Reduced-Complexity Digital Predistortion: Adaptive Spline-Based Hammerstein Approach

Pablo Pascual Campo, Alberto Brihuega, Student Member, IEEE, Lauri Anttila, Member, IEEE, Matias Turunen, Dani Korpi, Member, IEEE, Markus Allén, and Mikko Valkama, Senior Member, IEEE

Abstract—In this paper, a novel digital predistorter concept for power amplifier (PA) linearization is proposed, with particular emphasis on reduced processing complexity in future 5G and beyond wideband radio systems. The proposed method builds on a complex spline interpolated look-up table (LUT) followed by a linear finite impulse response (FIR) filter, comprising essentially a Hammerstein-type nonlinear system. For reliable parameter learning, gradient-descent based adaptive learning rules are derived, allowing for the estimation of the spline control points and the FIR filter parameters in a decoupled manner. Large set of experimental results are provided, with specific focus on 5G New Radio (NR) systems, showing successful linearization of multiple sub-6 GHz PA samples as well as a 28 GHz active antenna array, incorporating channel bandwidths up to 200 MHz. Explicit performance-complexity comparisons are also reported against two known reference solutions, namely a memory polynomial (MP) based DPD and a linear interpolated LUT. The results show that the linearization performance of the proposed method is very close to that of a memory polynomial while clearly outperforming the linear interpolated LUT. Additionally, it is shown that the processing complexity of the proposed DPD is commonly some 60 % lower than that of the MP based DPD, offering thus a very good complexity-performance tradeoff.

Index Terms—Digital predistortion, power amplifier, spline interpolation, Hammerstein model, look-up table, nonlinear distortion, behavioral modeling, EVM, ACLR

I. INTRODUCTION

MODERN radio communication systems, such as the 4G LTE/LTE-Advanced and the emerging 5G New Radio (NR) mobile networks, build on multicarrier modulation, most notably orthogonal frequency division multiplexing (OFDM) [1]. OFDM waveforms are known to contain high peak-to-average power ratio (PAPR) [2], [3], which complicates utilizing highly nonlinear power amplifiers (PAs) in transmitters operating close to saturation [2], [4], [5]. Digital predistortion (DPD) is, generally, a well-established approach to control the unwanted emissions and nonlinear distortion stemming from nonlinear PAs, see, e.g., [2], [4], [6]–[9] and references therein. Especially when combined with appropriate PAPR reduction methods [10], DPD based systems can largely improve the transmitter power efficiency while keeping the unwanted emissions within specified limits.

Some of the most common approaches in PA direct modeling as well as DPD processing are the memory polynomial (MP) [2], [9], [11] and the generalized memory polynomial (GMP) [2], [11]–[13], both of which can be interpreted to be special cases of the Volterra series [2], [14]–[16]. Such approaches allow for efficient direct and inverse modeling of nonlinear systems with memory, while also support straightforward parameter estimation, through, e.g., linear least-squares (LS), as they are known to be linear-in-parameters models [11]. However, the processing complexity per linearized sample is also relatively high, particularly with GMP and other more complete Volterra series type of approaches, though also some works exist where complexity reduction is pursued [15], [17].

In this paper, we develop and describe a DPD solution whose linearization capabilities are similar to those of the well-established memory polynomials, while at the same time offering a substantially reduced DPD main path processing and parameter learning complexities. Such reduced-complexity DPD solution is mainly motivated by the following four facts or tendencies. First, the channel bandwidths in NR are substantially larger than those in LTE-based systems. Specifically, up to 100 MHz and 400 MHz continuous channel bandwidths are already specified in NR Release-15 at frequency range (FR) 1 (below 6 GHz bands) and FR-2 (24-40 GHz bands), respectively, [18], which imply increased DPD processing rates. Second, the actual unwanted emission requirements, particularly in the form of adjacent channel leakage ratio (ACLR), are largely relaxed in NR FR-2 systems, being only in the order of 26–28 dB [18], increasing the feasibility of simplified DPD solutions. Third, the medium range and local area base-stations adopt substantially reduced transmit powers [18], compared to classical macro base-stations, hence the available power budget of the DPD solutions is also reduced. Finally, as observed recently in [15], continuous learning may be needed in FR-2 and other mmWave active array systems, hence developing methods which reduce the parameter learning complexity becomes important.

The DPD method proposed in this paper, referred to as the complex spline-based Hammerstein (SPH) approach, builds on complex spline-interpolated lookup table (LUT) followed by a linear finite impulse response (FIR) filter, and is illustrated conceptually in Fig. 1. The interpolation allows to use a small amount of points in the LUT, while the linear filter facilitates basic memory modeling. Gradient-based decoupled
learning algorithms are also derived, to efficiently estimate the LUT control points as well as the linear filter parameters. Additionally, a comprehensive computational complexity analysis is provided, and compared to MP DPD as well as to linear-interpolated LUT which can be seen as a special case of the proposed method. Then, an extensive set of RF measurement results are provided, covering several different FR-1 PA samples, channel bandwidth cases as well as base-station classes. Additionally, a state-of-the-art 28 GHz active antenna array, specifically Anokiwave AWMF-0129, is successfully linearized with 100 MHz and 200 MHz 5G NR channel bandwidths.

For clarity, it is acknowledged that LUT-based PA linearization, as such, is a well-known approach, see, e.g., [19–21] and the references therein. However, the PA memory aspects are not considered in [19], while a fairly sizeable LUT without interpolation is considered in [21]. Additionally, a LUT-type implementation of a memory polynomial is described in [20] while the learning is based on classical LS model fitting.

The rest of the paper is organized as follows. Section II describes the proposed complex spline-interpolated Hammerstein DPD system together with the parameter learning algorithms. Section III presents a complexity analysis and comparison of the proposed DPD solution and the two reference schemes, namely linear-interpolated LUT and MP DPD. Section IV describes the RF measurement setups, and presents the corresponding measurement results and their analysis. Finally, conclusions and drawn in Section V while the detailed derivations of the gradient-based parameter estimation algorithms are provided in Appendix A.

Throughout the rest of this article, matrices are denoted by boldface letters, e.g., $\mathbf{A} \in \mathbb{C}^{(M \times N)}$, while vectors are denoted by lowercase boldface letters, e.g., $\mathbf{v} \in \mathbb{C}^{M \times 1} = [v_1 \ v_2 \ \cdots \ v_M]^T$. Ordinary transpose and hermitian operators are represented as $(\cdot)^T$ and $(\cdot)^H$, respectively. Additionally, the absolute value, floor, and ceil operators are represented as $|\cdot|$, $\lfloor \cdot \rfloor$, and $\lceil \cdot \rceil$, respectively.

II. PROPOSED COMPLEX SPLINE-BASED HAMMERSTEIN DPD: PROCESSING AND PARAMETER LEARNING

In general, for notational convenience, the mathematical presentation in the following subsections is formulated in the context of indirect learning architecture (ILA) for the postdistortion processing, with $x_{SP}[n]$ denoting the postdistorter input, while in the actual predistortion stage the input signal is $x[n]$, as illustrated also in Fig. 1.

A. Spline Modeling Basics

Building on selected piece-wise polynomials, spline based modeling and interpolation seeks to determine a smooth curve that conforms or approximates a set of points, commonly known as control points. Consequently, the input signal range is divided into several pieces, and the polynomials model the nonlinear system behavior in the corresponding regions under continuity and smoothness constraints. With such approach, simple low-order functions can be adopted, per region, in contrast to more ordinary methods where a single high-order function or polynomial seeks to model the whole input range.

The continuous piece-wise polynomials are determined or constructed by the set of $Q$ control points that together with the polynomials describe the nonlinearity present in the system. In general, such construction can be expressed for arbitrary number of control points, $Q$, and polynomial degree, $P_{SP}$, where each region is basically the result of the interpolation of $P_{SP} + 1$ spline curves. For a comprehensive review of spline-based modeling for real-valued signals, please refer to [22].

As opposed to the traditional approach where spline modeling has been applied to real-valued signals and systems [23], [25], [26], in the context of radio communications complex I/Q signals are utilized. Therefore, the splines need to be extended to complex domain. Denoting the input signal by $x_{SP}[n]$, as in Fig. 1(b), two separate splines are adopted, one for the I and another one for the Q component. Additionally, the instantaneous magnitude $|x_{SP}[n]|$ is utilized as the input for both I and Q splines, and the regions are built by defining the span index $i_n$ and abscissa value $u_n$ at time instant $n$, expressed as

$$i_n = \left\lfloor \frac{|x_{SP}[n]|}{\Delta_x} \right\rfloor + 1,$$

$$u_n = \frac{|x_{SP}[n]|}{\Delta_x} - (i_n - 1),$$

Fig. 1. Illustration of the considered DPD system building on indirect learning architecture (ILA), in (a), and the proposed spline-interpolated adaptive Hammerstein model, in (b), with gradient-based adaptive parameter estimation.
where $\Delta_x > 0$ is the width or separation between the regions, and where we have assumed uniform equi-spaced splines for simplicity. Here, $i_n$ denotes the index of the selected region in the time instant $n$, and $u_n$, $0 \leq u_n < \Delta_x$, represents the normalized value of the corresponding input envelope within the current region $i_n$. A conceptual illustration in the context of mapping between the complex input $x_{SP}[n]$ and the output $s_{SP}[n]$, defined in \([3]\), as a function of the input envelope $|x_{SP}[n]|$ is shown in Fig. 2. It is noted that the range of the index variable $i_n$, and thus the number of regions, will depend on the range of the input envelope $|x_{SP}[n]|$ and $\Delta_x$, while $u_n$ will be always enclosed between 0 and $\Delta_x$.

**B. DPD Processing**

Following the processing architecture in Fig. 1(b), we first derive an expression for the complex I/Q spline-interpolated memoryless nonlinearity output, denoted by $s_{SP}[n]$. Specifically, following \([23]\) but extending to complex I/Q signals and complex splines, the signal $s_{SP}[n]$ can be expressed as

$$s_{SP}[n] = x_{SP}[n]g_n^T(1 + c),$$

where $I$ denotes a $Q \times 1$ column vector of all ones, while $c \in \mathbb{C}^{Q \times 1} = [c_0 \ c_1 \ \cdots \ c_{Q-1}]^T$ is the complex-valued set containing the control points for the I and Q components, representing essentially an interpolated lookup table (LUT) of size $Q$, with the control points being the table entries. We note that the number of control points $Q$ is given by $Q = i_{\text{max}} + P_{SP}$, where $i_{\text{max}}$ is the number of regions considered in the model. The vector $g_n \in \mathbb{R}^{Q \times 1}$, in turn, is defined as

$$g_n = [0 \ \cdots \ 0 \ u_n^T \ B_{P_{SP}} \ 0 \ \cdots \ 0]^T,$$

where

$$u_n = [u_{n}^{P_{SP}} \ u_{n}^{P_{SP}-1} \ \cdots \ 1]^T \in \mathbb{R}^{(P_{SP}+1) \times 1},$$

and $B_{P_{SP}}$ is the spline basis matrix of order $P_{SP}$. In the vector $g_n^T$, the term $u_n^T B_{P_{SP}}$ of size $1 \times (P_{SP} + 1)$ is located such that the starting index is $i_n$. Thus, at a given time instant $n$, only the control points $c_{i_n}, c_{i_n+1}, \ldots, c_{i_n+P_{SP}}$ contribute to the output. Hence, intuitively, the nonlinear mapping between the input $x_{SP}[n]$ and the output $s_{SP}[n]$ is approximated by first linearly combining different monomial transformations $u_n^{P_{SP}}, u_n^{P_{SP}-1}, \ldots, 1$, through spline basis matrix $B_{P_{SP}}$, which are then further combined together via weighting the control points.

Additionally, it is noted that the so-called basis matrix $B_{P_{SP}}$ can be precomputed for given type of splines and polynomial order, and it is therefore static. As a concrete example, in this article we especially focus on 3rd order ($P_{SP} = 3$, cubic interpolation) B-splines, although other spline orders are tested and demonstrated as well. In this case, the basis function matrix, $B_3$, can be expressed as \([23]\)

$$B_3 = \frac{1}{6} \begin{bmatrix} \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 \\ 1 & 4 & 1 \end{bmatrix}. \quad (6)$$

Next, after having derived the expression for the memoryless nonlinear signal model, the memory effects are incorporated through the FIR filter stage. Hence, the overall output signal $y_{SP}[n]$ can be directly expressed as

$$y_{SP}[n] = h^H s_n,$$

where $h \in \mathbb{C}^{M_{SP} \times 1} = [h_0 \ h_1 \ \cdots \ h_{M_{SP}-1}]^T$ contains the filter coefficients, with $M_{SP}$ denoting the number of memory taps included in the model, while $s_n \in \mathbb{C}^{M_{SP} \times 1} = [s_{SP}[n] \ s_{SP}[n-1] \ \cdots \ s_{SP}[n-M_{SP}+1]]^T$.

**C. Parameter Learning Rules**

We next derive efficient gradient-descent type learning rules, to adaptively estimate and track the unknown parameters, namely the vectors $c$ and $h$ containing the spline control points and the FIR memory filter coefficients. Notation-wise, to allow for sample-adaptive estimation, we denote the vectors by $c_n$ and $h_n$, from now on, to indicate their time-dependence.

To this end, the instantaneous error signal between $x_{\text{DPD}}[n]$ and $y_{SP}[n]$, in the context of the considered ILA-type architecture, is first defined as

$$e[n] = x_{\text{DPD}}[n] - y_{SP}[n] = x_{\text{DPD}}[n] - h_n^H s_n.$$ \quad (8)

Then, to facilitate the gradient-descent learning \([27]\), the cost function is defined as the instantaneous squared error, expressed as

$$J(h_n, c_n) = |e[n]|^2.$$ \quad (9)

The corresponding iterative learning expressions are then obtained through the partial derivatives of $J(h_n, c_n)$ with respect both parameters, expressed formally as

$$h_{n+1} = h_n - \mu_h[n] \frac{\partial J(h_n, c_n)}{\partial h_n}, \quad (10)$$

$$c_{n+1} = c_n - \mu_c[n] \frac{\partial J(h_n, c_n)}{\partial c_n}, \quad (11)$$

where $\mu_h[n]$ and $\mu_c[n]$ are the learning rates for $h_n$ and $c_n$, respectively, at time instant $n$. After straight-forward
derivations, provided in Appendix A, the resulting concrete learning rules read

\[
h_{n+1} = h_n + \mu_h[n]e^*[n]s_n, \quad (12)
\]

\[
c_{n+1} = c_n + \mu_c[n]e[n]\Sigma_n^T X_n^* h_n, \quad (13)
\]

where the \(M_{SP} \times M_{SP}\) diagonal matrix \(X_n\) reads \(X_n = \text{diag}\{x_{SP}[n], x_{SP}[n-1], \ldots, x_{SP}[n-M_{SP} + 1]\}\), and \(\Sigma_n\) contains \(M_{SP}\) previous instants of \(g_n\), defined as \(\Sigma_n \in \mathbb{R}^{M_{SP} \times Q} = [g_{n-1} \quad g_{n-2} \ldots \quad g_{n-M_{SP}+1}]^T\). These learning rules in (12) and (13) are executed in parallel such that both parameter vectors are updated simultaneously. For readers’ convenience, an example illustration of the structure of the matrix \(\Sigma_n\) is given in (14), for \(M_{SP} = 4, Q = 9, P_{SP} = 3,\) assuming representative example values of the index variable \(i_n\).

\[
\Sigma_n = \begin{bmatrix}
0 & 0 & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & * & * & * & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & * & * & * & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & * & * & * \\
\end{bmatrix}
\]

Note that the term \(u_n^T B_{PSP}\) is located in \(\Sigma_n\) at each iteration \(n\) according to the span index \(i_n\), as shown in (3).

III. COMPUTATIONAL COMPLEXITY ANALYSIS

In this section, a computational complexity analysis and comparison between the proposed SPH DPD, a linear interpolated LUT, and a widely-applied MP DPD with LMS Newton-like parameter adaptation is presented. LMS type adaptation is deliberately assumed also for MP DPD, for the fairness of the comparison. The complexity analysis is carried out in terms of floating point operations (FLOPs) per sample, essentially measuring the number of additions, subtractions and multiplications required to process one single linearized sample. Additionally, the required number of real multiplications is also shown, as multiplications are commonly more resource-intensive operations than additions in digital signal processing (DSP) implementations [11].

In general, the DPD main processing/transmit path is the most critical stage, complexity wise, as it needs to be executed continuously and in real time to predistort the transmit signal. However, especially if frequent or even continuous DPD coefficient learning/tracking is also required, the parameter learning complexity is also of large importance. Hence, in our complexity analysis, both the DPD main path and the DPD learning stage are addressed.

A. Complexity of Proposed SPH Method

The complexity analysis of the proposed gradient-adaptive SPH DPD is carried out by following the processing elements described in Section [II]. It is noted that the complexity expressions reported below basically represent an upper bound for the required arithmetical operations, as in real implementations some elementary or trivial operations such as multiplying by any integer power of 2 or 1/2 does not really reflect any actual FLOPs while are included as normal operations in the expressions for simplicity.

To this end, with reference to Fig. [1] and Section II, the generic complexity expressions can be stated as follows

- **DPD main path**, starting with the input signal \(x[n]\):
  1) \(s[n] \rightarrow 2P_{SP}^2 + 6P_{SP} + 18\) FLOPs.
  2) \(x_{DPD}[n] \rightarrow 8M_{SP} - 2\) FLOPs.

- **DPD learning**, for observed signal \(x_{SP}[n]\):
  1) \(y_{SP}[n] \rightarrow 2P_{SP}^2 + 6P_{SP} + 8M_{SP} + 16\) FLOPs.
  2) \(e[n] \rightarrow 2\) FLOPs.
  3) \(h_{n+1} \rightarrow 8M_{SP} + 2\) FLOPs.
  4) \(c_{n+1} \rightarrow 4P_{SP}M_{SP} + 10M_{SP} - 2P_{SP} + 9Q\) FLOPs.

It is noted that the amount of FLOPs in the DPD main path does not depend on the chosen number of control points \(Q\), or equivalently the number of regions, as the spline-interpolation algorithm basically utilizes \(P_{SP} + 1\) control points for any given region.

B. Complexity of Reference Methods

1) **Linear interpolated LUT**: The linear interpolated LUT is basically a special case of the SPH model, with \(P_{SP} = 1\) and \(M_{SP} = 0\), and thus its computational complexity can be directly obtained through the expressions shown in the previous subsection.

2) **Memory polynomial DPD**: When considering the LMS-adaptive MP DPD with monomial basis functions (BFs), in the context of ILA architecture in Fig. [1](a), we first write the postdistorter output sample as

\[
y_{DPD}[n] = w_{n}^H I_n, \quad (15)
\]

where \(w_n \in \mathbb{C}^{n \times 1}\) is the MP DPD coefficient vector, with \(m = [P_{MP}^2]/(1 + M_{MP})\) denoting the number of coefficients, while \(P_{MP}\) and \(M_{MP}\) are the assumed polynomial order and memory length (per nonlinearity order), respectively. Additionally, the vector of the basis function samples \(I_n\) used to calculate the current output is as defined in (16), where \(x_{MP}[n]\) denotes the observed feedback signal at postdistorter input.

Once \(y_{MP}[n]\) is calculated, the error signal can be directly obtained as

\[
e_{MP}[n] = x_{DPD}[n] - y_{MP}[n] = x_{DPD}[n] - w_{n}^H I_n, \quad (17)
\]

and the coefficient update can be written as

\[
w_{n+1} = w_n + \mu_e[n]e_{MP}[n]R^{-1}I_n, \quad (18)
\]

where \(\mu_e[n]\) is the learning rate, and \(R^{-1}\) is the inverse of the autocorrelation matrix of the PA output basis function samples [27]. We assume that a block of \(N_B\) samples is used to calculate the sample estimate of \(R\), and include below
TABLE I

Complexity expressions in terms of (i) FLOPs per sample and (ii) real multiplications per sample for the proposed SPH method and the reference MP method, covering both the DPD main path processing and the DPD parameter learning

| Operation     | FLOPs / sample | Real multiplications / sample |
|---------------|----------------|-----------------------------|
| **Main path** |                |                             |
| Nonlinearity  | $2P_{SP}^2 + 6P_{SP} + 18$ | $3\lceil P_{MP} \rceil - 1$ |
| Filtering     | $8M_{SP} - 2$  | $8m - 2$                    |
| **Total**     | $2P_{SP}^2 + 6P_{SP} + 8M_{SP} + 16$ | $3\lceil P_{MP} \rceil + 8m - 3$ |
| **Learning**  |                |                             |
| Error signal  | $2P_{SP}^2 + 6P_{SP} + 8M_{SP} + 18$ | $3\lceil P_{MP} \rceil + 8m - 1$ |
| Coeff. update | $2P_{SP}(2M_{SP} - 1) + 18M_{SP} + 9Q + 2$ | $8m^2 + 14m + 2$ |
| **Total**     | $P_{SP}(2P_{SP} + 4M_{SP} + 4)$ | $3\lceil P_{MP} \rceil + 8m^2 + 22m + 1$ |

$I_n = \begin{bmatrix} x_{MP}[n] & x_{MP}[n] |x_{MP}[n]|^2 & \ldots & x_{MP}[n] |x_{MP}[n]|^{P_{MP} - 1} & x_{MP}[n - 1] & x_{MP}[n - 1] |x_{MP}[n - 1]|^2 & \ldots & x_{MP}[n - 1] |x_{MP}[n - 1]|^{P_{MP} - 1} \\ x_{MP}[n - M_{MP} + 1] & x_{MP}[n - M_{MP} + 1] |x_{MP}[n - M_{MP} + 1]|^2 & \ldots & x_{MP}[n - M_{MP} + 1] |x_{MP}[n - M_{MP} + 1]|^{P_{MP} - 1} \end{bmatrix}^T$. 

TABLE II

Numerical complexity values, in terms of FLOPs per sample and real multiplications per sample, for $Q = 15$, $\Delta_x = 1$, $M_{SP} = 4$ (SPH), and $P_{MP} = 11$, $M_{MP} = 3$ (MP)

|                | FLOPs / sample | Real mul. / sample |
|----------------|----------------|--------------------|
|                 | SPH DPD | MP DPD | SPH DPD | MP DPD |
| **Nonlinearity** | 47     | 17     | 22      | 16     |
| **Filtering**   | 30     | 190    | 16      | 96     |
| **Total Main path** | 77     | 207    | 38      | 112    |
| **Reduction**   | 62.8%  | 66.1%  |         |        |
| **Error signal**| 79     | 209    | 38      | 112    |
| **Coeff. update**| 251    | 4754   | 128     | 2402   |
| **Total Learning** | 330    | 4963   | 166     | 2514   |
| **Reduction**   | 93.4%  | 93.3%  |         |        |

Next, to obtain concrete complexity numbers and to carry out a comparison, we study an example case where the SPH DPD spline polynomial order is $P_{SP} = 3$. Additionally, the number of control points is chosen to be $Q = 15$, the width of the regions is assumed to be $\Delta_x = 1$ (uniform splines), and the considered memory length is $M_{SP} = 4$. These constitute a total number of 22 free parameters to be estimated, in the SPH DPD case. Then, the MP DPD polynomial order is chosen as $P_{MP} = 11$, and the considered memory length per filter is $M_{MP} = 3$. This configuration leads to 24 free parameters in the MP DPD, constituting thus a fair starting point of the complexity comparison with roughly the same number of free parameters. Similar type parametrizations are used also in the actual DPD measurements and experiments, in Section IV.

The resulting exact numerical processing complexities, expressed in terms of FLOPs per sample and real multiplications per sample, are presented in Table II. In these numerical values, when it comes to the SPH DPD, we have excluded the trivial operations, i.e., multiplications by zeros, ones and
Host processing
- Generate I/Q baseband data samples.
- Transmit I/Q samples through RF out port.
- Receive I/Q samples through RF in port.
- Learn new DPD coefficients from the transmit and observed data. Repeat until the algorithm is converged.
- Deploy and quantify DPD performance.

(a) RF measurement setup at FR-1.

(b) General purpose wideband PA.

(c) Skyworks NR Band 78 PA.

(d) Proprietary LTE-A Band 3 LDMOS PA.

Fig. 3. Overall RF measurement setup and PA modules used in the Experiments 1-3.

integer powers of two or half, stemming from the structure of $B_3$ in (6). Overall, the results in Table II demonstrate the large complexity reduction provided by the proposed DPD approach, as nearly 63% less FLOPs per sample, and 66% less real multiplications per sample are needed in the DPD main path to predistort the input signal. Furthermore, the required parameter learning complexity is also very remarkably reduced, by 93% in terms of FLOPs per sample and real multiplications per sample, indicating that solutions like this might already facilitate even continuous learning in selected applications. Additionally, owing to the largely reduced learning complexity, the feasibility of implementing both the DPD parameter learning as well as the main path processing in the same chip increases.

IV. EXPERIMENTAL RESULTS

In order to evaluate and validate the proposed DPD concept, four separate linearization experiments are carried out. Three of the measurement scenarios are related to FR-1 (sub-6 GHz) power amplifiers and classical conducted measurements, including a general purpose wideband PA, a 5G NR Band 78 small-cell BS PA, and an LTE-Advanced Band 3 high-power macro BS PA. The fourth experiment is then related to FR-2 and over-the-air (OTA) measurements where a state-of-the-art 28 GHz active antenna array with 64 integrated PAs and antenna units is linearized.

A. FR-1 Measurement Environment and Figures of Merit

The FR-1 measurement setup utilized for the first three experiments is illustrated in Fig. 3 (a), and consists of a National Instruments PXIe-5840 vector signal transceiver (VST), facilitating arbitrary waveform generation and analysis between 0-6 GHz with instantaneous bandwidth of 1 GHz. This instrument is used as both the transmitter and the observation receiver, and includes also an additional host-processing based computing environment where all the digital waveform and DPD processing can be executed. In a typical conducted measurement, the baseband complex I/Q waveform is generated by MATLAB in the VST host environment, and fed to the device under test (DUT) through the VST transmit chain, incorporating also an external feeding or driver amplifier in high power measurements. The DUT output is then observed via the VST receiver, through an external attenuator. All DPD parameter learning and actual DPD main math processing stages are executed in the host environment. Finally, the actual DPD performance measurements are carried out where different random modulating data is used, compared to the learning phase.

In order to measure and quantify the performance of the DPD methods, selected metrics or figures of merit are needed. In this work, we adopt the well-established error vector magnitude (EVM) and ACLR metrics, as defined for 5G NR in [18]. The EVM focuses on the passband transmit signal...
quality, and can be defined as

\[
EVM (\%) = \sqrt{\frac{P_{\text{error}, \text{eq.}}}{P_{\text{ref.}}}} \times 100,
\]

(19)

where \( P_{\text{error}, \text{eq.}} \) denotes the power of the error signal calculated between the ideal subcarrier symbols and the corresponding observed subcarrier samples at the PA output after zero forcing equalization, removing the effects of the possible linear distortion \[18\]. Furthermore, \( P_{\text{ref.}} \) denotes the corresponding power of the ideal (reference) symbols. The ACLR, in turn, is defined as the ratio of the transmitted power within the desired channel \((P_{\text{desired ch.}})\) and that in the left or right adjacent channel \((P_{\text{adj. ch.}})\), expressed as

\[
\text{ACLR (dB)} = 10 \log_{10} \frac{P_{\text{desired ch.}}}{P_{\text{adj. ch.}}},
\]

(20)

measuring thus the out-of-band performance. While ACLR is, by definition, a relative measure, an explicit out-of-band spectral density limit, in terms of dBm/MHz measured with a sliding 1 MHz window in the adjacent channel region, is also defined for certain base-station types \[18\], referred to as the absolute basic limit in 3GPP terminology. Thus, the PA output spectral density in dBm/MHz is also quantified in the measurements, particularly in the context of local area and medium-range BS PAs \[18\].

All the forthcoming experiments utilize 5G NR Release-15 standard compliant OFDM downlink waveform and channel bandwidths \[18\], while the adopted carrier frequencies in each experiment are selected according to the available 5G NR bands and the available PA samples. In all experiments, the PAPR of the digital waveform prior to the DPD stage is limited to 7.0 dB, through well-known iterative clipping and filtering based processing, while also additional time-domain windowing is applied to suppress the inherent OFDM signal sidelobes. More specific waveform parameters such as the subcarrier spacing (SCS) and the occupied physical resource block (PRB) count are stated along the experiments.

B. Experiment 1: General Purpose PA

The first experiment focuses on a general purpose wideband PA (Mini-Circuits ZHL-4240), illustrated in Fig. 3(b), as the actual amplification stage. The amplifier has a gain of 41 dB, and a 1-dB compression point of +31 dBm, being basically applicable in small-cell and medium-range base-stations. The transmit signal is a 5G NR downlink OFDM waveform, with 30 kHz subcarrier spacing and 264 active PRBs \[18\], yielding an aggressive passband width of 95.04 MHz. The RF center frequency is 3.5 GHz and the assumed channel bandwidth is 100 MHz. The I/Q samples are transmitted through the VST RF output port directly to the PA, facilitating a maximum output power of +27 dBm. The proposed and the reference DPD schemes are then adopted, and the performance quantification measurements are carried out. In all results, five ILA learning iterations are adopted while the signal length within each ILA iteration is 100,000 samples. In this experiment, the VST observation receiver runs at 491.52 MHz (4x oversampling).

Fig. 4 shows a snap-shot linearization example, at PA output power of +27 dBm, when different spline orders \((P_{\text{SP}})\) of 2, 3 and 4 are applied in the proposed SPH DPD, while the number of control points is fixed to \(Q = 15\) and the memory filter order \(M_{\text{SP}} = 3\). Additionally, an LMS-based MP DPD of order nine \((P_{\text{MSP}} = 9)\) with memory filters of order \(M_{\text{MSP}} = 3\) as well as a linear interpolated LUT DPD are also adopted and presented for reference. We can observe that the performance of the proposed DPD is very close to that of the MP DPD, despite the substantially reduced complexity. The figure also illustrates that all DPD methods basically satisfy the absolute basic limit requirement of -25 dBm/MHz, which if less stringent than the
### Table III

|                | DPD running complexity | DPD performance |
|----------------|------------------------|-----------------|
|                | P | M | Q | $\Delta_v$ | # of coefficients | FLOPs/sample | Mul./sample | EVM (%) | Max. dBm/MHz |
| No DPD         |   |   |   |           |                  |             |            |         |             |
| SPH DPD        | 1 | - | 15| 1          | 16               | 30          | 15         | 6.54    | -29.20      |
|                | 2 | 3 | 15| 1          | 20               | 63          | 43         | 5.61    | -32.30      |
|                | 3 | 3 | 15| 1          | 21               | 68          | 46         | 5.54    | -36.30      |
|                | 4 | 3 | 15| 1          | 22               | 73          | 49         | 5.55    | -36.80      |
| MP DPD         |   | 9 | 3 | -          | 20               | 172         | 93         | 5.47    | -38.20      |

---

Fig. 6. Example illustration of linearization results in Experiment 2 (NR small-cell PA measured at 3.65 GHz), with 100 MHz channel bandwidth and PA output power of +24 dBm. DPD parameters are as shown in Table IV.

The classical 45 dB ACLR limit, applies in medium-range BS cases with transmit powers of higher than +24 dBm up to +38 dBm [18].

Fig. 5 then presents the behavior of the measured EVM and ACLR performance metrics, as functions of the PA output power, with the SPH DPD spline order now fixed to $P_{SP} = 3$, while otherwise following the same DPD parameterization. Again, we can observe that the proposed SPH DPD and the reference MP DPD behave very similarly. Table III then collects and summarizes the obtained DPD results in Experiment 1, at PA output power of +27 dBm, while also showing the DPD main path processing complexities. We can conclude that the proposed DPD offers a favorable performance-complexity trade-off compared to the reference MP DPD approach.

C. Experiment 2: 5G NR Band 78 Small-Cell PA

The second experiment includes the Skyworks SKY66293-21 PA module, illustrated in Fig. 3 (c), which is a low-to-medium power PA oriented to be used either in small-cell base-stations or in large antenna array transmitters. The PA module is specifically designed to operate in the NR Band 78 (3300-3800 MHz), having a gain of 34 dB, and a 1-dB compression point of +31.5 dBm. Similar 5G NR downlink signal corresponding to the 100 MHz channel bandwidth scenario, as in the Experiment 1, is adopted, while the considered RF center-frequency is 3.65 GHz. The test signal is again transmitted via the RF TX port of the VST directly to the PA module, while the considered PA output power is +24 dBm, corresponding to

---

Fig. 7. Measured AM-AM and AM-PM responses of SKY66293-21 PA at 3.65 GHz when excited with an NR 100 MHz transmit signal. DPD parameters are as shown in Table IV.
TABLE IV

**SUMMARY OF DPD MAIN PATH PROCESSING COMPLEXITY AND LINEARIZATION PERFORMANCE RESULTS OBTAINED IN EXPERIMENT 2, PA OUTPUT POWER IS +24 dBm**

| DPD running complexity | DPD performance |
|------------------------|-----------------|
| P M Q \( \Delta x \) | # of coefficients | FLOPs/sample | Mul./sample | EVM (%) | Max. dBm/MHz |
| No DPD | - - - - | - | 0 | 0 | 8.64 | -18.20 |
| SPH DPD | 1 - 15 1 | 16 | 30 | 15 | 6.02 | -26.90 |
| | 2 4 15 1 | 21 | 71 | 35 | 5.80 | -29.40 |
| | 3 4 15 1 | 22 | 77 | 38 | 5.57 | -32.60 |
| | 4 4 15 1 | 23 | 81 | 41 | 5.58 | -32.60 |
| MP DPD | 11 4 - - | 30 | 255 | 136 | 5.54 | -33.20 |

TABLE V

**SUMMARY OF DPD MAIN PATH PROCESSING COMPLEXITY AND LINEARIZATION PERFORMANCE RESULTS OBTAINED IN EXPERIMENT 3, PA OUTPUT POWER IS +48 dBm**

| DPD running complexity | DPD performance |
|------------------------|-----------------|
| P M Q \( \Delta x \) | # of coefficients | FLOPs/sample | Mul./sample | EVM (%) | ACLR (L/R) (dB) |
| No DPD | - - - - | - | 0 | 0 | 8.62 | 35.40 / 35.80 |
| SPH DPD | 1 - 15 1 | 16 | 30 | 15 | 7.54 | 40.20 / 41.80 |
| | 2 6 15 1 | 23 | 87 | 43 | 6.02 | 45.10 / 45.20 |
| | 3 6 15 1 | 24 | 92 | 46 | 5.60 | 45.90 / 46.10 |
| | 4 6 15 1 | 25 | 97 | 49 | 5.57 | 45.20 / 47.30 |
| MP DPD | 11 4 - - | 30 | 255 | 136 | 5.46 | 51.40 / 49.90 |

the maximum transmit power of a Local Area BS according to the NR regulations [13]. Again, five ILA learning iterations are adopted while the signal length within each ILA iteration is 100,000 samples. In this experiment, the VST observation receiver runs at 491.52 MHz (4x oversampling).

Fig. 6 and Table IV illustrate and summarize the obtained linearization results for the proposed and the reference DPD methods. Again, also comparative complexity numbers are stated in Table IV. As stated in [13], a 5G NR Local Area BS can operate within an absolute basic limit of -32 dBm/MHz in the adjacent channel region, assuming the considered PA output power of +24 dBm. As shown in Fig. 6 and Table IV, both the 3rd order SPH DPD and the memory polynomial DPD satisfy this limit, indicating successful linearization. For further visualization, the AM-AM and AM-PM responses are also illustrated in Fig. 7 for selected DPD cases. Again, as can be observed in Table IV, a remarkable complexity reduction is obtained through the proposed SPH DPD, compared to the reference MP DPD, while both providing a very similar linearization performance.

D. Experiment 3: LTE-Advanced Band 3 High-Power PA

The third experiment includes a wide-area/macro base-station PA module, illustrated in Fig. 3 (d), to demonstrate that the proposed DPD solution is also capable of linearizing high-power PAs. This LDMOS-based PA system is designed to operate at the LTE-A Band 3 (1805-1880 MHz), targeting Category A wide-area base-stations with output powers up to +48 dBm. A driver amplifier is now adopted before the DUT, providing a gain of 25 dB, while working in a relatively linear point to ensure that as little additional nonlinear distortion is injected into the transmit signal as possible. In this scenario, a 20 MHz channel bandwidth case is assumed and a 5G NR signal with SCS of 30 kHz and 50 active PRBs are assumed. The RF carrier frequency is 1.82 GHz.

Again, in addition to the proposed DPD, the MP DPD and linear-interpolated LUT methods are also tested, for reference purposes. Five ILA learning iterations are adopted while the signal length within each ILA iteration is 100,000 samples, and the VST observation receiver runs at 153.60 MHz (5x oversampling). Example measured linearization results are...
HMC943ALP5DE, with 17 dB and 23 dB gain, respectively, means of two Analog Devices’ driver PAs, HMC499LC4 and filters. The modulated RF waveform is then pre-amplified by running at 24.5 GHz, together with external mixers and other relevant instruments and equipment for signal generation. Anokiwave A WMF-0129 active antenna array together with measurement setup is depicted in Fig. 9, incorporating an MXG that generates the corresponding local oscillator signal directly the I/Q samples of a wideband modulated IF signal.

On the transmit chain side, the setup consists of a Keysight frequency with up to 3 GHz of instantaneous bandwidth. To the MP DPD approach. For fairness, it is additionally stated that the integrated PAs of the Anokiwave AWMF-0129 active antenna array are driven towards saturation. The transmit signal propagates over-the-air (OTA) and is captured by a horn antenna at the observation receiver, such that the receiving antenna system is well within the main transmit beam. At the receiver side, another Keysight N5183B-MXG and a mixing stage are used to downconvert the signal back to IF. Then, the Keysight DSOS804A oscilloscope is utilized as the actual digitizer, including also built-in filtering, and the signal is taken to baseband and processed in a host PC, where the DPD learning and predistortion are executed. The OTA measurement system is basically following the measurement procedures described in [18], [34], specifically the measurement option utilizing the beam-based directions. In these measurements, the observation receiver provides I/Q samples at 7x oversampled rate.

Fig. 9. RF measurement setup including the 64-element Anokiwave AWMF-0129 active antenna array, working at 28 GHz center frequency.

shown in Fig. 8 corresponding to the PA output power of +48 dBm, building on the parameter values and processing complexities indicated in Table V. It can be observed that the proposed SPH DPD can meet the macro BS 45 dB ACLR requirement, while in this case there is some performance gap to the MP DPD approach. For fairness, it is additionally stated that we measured the high-power PA also with 40 MHz and 60 MHz NR channel bandwidths, basically still fitting to the overall available bandwidth in LTE Band 3, and observed that the proposed SPH DPD cannot anymore provide the required ACLR of 45 dB, though lacking only by some 2-4 dB.

E. Experiment 4: FR-2 Environment and 28 GHz Active Array

In order to further demonstrate the applicability of the proposed spline-based DPD concept, the fourth and final experiment focuses on timely 5G NR mmWave/FR-2 deployments [18] with active antenna arrays. Unwanted emission modeling and DPD-based linearization of active arrays with large numbers of PA units is, generally, an active research field, with good examples of recent papers being, e.g., [5], [29]–[33]. Below we first describe shortly the FR-2 measurement setup, and then present the actual linearization results.

1) FR-2 Measurement Setup: The overall mmWave/FR-2 measurement setup is depicted in Fig. 9 incorporating an Anokiwave AWMF-0129 active antenna array together with other relevant instruments and equipment for signal generation and analysis, facilitating measurements at 28 GHz center-frequency with up to 3 GHz of instantaneous bandwidth. On the transmit chain side, the setup consists of a Keysight M8190 arbitrary waveform generator that is used to generate directly the I/Q samples of a wideband modulated IF signal centered at 3.5 GHz. The signal is then upconverted to the 28 GHz carrier frequency by utilizing the Keysight N5183B-MXG that generates the corresponding local oscillator signal running at 24.5 GHz, together with external mixers and filters. The modulated RF waveform is then pre-amplified by means of two Analog Devices’ driver PAs, HMC499LC4 and HMC943ALP5DE, with 17 dB and 23 dB gain, respectively, such that the integrated PAs of the Anokiwave AWMF-0129 active antenna array are driven towards saturation. The transmit signal propagates over-the-air (OTA) and is captured by a horn antenna at the observation receiver, such that the receiving antenna system is well within the main transmit beam. At the receiver side, another Keysight N5183B-MXG and a mixing stage are used to downconvert the signal back to IF. Then, the Keysight DSOS804A oscilloscope is utilized as the actual digitizer, including also built-in filtering, and the signal is taken to baseband and processed in a host PC, where the DPD learning and predistortion are executed. The OTA measurement system is basically following the measurement procedures described in [18], [34], specifically the measurement option utilizing the beam-based directions. In these measurements, the observation receiver provides I/Q samples at 7x oversampled rate.

In the DPD measurements, we adopt 5G NR FR-2 OFDM signal with SCS of 60 kHz, and consider active PRB counts of 132 and 264, mapping to 100 MHz and 200 MHz channel bandwidths, respectively [18]. In this case, 5 ILA iterations are adopted, each containing 50,000 samples. Example OTA linearization results are illustrated in Fig. 10 measured at EIRP of +40.5 dBm, where the received spectra with the proposed 3rd order SPH with Q = 15, Δx = 1 and the reference MP DPD are shown, while the no-DPD case is also shown for comparison. The order of the adopted MP DPD is P_{MP} = 9 while the memory parameters of the MP DPD and the SPH DPD are M_{MP} = 3 and M_{SP} = 3, respectively. These correspond to total of 20 and 21 free parameters in the MP DPD and SPH DPD cases, respectively. As mentioned already in the introduction, the OTA ACLR requirements at FR-2 are quite clearly relaxed, compared to the classical 45dB number at FR-1, with 28 dB defined as the ACLR limit in the current NR Release-15 specifications [18]. Additionally, 64-QAM is currently the highest supported modulation scheme at FR-2, heaving 8 % as the required EVM. In both channel bandwidth cases, the initial ACLR is around 26 dB at the right adjacent channel, hence linearization is indeed required if the same output power is to be maintained. Table VI shows the measured numerical ACLR and EVM values, indicating good amounts of linearization gain and that the EVM and ACLR requirements can be successfully met. It is noted that the initial ACLR of 26 dB corresponds already to a very nonlinear starting point, however, in our future work we pursue further measurements with even more saturated PA units.
Fig. 10. Illustration of OTA linearization of the Anokiwave AWMF-0129 active antenna array, when (a) NR 100 and (b) NR 200 MHz transmit signals are applied, measured at EIRP of +40.5 dBm. The SPH DPD number of control points \( Q = 15 \), the MP DPD order is \( P_{MP} = 9 \) while the memory parameters are \( M_{MP} = 3 \) and \( M_{SP} = 3 \).

### Table VI
Summary of linearization performance of the Anokiwave AWMF-0129 active antenna array, with 100 MHz and 200 MHz NR channel bandwidths, measured at +40.5 dBm EIRP

|                | DPD running complexity | DPD perf., 100 MHz | DPD perf., 200 MHz |
|----------------|------------------------|--------------------|--------------------|
|                | P M Q Δ\(x\) FLOPs/sample Mul/sample | EVM (%) ACLR L/R (dB) | EVM (%) ACLR L/R (dB) |
| No DPD         | - - - - - 0 0          | 12.10 27.50 / 26.31 | 12.13 28.40 / 26.10 |
| SPH DPD        | 3 3 15 1 68 46         | 6.30 40.00 / 37.50  | 6.25 38.60 / 35.00  |
| MP DPD         | 9 3 - - - 172 93       | 6.10 42.80 / 41.50  | 6.13 41.30 / 38.10  |

V. Conclusions

In this paper, a novel digital predistorter concept was proposed for power amplifier linearization, with particular emphasis on reduced main path and parameter learning complexities while still allowing for good linearization performance. The proposed predistortion concept divides the input envelope range into several pieces or ranges, and relies on complex spline interpolation to estimate or approximate the instantaneous nonlinear behavior. This is complemented with an FIR filter, for memory modeling purposes, composing thus as a whole a spline-interpolated Hammerstein like DPD solution. Gradient based iterative parameter learning algorithms were also derived, allowing to estimate the unknown spline control points and the unknown FIR filter parameters in a decoupled manner. A vast amount of different measurement-based experiments were provided, covering successful linearization of different local-area, medium range and wide-area/macro PA samples at sub-6 GHz bands. Additionally, a 28 GHz state-of-the-art active antenna array was successfully linearized. The measured linearization performance results, together with the provided explicit complexity analysis, show that the proposed spline-interpolated DPD concept provides a very appealing complexity-performance trade-off, compared to, e.g., a memory polynomial type DPD. The proposed method may find good applications especially in future mmWave/FR-2 5G NR networks, where the linearity requirements are relaxed but power efficiency is of utmost importance, thus calling for reduced-complexity processing solutions suitable for very nonlinear wideband systems and PAs. Our future work will focus on developing DPD solutions that allow for even deeper compression of the PA units in active antenna arrays.

### Appendix A
Derivation of the SPH DPD Learning Rules

In this Appendix, we provide detailed derivations for the gradient-based update rules for the linear memory filter \( h_n = h_n^{re} + jh_n^{im} \), and the spline control points \( c_n = c_n^{re} + jc_n^{im} \). First, the derivative presented in (10) can be expanded as

\[
\frac{\partial J(h_n, c_n)}{\partial h_n} = \frac{\partial e[n]}{\partial h_n} e^*[n]
\]

\[
= e^*[n] \frac{\partial e[n]}{\partial h_n} + e[n] \frac{\partial e^*[n]}{\partial h_n}
\]

\[
= -e^*[n] \frac{\partial ySP[n]}{\partial h_n} + e[n] \frac{\partial ySP^*[n]}{\partial h_n}
\]

\[
= e^*[n] \left[ \frac{\partial ySP[n]}{\partial h_n^{re}} + j \frac{\partial ySP^*[n]}{\partial h_n^{im}} \right]^{*} - e[n] \left[ \frac{\partial ySP[n]}{\partial h_n^{re}} \right]^{*} + \left( \frac{\partial ySP^*[n]}{\partial h_n^{im}} \right)^{*},
\]

and since \( h_n^H = h_n^{T, re} - jh_n^{T, im} \), the subderivatives with respect
the real and imaginary parts can be calculated as
\[
\frac{\partial y_{SP}[^n]}{\partial h_{re}[^n]} = \frac{\partial h_{re}[^n]s_n}{\partial h_{re}[^n]} = \frac{\partial h_{re}[^n]T^{-1}s_n}{\partial h_{re}[^n]} = s_n, \\
\frac{\partial y_{SP}[^n]}{\partial h_{im}[^n]} = \frac{\partial h_{im}[^n]s_n}{\partial h_{im}[^n]} = -j \frac{\partial h_{im}[^n]T^{-1}s_n}{\partial h_{im}[^n]} = -js_n.
\]

Gathering these terms together, the derivative in (10) reads
\[
\frac{\partial J(h_n, c_n)}{\partial h_n} = -e^*[n] \left( s_n + j(-j)s_n \right) - e[n] \left( s^*_n + js^*_n \right) \\
= -2e^*[n]s_n.
\]

Thus, it directly follows that
\[
h_{n+1} = h_n + \mu_h[e^*[n]]s_n.
\]

Next, the derivative in (11) against \( c_n \) can be expressed as
\[
\frac{\partial J(h_n, c_n)}{\partial c_n} = \frac{\partial e[n]e^*[n]}{\partial c_n} \\
= \ldots \\
= -e^*[n] \left[ \frac{\partial y_{SP}[n]}{\partial c_{re}[^n]} + j \frac{\partial y_{SP}[n]}{\partial c_{im}[^n]} \right] \\
- e[n] \left[ \left( \frac{\partial y_{SP}[n]}{\partial c_{re}[^n]} \right)^* + j \left( \frac{\partial y_{SP}[n]}{\partial c_{im}[^n]} \right)^* \right],
\]

where the individual derivatives against the real-part of \( c_n \) can be computed as
\[
\frac{\partial y_{SP}[n]}{\partial c_{re}[^n]} = \frac{\partial h_{re}[^n]s_n}{\partial c_{re}[^n]} = \frac{\partial h_{re}[^n]^T h_n^*}{\partial c_{re}[^n]} = \left( \frac{\partial y_{SP}[n]}{\partial c_{re}[^n]} \ldots \frac{\partial y_{SP}[n-M_{SP}+1]}{\partial c_{re}[^n]} \right) h_n^*,
\]

with
\[
\frac{\partial y_{SP}[n-k]}{\partial c_{re}[^n]} = \frac{\partial x_{SP}[n-k]g_{n-k}^T c_{re}[^n]}{\partial c_{re}[^n]} = x_{SP}[n-k]g_{n-k},
\]

for arbitrary \( k \). Therefore,
\[
\frac{\partial y_{SP}[n-k]}{\partial c_{re}[^n]} = \Sigma_n^T X_n h_n^*.
\]

The corresponding derivatives with respect the imaginary part of \( c_n \) can be obtained by following the same procedure. Hence, it follows that
\[
\frac{\partial y_{SP}[n]}{\partial c_{im}[^n]} = j \Sigma_n^T X_n h_n^*.
\]

Gathering all the terms together, the overall derivative reads
\[
\frac{\partial J(h_n, c_n)}{\partial c_n} = -e^*[n] \left( \Sigma_n^T X_n h_n^* + j \Sigma_n^T X_n h_n^* \right) \\
- e[n] \left( \Sigma_n^T X_n h_n - j \Sigma_n^T X_n h_n \right) \\
= -2e[n] \Sigma_n^T X_n h_n^*.
\]

Thus, it follows that
\[
c_{n+1} = c_n + \mu_c[e[n]] \Sigma_n^T X_n h_n.
\]
[25] M. Scarpiniti, D. Commiello, R. Parisi, and A. Uncini, “Hammerstein uniform cubic spline adaptive filters: Learning and convergence properties,” *Signal Processing*, vol. 100, pp. 112–123, 2014.

[26] —, “Novel cascade spline architectures for the identification of nonlinear systems,” *IEEE Trans. Circuits Syst. I*, vol. 62, no. 7, pp. 1825–1835, Jul. 2015.

[27] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 2001.

[28] R. Raich, Hua Qian, and G. T. Zhou, “Orthogonal polynomials for power amplifier modeling and predistorter design,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1468–1479, Sep. 2004.

[29] C. Mollen, E. G. Larsson, U. Gustavsson, T. Eriksson, and R. W. Heath, “Out-of-band radiation from large antenna arrays,” *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 196–203, Apr. 2018.

[30] M. Abdelaziz, L. Anttila, A. Brihuega, F. Tufvesson, and M. Valkama, “Digital Predistortion for Hybrid MIMO Transmitters,” *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 445–454, Jun. 2018.

[31] A. Brihuega, L. Anttila, M. Abdelaziz, and M. Valkama, “Digital Predistortion in Large-Array Digital Beamforming Transmitters,” in 2018 52nd Asilomar Conference on Signals, Systems, and Computers, Oct. 2018, pp. 611–618.

[32] X. Liu, Q. Zhang, W. Chen, H. Feng, L. Chen, F. M. Ghannouchi, and Z. Feng, “Beam-Oriented Digital Predistortion for 5G Massive MIMO Hybrid Beamforming Transmitters,” *IEEE Trans. Microw. Theory Tech.*, vol. 66, no. 7, pp. 3419–3432, Jul. 2018.

[33] N. Tervo, J. Aikio, T. Tuovinen, T. Rahkonen, and A. Parssinen, “Digital predistortion of amplitude varying phased array utilising over-the-air combining,” in 2017 *IEEE MTT-S International Microwave Symposium (IMS)*, Jun. 2017, pp. 1165–1168.

[34] 3GPP Tech. Spec. 38.141-2, “NR: Base Station (BS) conformance testing, Part 2,” v15.1.0 (Release 15), March 2019.