Beam-Dependent Active Array Linearization by Global Feature-Based Machine Learning

Mattia Mengozzi[ID], *Graduate Student Member, IEEE*, Gian Piero Gibiino[ID], *Member, IEEE*,
Alberto M. Angelotti[ID], *Member, IEEE*, Corrado Florian[ID], *Member, IEEE*, and Alberto Santarelli[ID], *Member, IEEE*

Abstract— **An approach based on machine learning is proposed for the global linearization of microwave active beamforming arrays. The method allows for the low-complexity real-time update of the digital predistortion (DPD) coefficients by exploiting order-reduced model features, hence avoiding the need for repeated local DPD identification steps across the various operating conditions of the beamformer (e.g., different beam angles or RF power levels). The validation is performed by over-the-air (OTA) measurements of a 1 × 4 array operating at 28 GHz across 100-MHz modulation bandwidth (BW).**

Index Terms— **Beamforming array, linearization, machine learning, over-the-air (OTA) measurements.**

I. INTRODUCTION

ACTIVE beamforming arrays, nowadays being adopted in telecom systems, pose several challenges with respect to linearity specifications and design of digital predistortion (DPD). Classical single-input–single-output (SISO) DPD is affected by the intrinsic differences among the power amplifiers (PAs) composing the array, and by the dynamic loading conditions involved as the beam direction changes [1]. These effects can cause nonoptimal performance of the SISO DPD trained along one single beam direction as the beam is oriented to other steering angles [2], [3].

One possible solution to address the nonlinear distortion due to beam dependency could consist of a real-time identification of the DPD coefficients across the varying operating conditions [4]. However, this generally requires high-performance feedback loops as well as high-complexity HW and SW, and might finally result unfeasible in actual deployments. The DPD architecture in [5] realizes the real-time linearization of a beamforming device without a feedback loop, but it requires to sense the output of each PA within the array during the DPD training phase. Moreover, a complex nonlinear predictive function needs to be calculated in real time for every DPD coefficient, increasing the burden on signal processing.

In this work, we realize a real-time beam-dependent DPD (BD-DPD) without any feedback or need to collect any information about the single PAs. A global DPD beam adapter (BA)
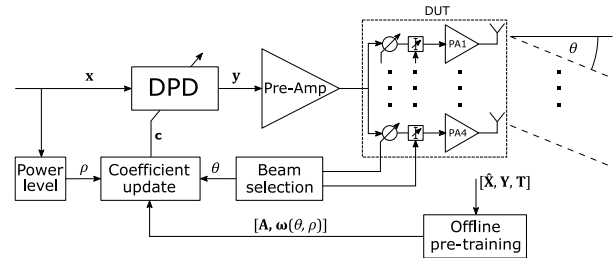
Fig. 1. Block diagram of the adopted DPD configuration for active beamforming arrays.

is preliminarily identified from a reduced set of measurements at the air interface, obtaining all the information needed for adjusting the DPD coefficients with a minimum number of predictive functions.

II. BEAM-DEPENDENT LINEARIZATION

A. DPD Architecture

The proposed BD-DPD architecture (see block diagram in Fig. 1) aims at identifying an open-loop BA. The purpose of the BA is to adjust the DPD coefficients along with the operating conditions in a predictive way, i.e., without any feedback from the output of the array. Given a DPD model that is linear-in-the-parameters, an open-loop BD-DPD could be represented as $y = Xc(\theta, \rho)$, where y is the output of the predistorter, X is the model regression matrix based on the predistorter input x, and c are the varying predistorter coefficients, which are a function of the beam angle (θ) and input signal peak-to-average power ratio (PAPR) (ρ).

The approach is deployed in two distinct phases. First, an offline *pre-training* aims at identifying the parameters of the open-loop BD-DPD generator model. In this phase, the PA array is linearized for a subset Γ of operating conditions, and the data obtained from this linearization procedure are used to extract the BA model $c(\theta, \rho)$. Then, the online linearization is performed, where the DPD coefficients are adjusted by the BA in a predictive way, based on the given beam angle and input signal PAPR. This DPD update is applied in open loop without any real-time feedback.

B. Feature-Based Model Reduction

Considering the many possible states involved in the operation of a beamformer array, extracting a global DPD model could in principle require a very large number of different sets of DPD parameters, potentially one for each state (θ, ρ). This option is prone to over-fitting issues, and its practical handling would be too complex for real-time BA. Therefore, the initial pre-training procedure targets the reduction of these coefficients by identifying a reduced set of so-called model

features. Indeed, it has been recently reported in [6] that the behavior of a PA can effectively be approximated across different operating conditions by a reduced set of features exploiting a transformation matrix A

$$y = Xc(\theta, \rho) \simeq XA\omega(\theta, \rho) \tag{1}$$

where $\omega(\theta, \rho)$ is a vector of coefficients acting as a BA and displaying a lower dimensionality with respect to $c(\theta, \rho)$. The beamformer array is first linearized through the iterative learning control (ILC) approach [7] across the pre-training subset, obtaining one DPD set c_k composed of Γ coefficients for each of the K different operating conditions

$$c_k = (X_k^H X_k)^{-1} X_k^H y_k \tag{2}$$

where y_k and X_k are the $N \times 1$ (where N is the number of time-domain complex samples) vectors representing, respectively, the optimal predistorted signal by ILC and the DPD regression matrix, both corresponding to the kth operating condition. Then, A can be identified by the following minimization problem:

$$\min_{(A)} \sum_k \|X_k A\omega_k - y_k\|^2 \tag{3}$$

where ω_k is the $S \times 1$ reduced set of coefficients for the kth operating condition. As reported in Algorithm 1, a generalized QR-SVD algorithm [6] is here used to solve (3) and obtain A by taking as inputs all DPD coefficients and regression matrices structured as

$$C = [c_1, \ldots, c_K]; \quad \hat{X} = [X_1^\top, \ldots, X_{\hat{K}}^\top]^\top \tag{4}$$

where $\hat{X}$ is the vertical concatenation of the nonrepeated $\hat{K}$ X_k matrices in the pre-training set.

The reduced set ω_k of DPD coefficients corresponding to the kth operating condition can be finally identified as

$$\omega_k = ((X_k A)^H X_k A)^{-1} (X_k A)^H y_k. \tag{5}$$

Eventually, a global BA model $\omega(\theta, \rho)$ can be obtained by interpolating ω_k across the values of beam angle (θ_k) and input power level (ρ_k). In this work, the interpolation has been performed either using a 2-D cubic spline model or a polynomial model. In order to linearize the array, the actual DPD coefficients can be easily updated by $c = A\omega(\theta, \rho)$, using the value of $\omega(\theta, \rho)$ at the given beam angle and input signal PAPR. Overall, the proposed SVD-based procedure allows to automatically select a number $S \ll K$ of significant features for the global BA behavior, allowing to greatly reduce the number of coefficients required in order to provide effective

---

Algorithm 1 Generalized QR-SVD Algorithm

Data: $\hat{X} \in \mathbf{C}^{(N\hat{K}) \times \Gamma}$, $C \in \mathbf{C}^{\Gamma \times K}$
Result: $A \in \mathbf{C}^{\Gamma \times S}$
1: Perform QR decomposition on $\hat{X}$: $\hat{X} = \hat{Q}\hat{R}$
2: $D = \hat{R}C$
3: Perform QR decomposition on D: $D = U\Sigma V$
4: Select an arbitrary number S of singular values of Σ
5: Form U': Compress U selecting the S columns that correspond to the selected singular values of Σ
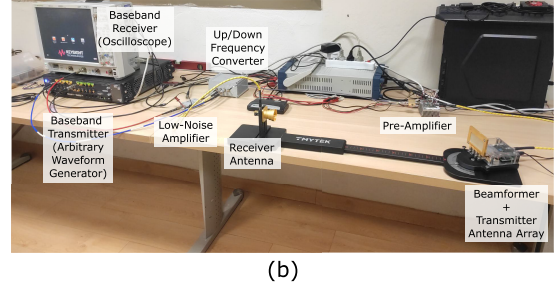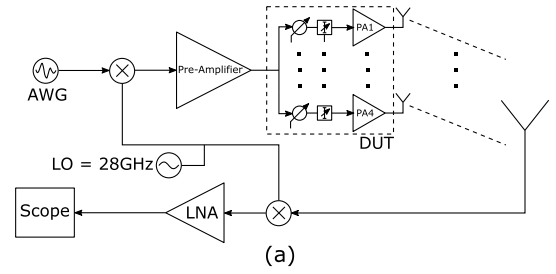6: $A = \hat{R}^{-1}U'$

Fig. 2. (a) Block diagram and (b) photograph of the OTA measurement setup.

beam-adaptation. Indeed, instead of storing and interpolating across a total of $\Gamma \times K$ DPD coefficients arising from pre-training at the K different operating conditions, it is sufficient to store a much smaller A matrix ($\Gamma \times S$) and the $S \times 1$ interpolated functions $\omega(\theta, \rho)$.

III. OTA MEASUREMENT SETUP

The implemented over-the-air (OTA) setup (see Fig. 2) operates at 28 GHz. Up/down conversion from an intermediate frequency (IF) of 1.55 GHz is performed by a two-way mixer (TMYTEK UDBox) with a shared local oscillator. A Keysight M8190 arbitrary waveform generator is used to generate modulated signals at IF. The IF receiver is composed of a Mini-Circuits ZRL-2150+ low-noise amplifier and a wideband oscilloscope (Keysight DSO9254A) allowing for 500-MHz acquisition bandwidth (BW).

The device-under-test (DUT) is a 1×4 active beamformer array (TMYTEK BBoard), which includes signal splitting and a separate power amplification/attenuation/phase control in each signal path. The beamformer is driven by an Analog Devices HMC943 preamplifier, pushing the four integrated PAs into their nonlinear region at up to 5-dB compression, resulting in 10 dBm of maximum available output power for each way. The signal is OTA-transmitted by a four-way patch antenna array and received in the far-field by a horn antenna. The excitation signals are random-phase multitone signals with BW = 100 MHz, whose statistics have been matched to a Gaussian OFDM-like signal. The linearization performance is measured in terms of adjacent channel power ratio (ACPR) and error vector magnitude (EVM).

IV. EXPERIMENTAL RESULTS

A. DPD Identification

The adopted pre-training set is a rectangular sampling of $K = 50$ different operating points, selected across the θ, ρ variables' space as depicted in Fig. 3 (blue circles). The beam angle is swept in 5° steps, whereas the input power accounts for signals with five different PAPR from 7 to 11 dB
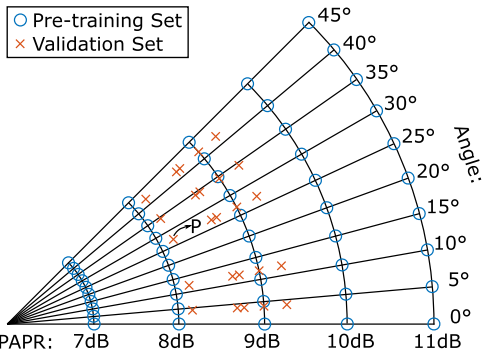
Fig. 3. Array operating conditions used as a pre-training set (blue circles) and as a validation set (red crosses).
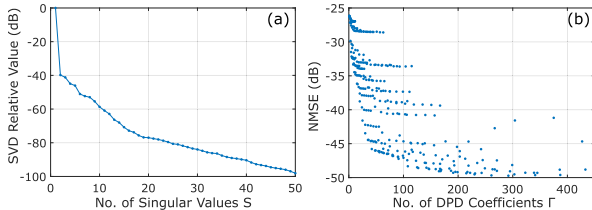

Fig. 4. (a) Singular values of matrix $\boldsymbol{D}$ as from Algorithm 1. (b) Model fitting error (NMSE) of the GMP-based predistorter.

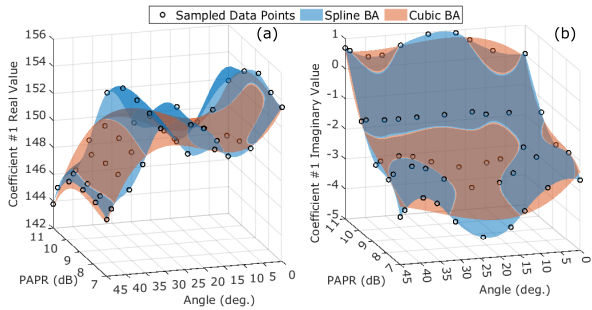Fig. 5. Cubic polynomial and spline interpolation of the first reduced feature across all tested conditions. (a) Real and (b) imaginary part.
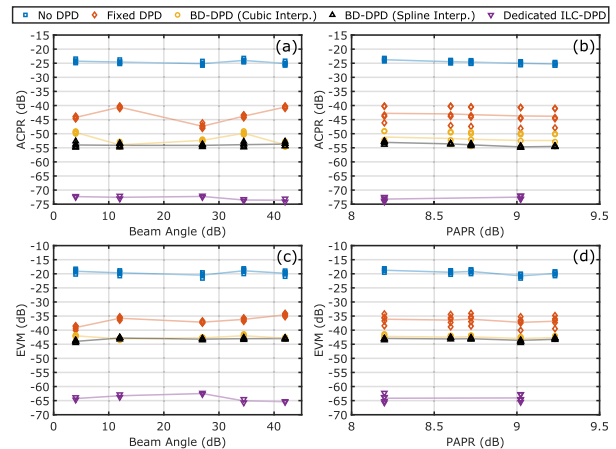

Fig. 6. DPD performance comparison. (a) ACPR and (c) EVM in the validation points plotted across beam angle. (b) ACPR and (d) EVM in the validation points plotted across PAPR. The continuous line indicates the average value in a given DPD configuration across all tested operating conditions.
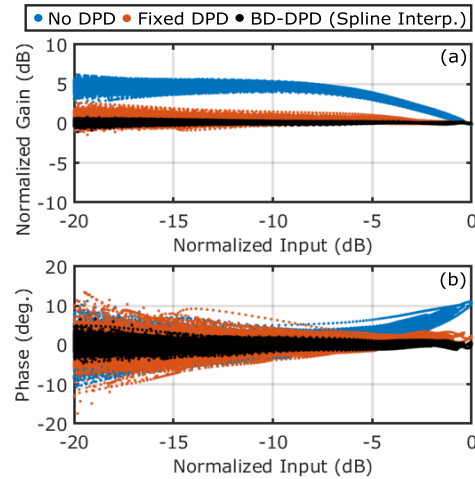

Fig. 7. (a) Gain and (b) phase characteristics of the predistorted and nonpredistorted array in the test condition P as from Fig. 3 ($\theta = 27°$ and PAPR = 8.2 dB).

and the same peak power. Beyond this particular choice, the pre-training set could be freely defined by adopting any other sampling or design-of-experiment strategy, e.g., latin hypercube, etc., [8].

By applying Algorithm 1 and according to the singular values extracted for $\boldsymbol{D}$ [see Fig. 4(a)], a dimensionality of $S = 5$ for $\boldsymbol{\omega}(\theta, \rho)$ has been found suitable for the array-under-test, as the residual error is less than the -45 dB of normalized mean square error (NMSE) targeted in this work. As an example, Fig. 5 reports the interpolation results across the beam angle and PAPR value for the first complex component of $\boldsymbol{\omega}(\theta, \rho)$, showing good fitting properties both for the adopted third order polynomial model and the 2-D cubic spline. The parametric structure adopted for the predistorter is the generalized memory polynomial (GMP) model [9]. The GMP nonlinear, memory, and cross memory orders have been, respectively, selected as 7, 9, and 1 to ensure -49 dB for the NMSE between the nonparametric predistorted signal (found with ILC) and the one modeled by the GMP-based predistorter [see Fig. 4(b)].

B. DPD Validation

The BD-DPD is validated in 25 independent operating regimes with different beam angles and PAPR values,

as depicted in Fig. 3 (red crosses). Fig. 6 depicts the performance of the proposed BD-DPD compared against a fixed DPD approach, i.e., DPD using a fixed set of coefficients extracted from a single operating regime ($\theta = 0°$, PAPR = 11 dB) and kept constant across beam direction and input signal PAPR. For reference purposes, Fig. 6 also reports the case without DPD as well as the case for best achievable DPD performance, obtained by means of a dedicated nonparametric ILC-DPD extracted at every single operating condition.

Fig. 6 clearly shows a general reduction of the distortion induced by the dependency on beam directions and PAPR for both implemented BA interpolations. As from Fig. 6(a) and (b), the BD-DPD leads to an improvement of ~5 and ~10 dB in terms of ACPR for the cubic and the spline BA, respectively. The improvement in terms of EVM is ~5 dB for both implemented methods, as shown in Fig. 6(c) and (d). Fig. 7(a) and (b) report the gain and AM/PM characteristics of the nonpredistorted and predistorted array in a particular operating condition of the validation set ($\theta = 27°$, PAPR = 8.2 dB), whereas the corresponding output spectra is shown in Fig. 8.
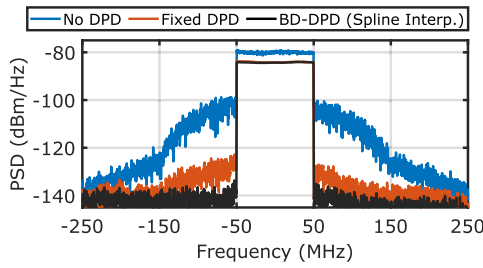
Fig. 8. Spectrum of the predistorted and nonpredistorted array in the condition P as from Fig. 3 ($\theta = 27°$ and PAPR = 8.2 dB).

V. CONCLUSION

A BD-DPD architecture exploiting feature-based model reduction for beamformer arrays has been proposed. The feature-based reduction applied at the array level does not depend on the number of PAs within the array, and it allows for a low-complexity update of the DPD of the array according to the beam direction and RF power level, without the need to extract a separate DPD coefficient set for every different operating regime. The proposed BD-DPD is effective in significantly reducing DUT in-band distortion and residual out-of-band regrowth across all tested beam directions and PAPR levels.

ACKNOWLEDGMENT

REFERENCES

[1] A. Brihuega, M. Turunen, L. Anttila, T. Eriksson, and M. Valkama, "On the power and beam dependency of load modulation in mmWave active antenna arrays," in *Proc. 50th Eur. Microw. Conf. (EuMC)*, Jan. 2021, pp. 680–683.

[2] E. Ng, Y. Beltagy, G. Scarlato, A. B. Ayed, P. Mitran, and S. Boumaiza, "Digital predistortion of millimeter-wave RF beamforming arrays using low number of steering angle-dependent coefficient sets," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 11, pp. 4479–4492, Nov. 2019.

[3] M. Mengozzi, G. P. Gibiino, A. M. Angelotti, C. Florian, and A. Santarelli, "Over-the-air digital predistortion of 5G FR2 beamformer array by exploiting linear response compensation," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2022, pp. 394–397.

[4] X. Liu et al., "Beam-oriented digital predistortion for 5G massive MIMO hybrid beamforming transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3419–3432, Jul. 2018.

[5] X. Wang et al., "Digital predistortion of 5G multiuser MIMO transmitters using low-dimensional feature-based model generation," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 3, pp. 1509–1520, Dec. 2022.

[6] Y. Li, X. Wang, and A. Zhu, "Complexity-reduced model adaptation for digital predistortion of RF power amplifiers with pretraining-based feature extraction," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 3, pp. 1780–1790, Mar. 2021.

[7] 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, Sep. 2016.

[8] P. Barmuta, G. P. Gibiino, F. Ferranti, A. Lewandowski, and D. M. M.-P. Schreurs, "Design of experiments using centroidal Voronoi tessellation," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 11, pp. 3965–3973, Nov. 2016.

[9] 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.