Predistortion of OFDM Waveforms using Guard-band Subcarriers

Chance Tarver*, Alexios Balatsoukas-Stimming†, and Joseph R. Cavallaro*
*Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA
†Electrical Engineering Department, Eindhoven University of Technology, The Netherlands

2020 54th Asilomar Conference on Signals, Systems, and Computers | 978-0-7381-3126-9/20/$31.00 ©2020 IEEE | DOI: 10.1109/IEEECONF51394.2020.9443468

***Abstract*—Digital predistortion (DPD) is an important technique that is commonly used in wireless transmitters to reduce the out-of-band emissions caused by power amplifier (PA) nonlinearities. This frequency-domain goal is generally achieved by learning a baseband equivalent time-domain inverse transfer function of the PA and applying it to the transmitted digital baseband signal. In this work, we take advantage of the frequency-domain nature of orthogonal frequency-division multiplexing (OFDM) signals by injecting cancellation tones into the guard-band subcarriers to perform DPD. We experimentally evaluate our OFDM-based DPD (ODPD) method using a Doherty PA and show that, when compared to a standard generalized memory polynomial solution, the ODPD can achieve better suppression of out-of-band emissions with lower complexity. Moreover, when combined with a neural network model of the PA, our proposed ODPD method only requires oversampling the transmitted signal by a factor of two, which has important implications for the analog transceiver front-end.**

***Index Terms*—Digital predistortion (DPD), orthogonal frequency-division multiplexing (OFDM), neural networks.**

I. Introduction

In wireless communications systems, power amplifiers (PAs) are essential components that allow us to increase the range and the signal-to-noise ratio (SNR) of the employed signals. However, PAs typically introduce nonlinearities which can cause adjacent channel leakage and can degrade the error vector magnitude (EVM). The acceptable adjacent channel leakage, which is measured using the adjacent channel leakage ratio (ACLR), is strictly defined in most standards, such as 5G New Radio (NR). Beyond complying with standards, regulatory bodies, such as the Federal Communications Commission (FCC), will also legally enforce certain ACLR requirements.

One way to avoid the PA nonlinearities is to back off the input power so that the PA operates in its linear region. However, this is typically undesired because PAs are less power efficient in this region of operation. Moreover, the high peak-to-average power ratio (PAPR) of orthogonal frequency-division multiplexing (OFDM) waveforms that most modern communications systems use requires extreme amounts of backoff to account for the worst-case scenario.

A popular alternative solution is to predistort the digital baseband signal with the inverse of the PA nonlinearities so that the cascade of the two systems becomes linear. This widely used technique is called digital predistortion (DPD), and it typically involves taking the digital baseband signal and passing it through some variant of a memory polynomial (MP) [1] or generalized memory polynomial (GMP) [2]. The proper DPD coefficients can be learned through a system such as an indirect learning architecture (ILA) [3]. Neural networks (NNs) have also been shown throughout the literature to be suitable for predistortion applications [4], [5].

Even though DPD typically operates in the time domain, most modern communications systems (such as 5G NR, 4G Long-Term Evolution (LTE), and Wi-Fi) use OFDM-based signals, which are constructed in the frequency domain. Moreover, the ACLR, which is the principal figure of merit in the context of DPD, is, in fact, a frequency-domain metric.

In this work, we introduce ODPD, a novel DPD method for OFDM waveforms that exploits the guard-band subcarriers that are typically present in OFDM-based systems. In particular, instead of transmitting zeros on the guard-band subcarriers, we iteratively tune their values in order to reduce the ACLR on a per-OFDM-symbol basis directly. To determine the appropriate values of the guard-band subcarriers, one needs only a forward model of the PA, as opposed to the inverse model needed in most DPD solutions. Our experimental measurements using a commercial Doherty PA platform, widely regarded for 5G applications such as massive MIMO [6], shows that we can achieve linearization that outperforms a state-of-the-art polynomial model. Moreover, when combined with a NN-based PA model, our OFDM-based DPD can perform DPD with as little as a 2x upsampling/oversampling rate[1] as opposed to the 5x upsample rate typically considered in polynomial-based DPD solutions [2], [7]. The lower sampling rate translates into lower energy consumption and reduced system complexity.

Similar ideas can be found in the literature, though not directly applied for DPD. For example, a similar technique is used for PAPR reduction in [8]. In [9], the authors also utilize the guardbands for *cancellation carriers*, but their goal is OFDM sidelobe suppression and not correction of PA nonlinearities. The learning iterations of our method are

This work was supported in part by Xilinx, Samsung, and by the US NSF under grants CNS-1717218, CNS-2016727, and CNS-1827940, for the "PAWR Platform POWDER-RENEW: A Platform for Open Wireless Data-driven Experimental Research with Massive MIMO Capabilities."

[1]While upsampling refers to the DSP process of interpolating and oversampling refers to the relative rate of the ADC sampling rate to a signal's Nyquist rate, we will use the terms interchangeably throughout the paper for brevity.

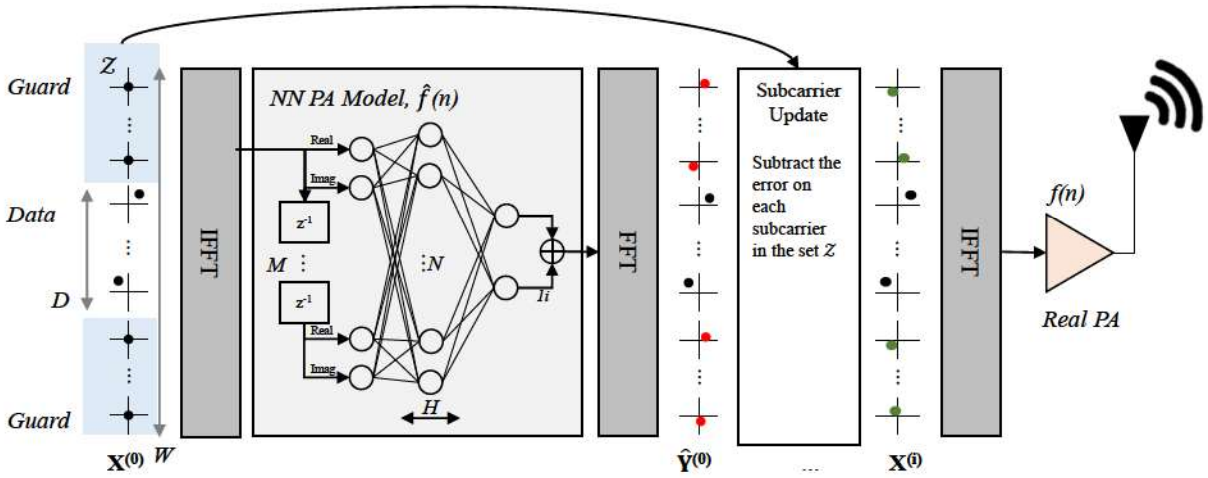

Figure 1. Overview of ODPD algorithm. For each symbol, we transmit through a PA model, $\hat{f}$, to see the error in the guardband subcarriers in the set $\mathcal{Z}$. The error is subtracted to form the new frequency domain input. After sufficient iterations, the symbol can be transmitted through the real PA. In this figure, we omit the DAC and up-converter after the final IFFT for simplicity.

similar to the iterative learning control (ILC) DPD method used in [10]. However, in our method, we adapt it to operate directly on the guard band subcarriers in the frequency domain.

II. Digital Predistortion

Time-domain DPD is widely used to linearize PAs, which reduces out-of-band (OOB) emissions and improves the power-added efficiency. While there are various algorithms and inverse models that can be used, one popular choice is the GMP [2], which is defined as

$$\begin{aligned}\hat{x}(n) = &\sum_{p=1}^{P}\sum_{m=0}^{M-1} a_{p,m}x(n-m)\,|x(n-m)|^{p-1} + \\ &\sum_{p=1}^{P}\sum_{m=0}^{M-1}\sum_{l=1}^{L} b_{p,m,l}x(n-m)\,|x(n-m-l)|^{p-1} + \\ &\sum_{p=1}^{P}\sum_{m=0}^{M-1}\sum_{l=1}^{L} c_{p,m,l}x(n-m)\,|x(n-m+l)|^{p-1}.\end{aligned} \tag{1}$$

In (1), p indexes polynomial orders, m indexes a memory depth, and l indexes a lag/lead cross-term. The highest nonlinearity order is P with a maximum of M memory taps and a maximum crossterm lag/lead depth of L samples. This model has remained dominant in industry and academia because it is linear in terms of the coefficients, $a_{p,m}$, $b_{p,m}$, and $c_{p,m}$.

To solve for the coefficients, a common technique is the ILA [1]–[3]. With an ILA, it is assumed that the GMP model from the PA output, $y(n)$ to the PA input is equivalent to the system's inverse. Hence the linearity of the model in terms of the parameters can be leveraged through a least-squares fit from $y(n)/G$ to $x(n)$, where G is the gain of the system. However, there are shortcomings to this approach in that when sent through the GMP, a signal with noise, $y(n)$, enters an absolute value function leading to the possibility of bias [11].

As the PA enters deeper levels of saturation, higher-order polynomials are typically needed. However, the bandwidth of a signal expands with the polynomial order leading to aliasing of the signal if sufficient oversampling is not applied [2]. For this reason, it is typical to see GMPs deployed with at least 5x oversampling, which can lead to high system complexity for wide bandwidth signals considered in 5G NR and beyond.

III. OFDM Digital Predistortion

Let $\mathbf{X} \in \mathbb{C}^W$ denote a vector of symbols that correspond to one OFDM symbol. $W \in \mathbb{N}$ denotes the total number of subcarriers, which is typically a power of two for efficient fast Fourier transform (FFT) computations. In OFDM systems, $D \in \mathbb{N}$ where $D < W$ subcarriers carry information, and there are typically $K = W - D$ subcarriers that map to the edge of the spectrum, which are zero-filled and which form the *guard band*. Let the set of zero-filled subcarriers in the guard band be denoted by $\mathcal{Z} \subset \{0, \ldots, W-1\}$. Our proposed method's key idea is to replace the zero-filled subcarriers with tuned values that depend on the remaining D data subcarriers to reduce the OOB emissions directly.

For simplicity, we restrict our description to a single OFDM symbol. However, our method can be extended to multiple symbols by applying ODPD for each symbol and relying on the windowing technique typically applied in OFDM systems to improve the spectrum at symbol boundaries [12]. Let $f(\cdot)$ denote the baseband equivalent of the nonlinear PA transfer function. Then, the frequency domain output of the PA, denoted by $\mathbf{Y}$, is [2]

$$\mathbf{Y} = \text{FFT}\left(f\left(\text{IFFT}(\mathbf{X})\right)\right). \tag{2}$$

Assuming that an estimate of the PA transfer function, $\hat{f}(\cdot)$, is created, we can iteratively estimate for each subcarrier $k \in \mathcal{Z}$ the PA output, $\hat{Y}_k$, and predistort it to cancel out the tone. By iterating on each subcarrier, we heuristically can

[2]The cyclic prefix (CP) is also added after the IFFT and removed before the FFT. We do not model the CP here, though a 4.7 μs CP is used in the final results.

think of this as injecting the subcarrier with the energy of the opposite phase so that there can be a net cancellation. However, the exact value depends on the intermodulations of all subcarriers. Hence multiple iterations may be needed to account for new intermodulations due to previous iterations. At iteration $i \in \{0, \ldots, I-1\}$, in our proposed method, we first use (2) to calculate $\hat{\mathbf{Y}}^{(i)}$ based on $\mathbf{X}^{(i)}$. Then, we calculate the differences between each $\hat{Y}_k^{(i)}$, $k \in \mathcal{Z}$, and the desired output $\hat{X}_k$, $k \in \mathcal{Z}$. Finally, we adapt each guard tone $X_k^{(i+1)}$, $k \in \mathcal{Z}$, as follows

$$X_k^{(i+1)} = X_k^{(i)} - \mu \hat{Y}_k^{(i)}, \quad \forall k \in \mathcal{Z}, \tag{3}$$

where μ is a step size and $\mathbf{X}^{(0)} = \mathbf{X}$. This algorithm is illustrated in Fig. 1, where a NN is used to model the PA.

The guard band subcarriers' energy after the PA nonlinearities depends on the intermodulation products of the data subcarriers in each OFDM symbol. As such, this learning algorithm must be executed for each OFDM symbol. However, the learning by broadcasting through a PA model is conceptually no different than broadcasting a signal through a trained DPD block to apply predistortion. The difference here is the addition of an IFFT and an FFT before and after this PA model, which can be efficiently implemented in various platforms potentially through the reuse of existing accelerators already in the ODFM signal processing chain, and the addition of a subtraction on each subcarrier. The complexity increases linearly with the number of iterations, but excellent performance can be achieved even with a single iteration, as we will show in Section V.

There are multiple advantages to our method over other DPD methods. In particular, our ODPD does not require an inverse model of the PA. Instead, an easily obtained forward model of the device can be used. Moreover, we eliminate sensitive ILA systems that may converge to biased DPD training solutions [11]. When combined with a NN, additional benefits can be achieved, such as predistortion at a lower sampling rate than conventional DPD solutions.

IV. Forward PA Models

The PA model $\hat{f}(\cdot)$ can be constructed through various methods. For example, a GMP or a NN [5] may be attractive solutions. Moreover, in certain time division duplex (TDD) systems, it could be possible to train symbols using the actual PA while the radio is listening. While in this work we highlight the use of an NN-based PA model, we briefly discuss a GMP implementation in the following subsection.

A. GMP PA Model

The GMP from (1) can be used as a forward model of the PA, $\hat{f}(n)$. When using the GMP, a least-squares model can also be used to learn the set of parameters. However, contrary to the ILA approach, the forward model is not as susceptible to noise. A fundamental limit of ODPD performance is the accuracy of, $\hat{f}$. Hence, sufficient upsampling would be required to avoid aliasing of high order terms when using a GMP.

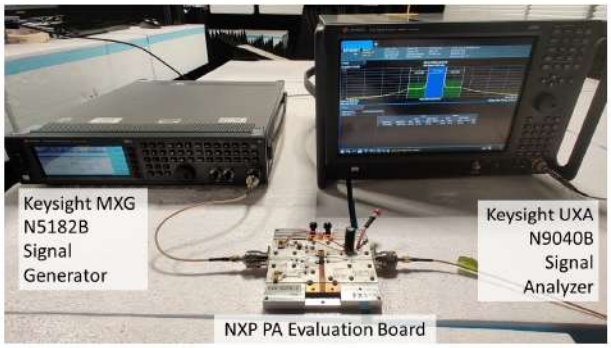

Figure 2. Photo of the measurement setup. A signal generated in MATLAB is uploaded to the signal generator, where it is transmitted at 3.5 GHz through the PA evaluation board into the UXA where the ACLR is measured.

B. Neural Network PA Model

While high-order polynomials suffer from aliasing when using low sample rates [2], [7], NNs do not necessarily have the same limitations as they are model-free.

We consider a multilayer feedforward NN with H hidden layers and N neurons in each hidden layer. M time-domain inputs are given to the network to account for memory effects in the PA. For each sample, the real and imaginary components enter the NN on separate neurons. The general architecture is shown within the ODPD system in Fig. 1.

Let g denote a nonlinear activation function, and let $\mathbf{W}_i$ and $\mathbf{b}_i$ denote the weights matrices and bias vectors corresponding to the ith layer in the NN. The output of the first hidden layer at time instant n is

$$\mathbf{h}_1(n) = g\left(\mathbf{W}_1 \begin{bmatrix} \Re(x(n)) \\ \Im(x(n)) \\ \vdots \\ \Re(x(n-M+1)) \\ \Im(x(n-M+1)) \end{bmatrix} + \mathbf{b}_1\right). \tag{4}$$

The output of hidden layer $i \geq 2$ is

$$\mathbf{h}_i(n) = g\left(\mathbf{W}_i \mathbf{h}_{i-1}(n) + \mathbf{b}_i\right). \tag{5}$$

Finally, the output of the network after hidden layer H is

$$\hat{\mathbf{x}}(n) = \mathbf{W}_{H+1}\mathbf{h}_H + \mathbf{b}_{H+1}, \tag{6}$$

where the first and second elements of $\hat{\mathbf{x}}$ represent the real and imaginary part of the signal, respectively.

Complexity remains low when considering a ReLU activation function, which can be implemented with a simple multiplexer. To further reduce the computational burden, a designer could consider options such as pruning and quantization.

C. Computational Complexity

The computational complexity of the ODPD scheme can be divided into two components, the complexity of the iterative application of IFFTs and FFTs, and the complexity of the forward model, $\hat{f}$. We consider the number of real multiplications

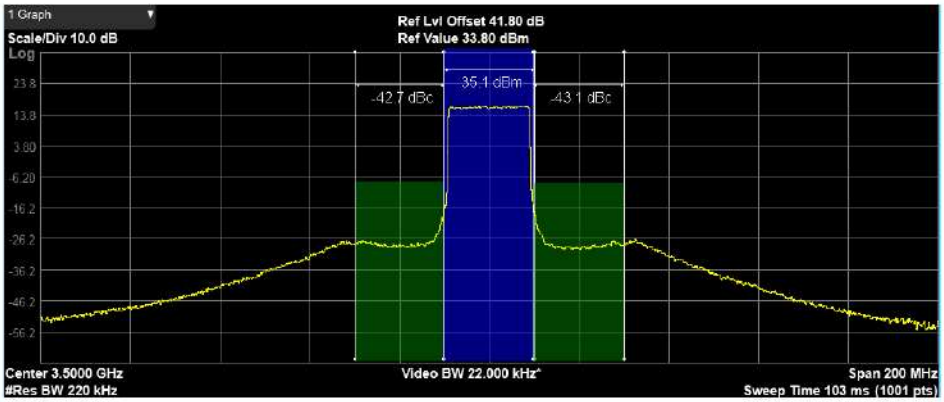

Figure 3. Measurement result from the Doherty PA for the ODPD with $I = 2$ iterations and $N = 40$ neurons. The blue is the 20 MHz associated with the main carrier while the green on the left and right correspond to the 20 MHz adjacent channels.

Table I
ACLR MEASUREMENTS FROM PA TESTBED

Case	L1 (dBc)	Main (dBm)	R1 (dBc)
No DPD	-30.8	35.3	-30.6
GMP			
$U = 2$	-30.5	35.2	-29.6
$U = 3$	-37.8	35.2	-36.4
$U = 4$	-41.3	35.2	-41.0
$U = 5$	-41.3	35.2	-41.3
NN, $U = 2$			
$N = 10, I = 1$	-37.2	35.2	-36.7
$N = 10, I = 2$	-38.0	35.2	-37.8
$N = 20, I = 1$	-39.1	35.3	-38.7
$N = 20, I = 2$	-40.1	35.2	-40.3
$N = 40, I = 1$	-41.7	35.2	-41.2
$N = 40, I = 2$	-42.7	35.1	-43.1

as a proxy for the complexity of the ODPD and count them as

$$n_{\text{mults, FFT}} = 4IW\log_2(W), \tag{7}$$

$$n_{\text{mults, ODPD}} = n_{\text{mults, FFT}} + n_{\text{mults},\hat{f}}. \tag{8}$$

Here, $n_{\text{mults, FFT}}$ counts the number of multiplications due to the added FFTs and $n_{\text{mults},\hat{f}}$ is the forward model complexity. We assume four real multiplies per complex multiply.

V. RESULTS

In this section, we present experimental results to showcase the performance of our proposed ODPD method, and we compare it with a standard polynomial-based DPD method.

A. Measurement Setup

We test our proposed algorithm using a Doherty PA from NXP [13], transmitting at 3.5 GHz. We use a Keysight MXG signal generator to drive the PA and a Keysight UXA signal analyzer to record the IQ data for DPD learning and ACLR measurements. A photo of the PA testbed is shown in Fig. 2. For the ACLR, we measure the integrated power of the 20 MHz band on each side of the primary carrier, corresponding to the first channel to left and right of the main carrier, denoted as the L1 and U1 channels, respectively.

We test the algorithm for an input power of 6 dBm on the signal generator. IQ data was collected at oversampling rates of $U = 2, 3, 4$, and 5. This data was used to learn a GMP predistorter with $P = 7, M = 4$, and $L = 1$.

The $U = 2$ IQ data was used to build three separate NN-based forward models, $\hat{f}(\cdot)$, of the PA. The sampling rate of the ADC measuring the PA output is set to 61.44 MSps, which corresponds to a 2x oversampling rate from the modulator's 30.72 MSps rate. For each NN, we use $M = 4$ and $H = 1$. We set N to 10, 20, and 40 to partially explore possible complexity versus performance trandeoffs in the NN architectures. We then applied the ODPD using $I = 1$ and $I = 2$ with $\mu = 1$. For the transmit signal, we use $D = 1200$ and $W = 4096$ for 14 symbols that are distinct from the data used while training. A cyclic prefix of length 4.6 μs is also applied on each symbol.

B. Measurement Results

An example measurement using the ODPD method with a $N = 40$ NN is shown in Fig. 3 where $I = 1$. The input power for this test was 6 dBm, the in-band PA output power was 35.3 dBm (which corresponds to 29.3 dB gain from the PA).

Table I shows the full measurement results for each DPD. Here, we show that the Doherty PA was highly nonlinear with a starting ACLR of -30.8 dBc on the left adjacent carrier. At a low oversampling rate of $U = 2$, the GMP is unable to resolve the high-order nonlinearities leading to an overall poor fit that causes the ACLR to degrade. At $U \geq 3$ the ACLR was able to improve with the GMP.

However, the NN-based ODPD was able to improve the ACLR for each considered architecture while only using an oversampling rate of $U = 2$. Moreover, the $N = 40$ architecture is able to outperform the best $U = 5$ case from the GMP.

The performance of the ODPD depends on the precision of the forward model. However, there is a tradeoff between complexity and precision. Therefore, careful tuning of the

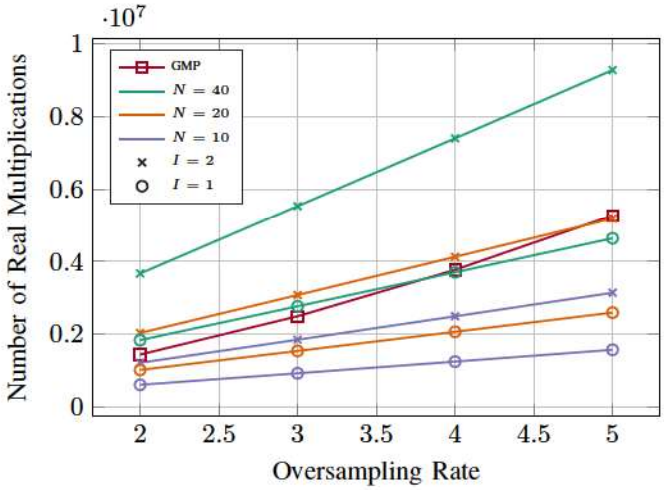

Figure 4. Complexity per symbol versus oversampling rate. Each NN uses $M = 4$ and $H = 1$. The GMP uses $P = 7, M = 4, L = 1$.

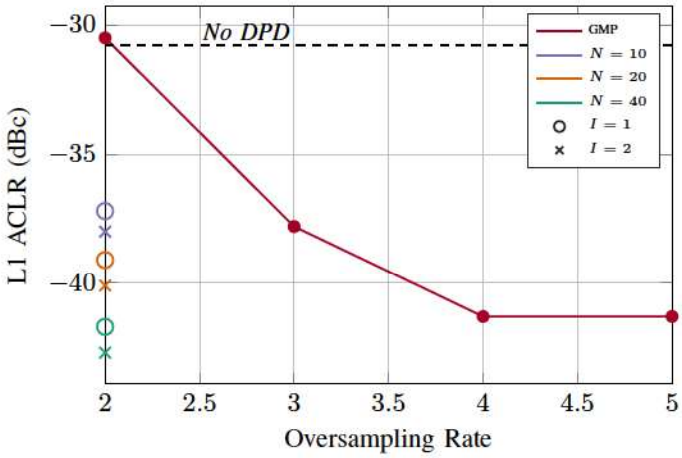

Figure 5. ACLR performance for each scheme. The NN-based ODPD is able to achieve an ACLR improvement with only 2x upsampling.

forward model is necessary. While there are many viable architectures of NNs that can be tested, we restrict our analysis to three. We restrict the NN so that $M = 4$ and $H = 1$ and vary the number of neurons in the hidden layer to be 10, 20, 40. Using the PA input/output data collected from the testbed, we train each neural network in MATLAB.

In Fig. 4, we show the complexity in terms of the number of real multiplications per OFDM symbol of the ODPD algorithm for the three considered NNs with $I = 1$ and $I = 2$ total iterations. We compare this to the GMP ILA-DPD application complexity in red which is obtained by summing the number of multiplications in (1). The ODPD cases include the complexity of each new FFT and the PA model. The GMP ILA-DPD case includes the complexity of the GMP as well as the upsampling complexity. Upsampling is assumed to be done via filling with zeros and passing through a low-pass-filter with 51 taps.

Fig. 5 shows the performance from Table I for the sake of comparison to Fig. 4. When trying to get the most performance per computation, there are a few takeaways. Firstly, the performance of the $U = 5$ GMP can be matched by a $U = 2$ NN-based ODPD with 34.8% the number of multiplications. Secondly, it can be seen in Fig. 4 that for the considered NN architectures, a larger neural network architecture was often less computationally intensive than more ODPD iterations on a less complex NN. It can be seen in Fig. 5 that the more complex NN with $I = 1$ gave a better result than the less complex NN with $I = 2$. Hence, improving the neural network (or, more generally, the forward PA model) is more worthwhile than performing additional ODPD iterations.

VI. Conclusions

In this work, we introduced an OFDM-based DPD (ODPD) method that takes advantage of the guard-band subcarriers to predistort in the frequency domain. Our proposed method of predistortion does not require the estimation of an inverse PA model and was able to linearize our test PA as effectively as state-of-the-art methods. Using an NN-based forward model, we showed that this performance could be achieved with 34.8% fewer multiplications and a lower oversampling rate for the DPD application.

References

[1] L. Ding *et al.*, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan 2004.

[2] 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, 2006.

[3] Changsoo Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. on Signal Process.*, vol. 45, no. 1, pp. 223–227, 1997.

[4] F. Mkadem and S. Boumaiza, "Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 4, pp. 913–923, 2011.

[5] C. Tarver, L. Jiang, A. Sefidi, and J. R. Cavallaro, "Neural network DPD via backpropagation through a neural network model of the PA," in *53rd Asilomar Conf. on Signals, Syst., and Comput.*, 2019, pp. 358–362.

[6] W. Chen, G. Lv, X. Liu, D. Wang, and F. M. Ghannouchi, "Doherty PAs for 5G massive MIMO: Energy-efficient integrated DPA MMICs for sub-6-GHz and mm-wave 5G massive MIMO systems," *IEEE Microw. Mag.*, vol. 21, no. 5, pp. 78–93, 2020.

[7] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, 2009.

[8] C. Hsu and H. Liao, "PAPR reduction using the combination of precoding and mu-law companding techniques for OFDM systems," in *IEEE 11th Int. Conf. on Signal Process.*, vol. 1, 2012, pp. 1–4.

[9] S. Brandes, I. Cosovic, and M. Schnell, "Reduction of out-of-band radiation in OFDM systems by insertion of cancellation carriers," *IEEE Commun. Lett.*, vol. 10, no. 6, pp. 420–422, 2006.

[10] J. Chani-Cahuana, P. N. Landin, C. Fager, and T. Eriksson, "Iterative learning control for RF power amplifier linearization," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 9, pp. 2778–2789, 2016.

[11] D. R. Morgan, Z. Ma, and L. Ding, "Reducing measurement noise effects in digital predistortion of RF power amplifiers," in *IEEE Int. Conf. on Commun.*, vol. 4, 2003, pp. 2436–2439 vol.4.

[12] Mathworks, "OFDM modulation." [Online]. Available: https://www.mathworks.com/help/lte/ref/lteofdmmodulate.html

[13] *Power Amplifier Module for LTE and 5G*, NXP Semiconductors, 5 2019, rev. 3. [Online]. Available: https://www.nxp.com/docs/en/data-sheet/AFSC5G35D37.pdf