M} c_{1,km,\mathrm{cross},4}\left|\sqrt{\left|\tilde{x}_1(n-m)\right|^2+\left|\tilde{x}_2(n-m)\right|^2}-\beta_{2,k,\mathrm{cross},4}\right|\tilde{x}_1(n-m) \quad (3)
\end{aligned}
$$

indirect learning (IDLA) [17], [18]. IDLA 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 two to three 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 modeling errors. This leads to a degradation in the extracted coefficients accuracy.

Figure 5. *The dual-loop extraction architecture [19].*

By contrast, the DLA 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 the disadvantages in both IDLA and DLA, thus constructing a more accurate and stable model extraction [19]. Pictured in Figure 5, the dual-loop extraction procedure uses the IDLA for the initial coarse coefficient extraction and then switches to the DLA 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 competition's training phase. 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 PA should be driven at the maximum possible output power. It is well known that PA linearity characteri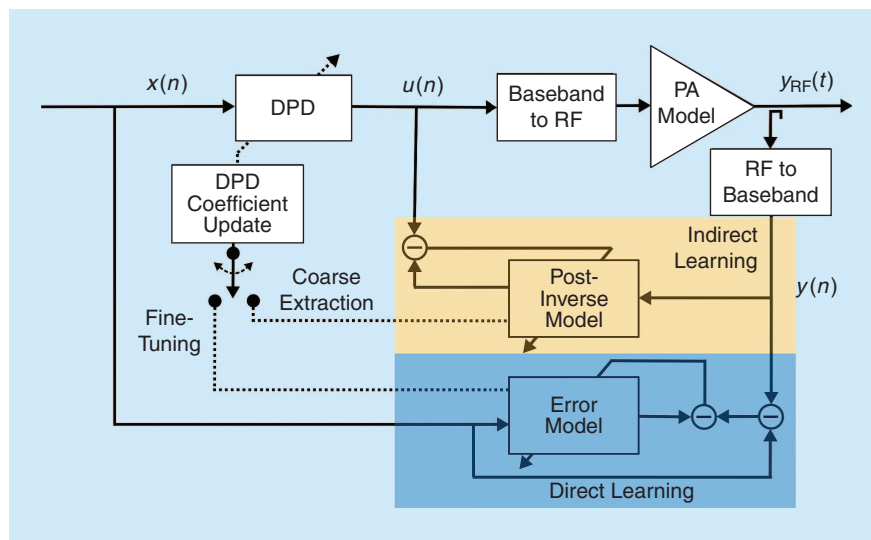stics change at different drive levels. For DPD, this means that a different set of coefficients is required for each different drive level. Given the limited training time available in competition, training multiple coefficient sets for each of the three possible signal formats is not considered feasible. 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 PA. 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's PAPR is important to ensure that the peak input power limit is not exceeded.
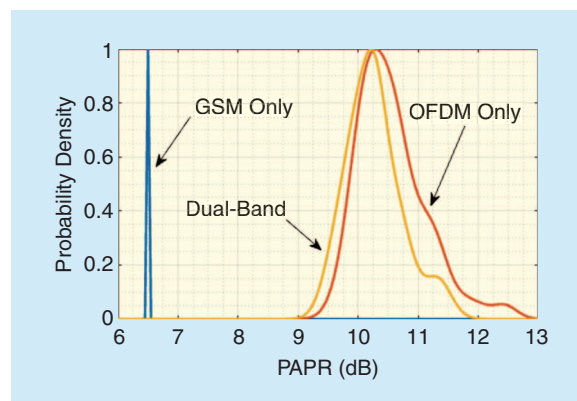

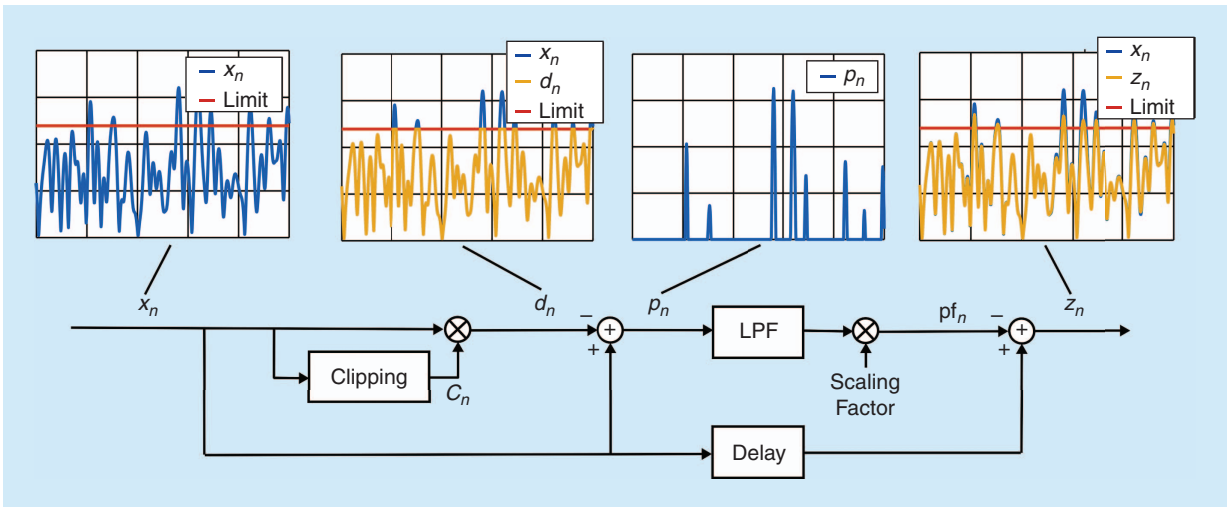**Figure 6.** *The PAPR probability distributions for each signal type.*

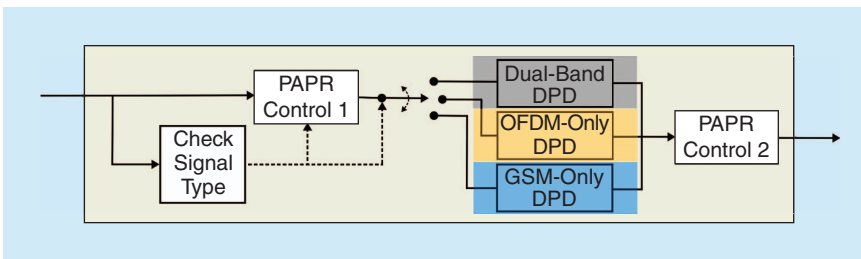**Figure 7.** *The CFR using the SPC method.*

Figure 8. *The complete DPD block.*

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, pf_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 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 that the peak power limitations are met before the signal is sent to the PA. 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 that in [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 tradeoff

TABLE 2. The DPD model parameters for each signal format.

Signal Type	Linear Memory	Nonlinear Memory	Number of Hyperplanes	Total Number of Model Terms
Dual band	3	3	8	164
OFDM only	3	3	10	154
GSM only	3	3	10	154

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, both of which are impacted by intense clipping.

Final Predistortion Solution

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

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 as measured at the VSA (after 40-dB 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 precompetition 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 precompetition testing, the likely cause being insufficient
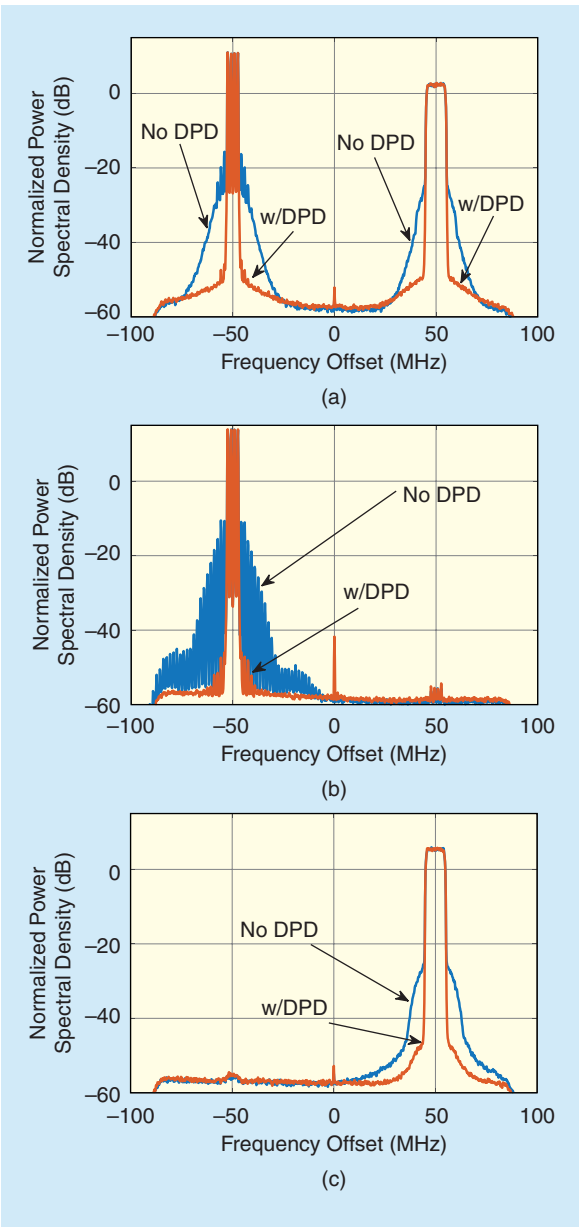

Figure 9. *The linearized spectra for (a) dual-band, (b) GSM-only, and (c) OFDM-only signals.*

cancellation of out-of-band spectral regrowth 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

---

TABLE 3. The ACPR and NMSE improvement with DPD for each signal type.

	Dual-Band [Figure 9(a)] (Carrier 1/Carrier 2)		GSM Only [Figure 9(b)]		OFDM Only [Figure 9(c)]	
	No DPD	**With DPD**	**No DPD**	**With DPD**	**No DPD**	**With DPD**
Output power (dBW)	3.8/3.8	3.6/3.7	7.6	7.4	6.4	6.3
ACPR (dB)	25.6/34.4	52.7/52.5	24.3	55.7	36.2	56.3
NMSE (dB)	−21.6/−23.6	−31.0/−33.5	−20.2	−45.8	−26.8	−38.4

The final PAPR value represents a tradeoff 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, both of which are impacted by intense clipping.

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 in which, 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 maximize extraction accuracy in the available training time. In competition at IMS2016, the complete DPD solution performed strongly and successfully linearized the provided Doherty PA across all three single configurations.

Acknowledgements

References

[1] G. Yuan, X. Zhang, W. Wang, and Y. Yang, "Carrier aggregation for LTE-advanced mobile communication systems," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 88–93, Feb. 2010.

[2] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 36–43, May 2014.

[3] P. Asbeck and Z. Popovic, "ET comes of age: Envelope tracking for higher-efficiency power amplifiers," *IEEE Microwave Mag.*, vol. 17, no. 3, pp. 16–25, Mar. 2016.

[4] T. Barton, "Not just a phase: Outphasing power amplifiers," *IEEE Microwave Mag.*, vol. 17, no. 2, pp. 18–31, Feb. 2016.

[5] B. Kim, J. Kim, I. Kim, and J. Cha, "The Doherty power amplifier," *IEEE Microwave Mag.*, vol. 7, no. 5, pp. 42–50, Oct. 2006.

[6] P. M. Lavrador, T. R. Cunha, P. M. Cabral, and J. C. Pedro, "The linearity-efficiency compromise," *IEEE Microwave Mag.*, vol. 11, no. 5, pp. 44–58, Aug. 2010.

[7] J. Wood, *Behavioral Modeling and Linearization of RF Power Amplifiers*. Norwood, MA: Artech House, 2014.

[8] L. Guan and A. Zhu, "Green communications: Digital predistortion for wideband RF power amplifiers," *IEEE Microwave Mag.*, vol. 15, no. 7, pp. 84–99, Nov./Dec. 2014.

[9] A. Zhu, "Decomposed vector rotation-based behavioral modeling for digital predistortion of RF power amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 63, no. 2, pp. 737–744, Feb. 2015.

[10] S. A. Bassam, W. Chen, M. Helaoui, and F. M. Ghannouchi, "Transmitter architecture for CA: Carrier aggregation in LTE-advanced systems," *IEEE Microwave Mag.*, vol. 14, no. 5, pp. 78–86, July/Aug. 2013.

[11] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *IEEE Trans. Microwave Theory Tech.*, vol. 59, no. 10, pp. 2547–2553, Oct. 2011.

[12] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[13] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[14] A. Zhu, J. C. Pedro, and T. J. Brazil, "Dynamic deviation reduction-based Volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 12, pp. 4323–4332, Dec. 2006.

[15] Y. Guo, "Digital predistortion techniques for RF power amplifiers in LTE-Advanced mobile communication systems," Ph.D. dissertation, School of Electron. Elect. Eng., Univ. College Dublin, Dublin, Ireland, 2016.

[16] C. Wang and X. Zhu, "A nonlinear filter-based model for concurrent dual-band power amplifiers," in *Proc. IEEE Int. Wireless Symp.*, Shenzhen, 2015, pp. 1–4.

[17] P. Draxler, J. Deng, D. Kimball, I. Langmore, and P. M. Asbeck, "Memory effect evaluation and predistortion of power amplifiers," in *Proc. IEEE MTT-S Int. Conf. Dig.*, Long Beach, CA, 2005, pp. 1549–1552.

[18] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[19] L. Guan and A. Zhu, "Dual-loop model extraction for digital predistortion of wideband RF power amplifiers," *IEEE Microwave Compon. Lett.*, vol. 21, no. 9, Sept. 2011.

[20] W.-J. Kim, K.-J. Cho, S. P. Stapleton, and J.-H. Kim, "An efficient crest factor reduction technique for wideband applications," *Analog. Integr. Circuits Signal Process.*, vol. 51, no. 1, pp. 19–26 Jan. 2007.

[21] R. N. Braithwaite, "A combined approach to digital predistortion and crest factor reduction for the linearization of an RF power amplifier," *IEEE Trans. Microwave Theory Tech.*, vol. 61, no. 1, pp. 291–302, Jan. 2013.