electronics

MDPI

Article

GaN Power Amplifier Digital Predistortion by Multi-Objective Optimization for Maximum RF Output Power

Mattia Mengozzi [1,*], Gian Piero Gibiino [*], Alberto M. Angelotti , Corrado Florian and Alberto Santarelli

Department Electrical, Electronic, and Information Engineering (DEI) 'G. Marconi', University of Bologna, 40126 Bologna, Italy; alberto.angelotti@unibo.it (A.M.A.); corrado.florian@unibo.it (C.F.); alberto.santarelli@unibo.it (A.S.)

* Correspondence: mattia.mengozzi3@unibo.it (M.M.); gianpiero.gibiino@unibo.it (G.P.G.)

Abstract: While digital predistortion (DPD) usually targets only the linearity performance of the radio–frequency (RF) power amplifier (PA), this work addresses more than a single PA performance metric exploiting a multi-objective optimization approach. We present a predistorer learning procedure based on a constrained optimization algorithm that maximizes the RF output power, while guaranteeing a prescribed linearity level, i.e., a maximum normalized mean square error (NMSE) or adjacent-channel power ratio (ACPR). Experimental results on a Gallium Nitride (GaN) PA show that the proposed approach outperforms the classical indirect learning architecture (ILA), yet using the same predistorter structure with predetermined nonlinearity and memory orders.

Keywords: digital predistortion; multi-objective optimization; power amplifier; linearization techniques

Citation: Mengozzi, M.; Gibiino, G.P.; Angelotti, A.M.; Florian, C.; Santarelli, A. GaN Power Amplifier Digital Predistortion by Multi-Objective Optimization for Maximum RF Output Power. *Electronics* **2021**, *10*, 244. https://doi.org/10.3390/electronics10030244

Received: 30 November 2020
Accepted: 18 January 2021
Published: 21 January 2021

1. Introduction

Digital predistortion (DPD) is a widely used technique to improve the linearity performance of a power amplifier (PA) at radio–frequency (RF), and it is increasingly becoming a fundamental part of the RF transmitting system. In recent years, more advanced PA topologies are being adopted for improved power efficiency, e.g., load- and supply-modulated PAs, or PA arrays, which inevitably introduce higher levels of distortion [1–6]. Moreover, PAs for microwave and millimeter-wave applications are nowadays implemented in Gallium Nitride (GaN) technology, which is affected by spurious dispersive phenomena causing peculiar behaviors like soft compression and long-term memory effects [7–9]. Indeed, the continuous push for broadband operation will incrementally expose to the presence of nonlinear dynamic effects [10]. Hence, novel PA modeling and linearization approaches must be envisioned [11–15].

In particular, DPD model formulation and coefficient identification techniques play a key role in the overall performance of the transmitter. Two main measurement-based identification methods are commonly adopted: the indirect learning architecture (ILA) [16,17] and the direct learning architecture (DLA) [18]. In the ILA, shown in Figure 1a, a postinverse model of the PA is identified, and its coefficients are then copied for the predistorter. This method, which leverages on a straightforward linear least-square estimation, is based on the assumption that the postinverse model of the PA is identical to its preinverse [19]. In the DLA, shown in Figure 1b, the predistorter is directly identified from PA stimulus–response pairs. In this case, the identification of the preinverse is obtained from a preliminary identification of the forward PA model, and from the subsequent model inversion. However, given the presence of nonlinear dynamic effects, such identification is often not trivial. In fact, the forward model inversion is usually achieved by means of nonlinear optimization algorithms [18].

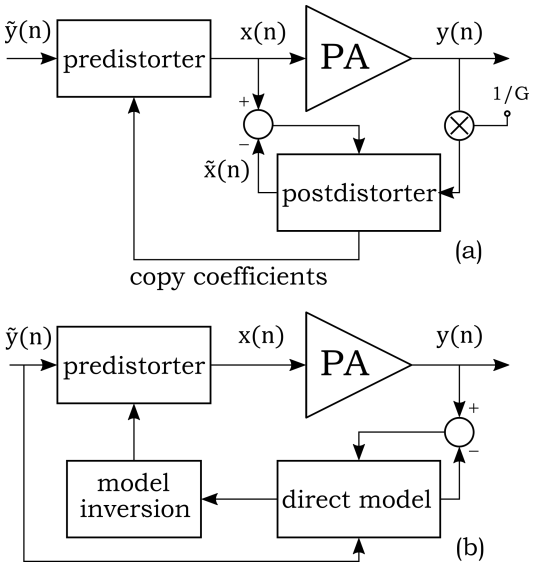

Figure 1. Block diagram for digital predistortion using (**a**) Iterative Learning Architecture (ILA) and (**b**) Direct Learning Architecture (DLA).

Both DLA and ILA architectures can also be implemented using adaptive iterative procedures involving multiple acquisitions, which might allow for the convergence to an improved set of DPD coefficients [20,21]. In this context, it is worth mentioning an alternative architecture referred to as iterative learning control (ILC) [22,23]. With ILC, the optimum predistorted signal is obtained at first (Figure 2a), while the identification of the actual parametric predistorter is performed in a subsequent step (Figure 2b). The first step involves the iterative identification of the predistorted signal in a non-parametric way, meaning that the algorithm does not directly target the DPD coefficients but the optimum input signal realizing the desired output. Despite requiring an increased number of measurements (one for each iteration), this approach allows for a model-independent solution. In the second step, the DPD coefficients can be eventually identified from the optimal input signal.

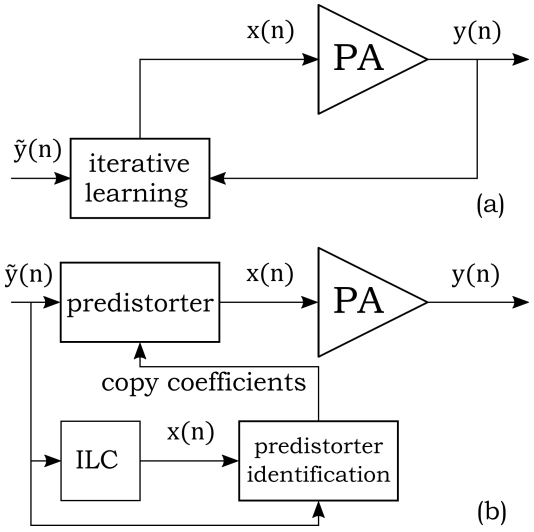

Figure 2. Block diagrams for DPD by Iterative Learning Control (ILC). (**a**) Identification of the predistorted signal (non-parametric) through iterative optimization; (**b**) DPD coefficients identification (parametric) using the optimum predistorted signal.

For all the mentioned architectures, the learning procedure only accounts for one single figure-of-merit (FoM), i.e., a linearity FoM, usually the time-domain normalized mean squared error (NMSE) between the actual and desired output signals. Indeed, these architectures can be generally seen as implementing unconstrained optimization, NMSE minimization being the only target. Indeed, this classical approach allows for operating linearly yet under deeper compression (i.e., with higher power efficiency), already enhancing the overall PA behavior with respect to the un-predistorted case. However, there is no guarantee that DPD will actually result in the best global PA performance achievable for the specifications of a given modulation standard. For example, there is no controlled way to limit the linearity performance up to a prescribed and sufficient level, although such a control would be useful to make room for maximizing another PA FoM of interest, e.g., the RF output power or power efficiency. In this context, it is reasonable to assume that multi-objective optimization approaches could deliver a better compromise among the conflicting PA specifications.

In this work, we propose a novel DLA-based DPD identification approach aimed at optimizing one (possibly, more than one) PA FoM, yet respecting a constraint. To this aim, we adopt the general architecture in Figure 3, leveraging on an iterative optimization routine. In particular, we implement a flexible constrained optimization algorithm capable of steering the learning process towards the desired joint optimum. This work is among the first ones [24–26] using multi-objective optimization for coefficients identification, while keeping a predetermined pre-inverse formulation and fixed model order.

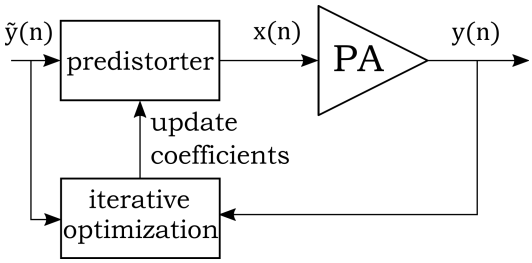

Figure 3. Block diagram for DLA-based DPD using an iterative nonlinear optimization.

The article is organized as follows: Section 2 describes in detail the proposed multi-objective algorithm. Section 3 reports the implementation of the algorithm in the case of a constrained optimization, showing the linearization of a GaN PA by maximizing the RF output power (hence, efficiency), yet respecting a prescribed linearity constraint. Conclusions are drawn in Section 4.

2. Multi-Objective DPD Optimization

Let us consider a predistorter to be implemented with a generic linear-in-the-parameters model. Labeling $\tilde{y}(n)$ and $x(n)$ as the input and output respectively (Figure 3), the generic predistorter structure can be expressed as

$$x(n) = \sum_{p=1}^{P} \phi_p h_p[\tilde{y}(n), \tilde{y}(n-1), \ldots, \tilde{y}(n-M)], \tag{1}$$

where ϕ_p are the model coefficients and h_p are suitable non-analytic functions of the past values of the complex signal $\tilde{y}(n)$, up to a memory duration M. The specific form of each function is fixed by the adopted model structure, with multivariate polynomials being particularly used in DPD applications [17]. As the nonlinear functions are pre-determined, the model can be fully identified once the coefficients ϕ_p are known. In this respect, Equation (1) can be recast in matrix form as

$$\mathbf{x} = \mathbf{H}\boldsymbol{\phi}, \tag{2}$$

where the following vector notation is introduced:

$$\mathbf{x} \stackrel{\text{def}}{=} \begin{bmatrix} x(n) & x(n-1) & \dots & x(n-N) \end{bmatrix}^T; \tag{3}$$

$$\boldsymbol{\phi} \stackrel{\text{def}}{=} \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_P \end{bmatrix}^T; \tag{4}$$

$$\mathbf{H} \stackrel{\text{def}}{=} \begin{bmatrix} h_{k,l} \end{bmatrix} = h_l[\tilde{y}(n-k), \tilde{y}(n-k-1), \dots, \tilde{y}(n-k-M)]; \quad \begin{matrix} k = 0, \dots, N \\ l = 1, \dots, P. \end{matrix} \tag{5}$$

Differently from the ILA approach, in which Equation (2) is directly solved for the coefficients of the postdistorter from the input–output measurements, the strategy adopted in this work leverages on a DLA identification of the coefficients in order to jointly optimize different performance metrics. Indeed, for the PA under test, all of the FoMs of interest are a function of the applied input signal $x(n)$ and, ultimately, for a fixed reference $\tilde{y}(n)$, depend just on the predistortercoefficients $\boldsymbol{\phi}$ through Equation (2).

Multi-objective optimization techniques have the goal of identifying Pareto-optimal solutions for $\boldsymbol{\phi}$ [24], in which no single objective can be improved without making at least another objective worse. This type of trade-off is commonly encountered in PA design, as an increase in linearity due to predistortion is usually accompanied by a reduction in the output power.

In this framework, standard DPD approaches can be seen as optimization of just a single linearity FoM, without taking explicitly into account output power or other (e.g., efficiency) metrics. Instead, in this work, we address the case where two objective functions f and g correspond to two conflicting PA specifications which are to be jointly optimized. In particular, we take g to represent a waveform linearity constraint while f corresponds to another conflicting performance metric, such as output power.

The general multi-objective optimization problem can be equivalently recast as a series of successive constrained optimizations [27], in which a single objective is maximized using standard methods, while fixed constraints are imposed on all the other objectives. In this case, $f(\boldsymbol{\phi})$ is taken to be the objective function to be maximized, and $g(\boldsymbol{\phi})$ as the constrained one. Then, the following optimization problem must be solved:

$$\max_{\boldsymbol{\phi}} \; f(\boldsymbol{\phi}) \quad \text{subject to} \;\; g(\boldsymbol{\phi}) < \eta \tag{6}$$

where η is a scalar quantity, representing the maximum allowable value for the constraint g. The solution to this problem explores just a single trade-off point between the two FoMs of interest. The full Pareto front of optimal solutions can then be explored by increasing or reducing the value of η and running the corresponding constrained optimization, in order to reconstruct the compromise between the two metrics under examination.

The constrained optimization algorithm proposed in this work for the solution of (6) is depicted in the flowchart in Figure 4. As a first step, if the constraint is not satisfied ($g(\boldsymbol{\phi}_i) < \eta$), an optimization direction $\mathbf{D}_g$ (relevant to the constraint g) must be obtained to resolve the next iteration.

In the proposed algorithm, the calculation of $\mathbf{D}_g$ follows the first step of the ILC approach by minimizing the instantaneous complex error between the (scaled) desired signal $\tilde{\mathbf{y}}$ and the measured output signal $\mathbf{y}$:

$$\mathbf{e}_i = \tilde{\mathbf{y}} \frac{\max\{\mathbf{y}_i\}}{\max\{\tilde{\mathbf{y}}\}} - \mathbf{y}_i. \tag{7}$$

Such a minimization can be achieved by an iterative procedure based on the Newton's method [22], where the optimized PA input at the ith iteration is found as:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \boldsymbol{\Gamma}_i \mathbf{e}_i \tag{8}$$

where $\mathbf{\Gamma}_i$ is a matrix containing the values of the instantaneous gain for each acquired sample $G_i(n) = \frac{y_i(n)}{x_i(n)}$:

$$\mathbf{\Gamma}_i = \begin{pmatrix} G_i(1) & 0 & 0 & 0 \\ 0 & G_i(2) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & G_i(N) \end{pmatrix}^{-1}. \tag{9}$$

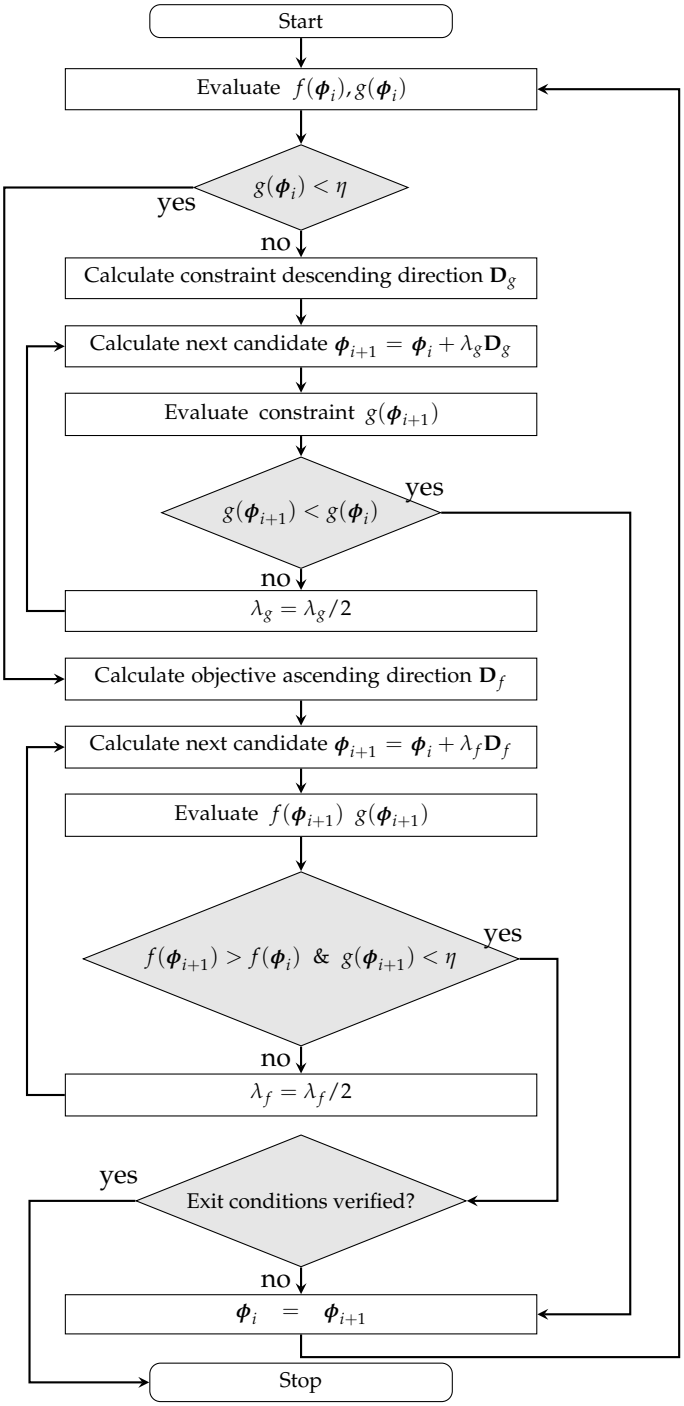

Figure 4. Flow chart of the constrained optimization algorithm exploited for DPD.

The gain matrix in (9) represents a first-order approximation of the Jacobian matrix as required by the Newton's method. However, other choices for $\mathbf{\Gamma}$ could also be adopted, e.g., a fixed value matrix or a finite-difference approximation of the Jacobian. While the

former could provide a more stable convergence of the iterative procedure, the latter could speed-up the algorithm at the cost of an higher number of function evaluations (i.e., measurements) at each iteration, necessary for the calculation of the finite differences.

In this work, differently from the classical ILC, $\mathbf{D}_g$ should be defined with respect to the coefficients $\boldsymbol{\phi}$. Then, we express the $\boldsymbol{\phi}$ in terms of $\mathbf{x}$ and $\mathbf{H}$ by inverting the matrix form in (2), obtaining

$$\boldsymbol{\phi} = \mathbf{H}^{\dagger}\mathbf{x}; \qquad \text{with} \quad \mathbf{H}^{\dagger} = (\mathbf{H}^{+}\mathbf{H})^{-1}\mathbf{H}^{+} \tag{10}$$

$\mathbf{H}^{\dagger}$ being the pseudoinverse of $\mathbf{H}$. By substituting this expression in (8), the formula for the ith iteration and the relative direction results in:

$$\boldsymbol{\phi}_{i+1} = \boldsymbol{\phi}_i + \lambda_g \mathbf{D}_g; \quad \mathbf{D}_g = \mathbf{H}^{\dagger}\boldsymbol{\Gamma}_i \mathbf{e}_i, \tag{11}$$

where the step size $\lambda_g \in [0, 1]$ is bisected until the constraint function g is improved, i.e., $g(\boldsymbol{\phi}_{i+1}) < g(\boldsymbol{\phi}_i)$. The progressive reduction of λ_g avoids overshoots, while improving the verification of the constraint. As soon as the constraint is satisfied, the next iteration goes towards the maximization of f.

For this second function, the optimization is based on a gradient descent:

$$\boldsymbol{\phi}_{i+1} = \boldsymbol{\phi}_i + \lambda_f \mathbf{D}_f; \quad \mathbf{D}_f = \boldsymbol{\nabla} f(\boldsymbol{\phi}_i), \tag{12}$$

where the gradient, corresponding to the optimization direction $\mathbf{D}_f$, is obtained by measuring a finite-differences approximation. The choice of a gradient-based approach for targeting the optimum allows for the maximum convergence speed in terms of number of iterations. However, it should be noted that estimating the partial derivatives requires two function evaluations for each complex DPD coefficient to be optimized. Therefore, an excessive number of DPD coefficient might offset the relatively low number of iterations. In Equation (12), the step size $\lambda_f \in [0, 1]$ is bisected until f improves, i.e., $f(\boldsymbol{\phi}_{i+1}) > f(\boldsymbol{\phi}_i)$, provided that the constraint remains valid. Indeed, depending on the regularity of the function to optimize, the gradient descent might not always identify the optimum direction. Hence, the progressive reduction of λ_f drives the method to a smoother convergence, while avoiding excessive oscillating behavior.

As far as the practical evaluation of the achieved optimum is concerned, it should be noted that the considered optimization aims at finding the maximum point of an unknown, possibly non-convex function. In this case, the definition of the stopping criteria is not trivial, considering that it is not possible to define a precise metric measuring the degree of global optimality [28]. Given these premises, the following approach is followed. As long as f is improved and $g < \eta$ satisfied, the algorithm evaluates a tolerance threshold $|f(\boldsymbol{\phi}_{i+1}) - f(\boldsymbol{\phi}_i)| \leq f_{\text{tol}}$. When exceeded, this threshold indicates that no further improvement is available for f, triggering the exit condition. Possible additional exit conditions, which can be combined with the main one, include reaching a maximum number of iterations ($N_i^{\max}$) or function evaluations ($N_f^{\max}$).

3. Measurement Results

While many different trade-offs could be considered [24], in this work, we have chosen the PA average RF output power and the linearity as conflicting FoMs. The average RF output power will correspond to the function to be maximized (f):

$$\bar{P}_{OUT} = \int_{\text{BW}} \frac{|Y(f)|^2}{2R_L} df \tag{13}$$

where BW is the signal bandwidth, $Y(f)$ is the spectrum of the measured output signal, and $R_L = 50\ \Omega$ is the matched-PA load resistance. The linearity constraint (g) will be

weighted by imposing a maximum limit to one of the following FoM, i.e., the Adjacent Channel Power Ratio (ACPR), or the NMSE:

$$\text{ACPR} = \frac{\int_{\text{SB}} |Y(f)|^2 df}{\int_{\text{BW}} |Y(f)|^2 df}; \quad \text{NMSE} = \frac{\sum_n |\tilde{G}\tilde{y}(n) - y(n)|^2}{\sum_n |\tilde{G}\tilde{y}(n)|^2}; \tag{14}$$

where SB stands for side-band, $\tilde{G}\tilde{y}(n)$ is the desired output, $y(n)$ is the measured output, and $\tilde{G}$ is the complex desired gain. Contrary to the ILA case, the proposed method does not fix an a priori value for $\tilde{G}$, which remains a degree of freedom for the optimization method.

The approach has been applied by means of the remote setup in Landin et al. [29]. Such a setup is based on a single benchtop instrument, the PXIe-5646R Vector Signal Transceiver (VST) by NI, featuring 200-MHz instantaneous bandwidth. With this setup, whose block diagram is shown in Figure 5, the user can set the baseband input signal using the vector signal generator within the VST, and measure the baseband output signal from the VST receiver, with a sample rate of $f_s = 200$ MSa/s for both input and output. The DUT is the cascade of a 40-dB linear driver amplifier and a GaN PA (Cree CGH40006-TB) biased in class-AB at 100 mA, operating at 2-GHz carrier frequency. The output of the PA is attenuated by 30 dB before reaching the VST receiver.

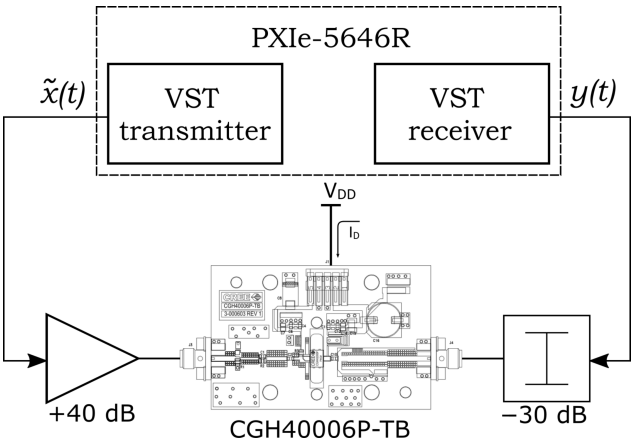

Figure 5. Measurement setup available at Chalmers University (RF WebLab) [29], used for the demonstration of the multi-objective linearization procedure.

The adopted input signal is a random-phase 2 k-tone with 20-MHz bandwidth and peak-to-average power ratio (PAPR) of ≃9 dB, resulting in $N = 20$ k samples for each acquisition, while the predistorter is described by a classical MP model formulation:

$$x(n) = \sum_{m=0}^{M} \sum_{k=1}^{K} \phi_{mk} \tilde{y}(n-m) |\tilde{y}(n-m)|^{k-1}, \tag{15}$$

where ϕ_{mk} are the DPD coefficients, K is the nonlinearity order, and M is the memory depth. The model in Equation (15) can be written in matrix form as in Equation (2), where $\mathbf{H}$ corresponds to:

$$\mathbf{H} = \begin{bmatrix} \tilde{y}(n) & \dots & \tilde{y}(n)|\tilde{y}(n)|^{(K-1)} & \dots & \tilde{y}(n-(M-1)) & \dots & \tilde{y}(n-(M-1))|\tilde{y}(n-(M-1))|^{(K-1)} \\ \tilde{y}(n-1) & \dots & \tilde{y}(n-1)|\tilde{y}(n-1)|^{(K-1)} & \dots & \tilde{y}(n-1-(M-1)) & \dots & \tilde{y}(n-1-(M-1))|\tilde{y}(n-1-(M-1))|^{(K-1)} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{y}(n-N) & \dots & \tilde{y}(n-N)|\tilde{y}(n-N)|^{(K-1)} & \dots & \tilde{y}(n-N-(M-1)) & \dots & \tilde{y}(n-N-(M-1))|\tilde{y}(n-N-(M-1))|^{(K-1)} \end{bmatrix}, \tag{16}$$

N being the total number of samples.

Throughout this work, K and M are predetermined and fixed values chosen as $K = 5$ and $M = 3$, which are not subject to optimization. This preliminary selection is performed by successively increasing the model order until the DPD (using ILA) achieved a sufficiently

low NMSE ($\simeq -30$ dB), resulting in a reasonable compromise between the necessary model complexity and the avoidance of over-fitting issues.

In addition, the total number of coefficients ($K \times M = 15$) is compatible with the dimensionality to be managed by the optimization algorithm in Section 2. However, the algorithm itself does not depend on the specific predistorter model, and it could be seamlessly used with other linear-in-the-parameters formulations, provided that a reasonably low number of DPD coefficients is considered.

Two separate datasets of signal realizations are acquired, one used for model coefficient learning by means of the optimization algorithm in Figure 4, and the other for DPD validation. Figures 6a,b report the linearization results when the ACPR (Figure 6a) and the NMSE (Figure 6b) are used as constraints within the algorithm. The blue curves represent the proposed multi-objective optimization results, obtained by sweeping the constraint value (η) for the ACPR (Figure 6a) and the NMSE (Figure 6b). The curves for the ILA case (red) are obtained by sweeping the target linearization gain, which is the only control variable in the ILA architecture. The linearization gain is swept up to an input signal peak power of -8 dBm, which corresponds to the maximum available power by the vector signal generator within the VST [29]. The black curves correspond to the non-predistorted case at different input powers.

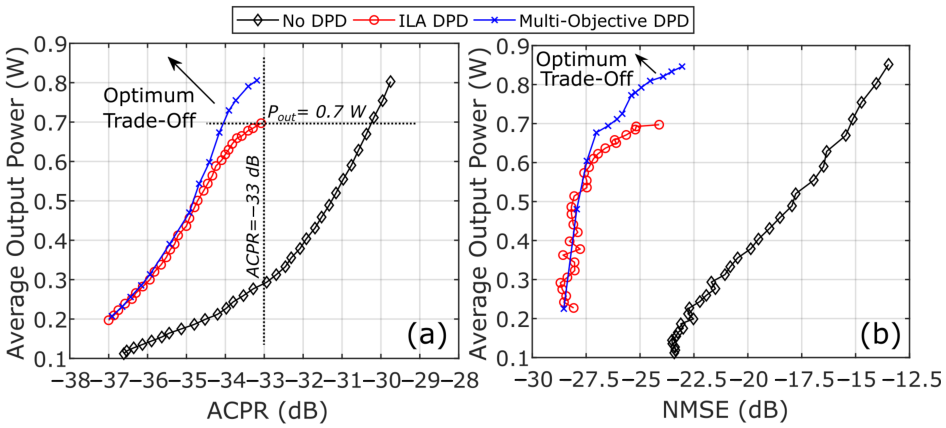

Figure 6. DPD performance when $\bar{P}_{OUT}$ is maximized and the ACPR (**a**) or the NMSE (**b**) is a constraint.

For both cases, the curve obtained with the proposed multi-objective algorithm lays closer to the $\bar{P}_{OUT}$—linearity optimum, substantially outperforming the classical ILA architecture for intermediate linearity levels (i.e., -35 to -33 dB of ACPR, or -27.5 to -23 dB of NMSE), yet using the same number of DPD coefficients. This demonstrates that the classical ILA method, which only accounts for linearity by NSME minimization and does not embed capabilities for Pareto front exploration, delivers inferior PA global performance. At the same time, for small values of ACPR and NMSE (i.e., high-linearity), the fundamental trade-off between linearity and RF output power gets increasingly narrow, so that no space for improvement is left for the output power maximization capability of the multi-objective approach. In this case, both methods will tend to show similar performance.

Figure 7a compares the acquired output spectra relative to the two different DPD configurations, as well as the case of the un-predistorted PA, for the same level of ACPR $\simeq -33$ dB. It can be clearly seen that, while the linearity performance is nearly the same for the three cases, the RF output power is maximum for the multi-objective DPD (corresponding to an average RF output power of 0.8 W), providing up to 14% $\bar{P}_{OUT}$ improvement with respect to ILA ($\bar{P}_{OUT} = 0.7$ W), whereas it is much smaller for the un-predistorted PA. In Figure 7b, instead, the three situations are shown for the same level of RF average output power ($\bar{P}_{OUT} = 0.7$ W), showing a better linearity performance for the multi-objective case (ACPR = -34 dB versus ACPR = -33 dB for ILA). For reference purposes, Figure 7 also includes the maximum output power performance without DPD at -4 dB compression

($\bar{P}_{OUT} \simeq 1.3$ W, ACPR $\simeq -26.5$ dB). Given the high level of distortion involved, this case falls out of the trade-off representation shown in Figure 6.

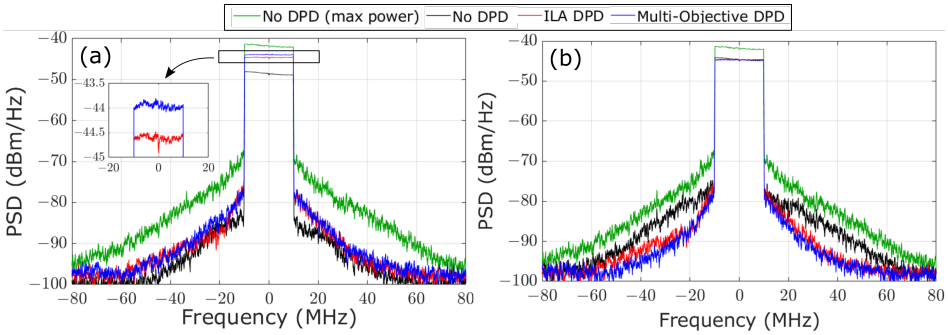

Figure 7. Spectra (not normalized) obtained with the proposed multi-objective optimization (blue), compared with the classical ILA approach (red) and the case without DPD (black). (**a**) same level of ACPR $\simeq -33$ dB, corresponding to the intercepts of the vertical dashed line in Figure 6a; (**b**) same level of $\bar{P}_{OUT} \simeq 0.7$ W, corresponding to the intercepts of the horizontal dashed line in Figure 6a. Green lines show the maximum output power performance without DPD at -4 dB compression ($\bar{P}_{OUT} \simeq 1.3$ W, ACPR $\simeq -27$ dB).

The same performance can also be observed in the instantaneous gain characteristics shown in Figure 8a,b, corresponding to Figure 7a,b, respectively. The PA is effectively linearized at a 4-dB compression point, with the multi-objective optimization further improving the performance in terms of RF output power. For completeness, Figure 9a,b show the corresponding instantaneous AM/AM characteristics.

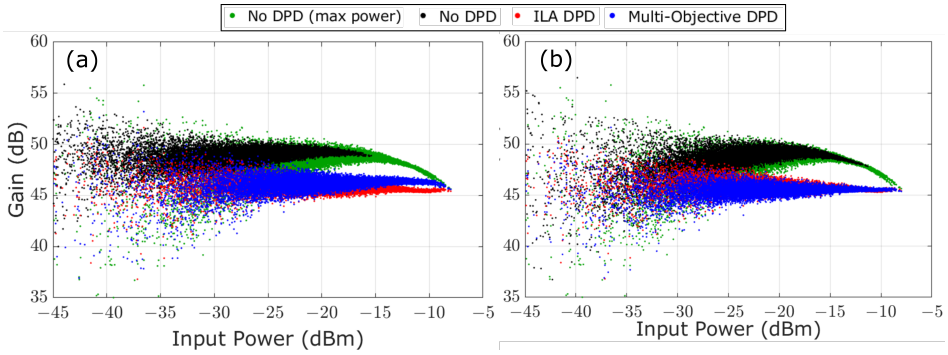

Figure 8. Instantaneous gain characteristics obtained with the proposed multi-objective optimization (blue), compared with the classical ILA approach (red) and the case without DPD (black). (**a**) same level of ACPR $\simeq -33$ dB, corresponding to the intercepts of the vertical dashed line in Figure 6a; (**b**) same level of $\bar{P}_{OUT} \simeq 0.7$ W, corresponding to the intercepts of the horizontal dashed line in Figure 6a. Green lines show the maximum output power performance without DPD at -4 dB compression ($\bar{P}_{OUT} \simeq 1.3$ W, ACPR $\simeq -27$ dB).

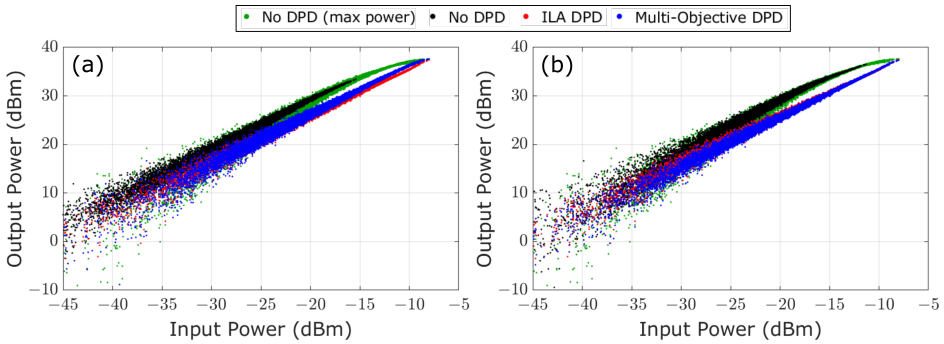

Figure 9. AM/AM characteristics obtained with the proposed multi-objective optimization (blue), compared with the classical ILA approach (red) and the case without DPD (black). (**a**) same level of ACPR $\simeq -33$ dB, corresponding to the intercepts of the vertical dashed line in Figure 6a; (**b**) same level of $\bar{P}_{OUT} \simeq 0.7$ W, corresponding to the intercepts of the horizontal dashed line in Figure 6a. Green lines show the maximum output power performance without DPD at -4 dB compression ($\bar{P}_{OUT} \simeq 1.3$ W, ACPR $\simeq -27$ dB).

Finally, Figure 10a reports the iterative behavior of the algorithm, highlighting the dynamic change of the optimization direction between the one kept to maintain the ACPR below the limit ($\mathbf{D}_g$), imposed using the instantaneous gain normalization, and the one for maximizing $\bar{P}_{OUT}$ ($\mathbf{D}_f$) by gradient descent. Figure 10b reports the same iterative trajectory of Figure 10a in the same domain as Figure 6. It can be seen that the found optimum can be reached within a limited number of iterations.

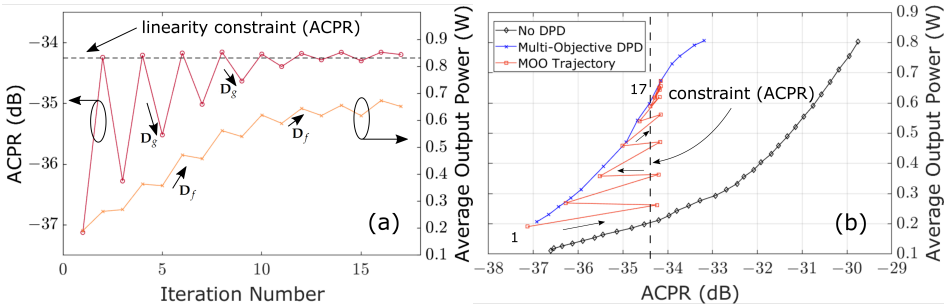

Figure 10. (**a**) Iterative behavior of the algorithm, where the conflicting optimization directions $\mathbf{D}_g$ (constraint) and $\mathbf{D}_f$ (maximization) are indicated; (**b**) iterative optimization trajectory (iterations 1 to 17) of the proposed multi-objective algorithm on the same domain as Figure 6.

4. Conclusions

The proposed multi-objective optimization method allows for exploring the trade-off between linearity and RF output power of a PA, resulting in a generalized DPD approach that can better enhance the global performance with respect to the classical ILA-based DPD. In particular, this is obtained by constraining one PA metric, and concurrently maximizing another one in an iterative fashion, eventually resulting in a tailored control of the predistorted input. Despite the impossibility to assess if the achieved maximization actually intercepts the global optimum of the complex multidimensional object function linking the FoMs with the DPD coefficients, the reached solution effectively delivers improved performance with respect to the classical ILA-based, single-objective DPD.

It is worth highlighting that this improvement just derives from the application of a different learning algorithm for the identification of the DPD coefficients, and that the same fixed DPD model order is used in all cases. Therefore, adopting this technique does not require any change in the DPD system configuration, nor any additional hardware.

The method should be seamlessly applicable in other test conditions and using different FoMs. As an example, a preliminary simulation-based evaluation of the trade-off between linearity and efficiency in supply-modulated PAs is reported in Mengozzi et al. [30]. Moreover, the proposed optimization framework is promising for the synthesis of opti-

mal input signals of next generation multiple-input transmitters, including PA arrays or load-modulated amplifiers [5,11]. Indeed, when multiple digitally-controlled inputs are involved, many different combinations of the inputs could reach a given performance metric, implying non-trivial preinverse identification and, possibly, sub-optimal performance. In this context, the proposed optimization framework can provide a useful numerical solution to exploit the higher dimensionality of multi-input systems. Future work will also include the enhancement of the method by combining PA incremental modeling techniques [31] to minimize the measurement burden.

Author Contributions: Conceptualization, M.M., G.P.G., and A.M.A.; Investigation, M.M., G.P.G., and A.M.A.; Methodology, M.M., G.P.G., A.M.A., and C.F.; Software, M.M.; Supervision, C.F. and A.S.; Validation, M.M. and G.P.G.; Visualization, M.M., G.P.G., C.F., and A.S.; Writing—original draft, M.M. and G.P.G.; Writing—review and editing, A.M.A., C.F., and Alberto Santarelli. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to thank Chalmers University for providing the remote RF weblab [29].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Darraji, R.; Ghannouchi, F.; Hammi, O. A dual-input digitally driven Doherty amplifier architecture for performance enhancement of Doherty transmitters. *IEEE Trans. Microw. Theory Tech.* **2011**, *59*, 1284–1293. [CrossRef]
2. Hausmair, K.; Gustafsson, S.; Sánchez-Pérez, C.; Landin, P.N.; Gustavsson, U.; Eriksson, T.; Fager, C. Prediction of nonlinear distortion in wideband active antenna arrays. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 4550–4563. [CrossRef]
3. Cappello, T.; Barton, T.; Florian, C.; Litchfield, M.; Popovic, Z. Multilevel supply-modulated Chireix outphasing with continuous input modulation. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 5231–5243. [CrossRef]
4. Cappello, T.; Pednekar, P.; Florian, C.; Cripps, S.; Popovic, Z.; Barton, T. Supply- and load-modulated balanced amplifier for efficient broadband 5G base stations. *IEEE Trans. Microw. Theory Tech.* **2019**, *67*, 3122–3133. [CrossRef]
5. Cappello, T.; Florian, C.; Niessen, D.; Paganelli, R.P.; Schafer, S.; Popovic, Z. Efficient X-band transmitter with integrated GaN power amplifier and supply modulator. *IEEE Trans. Microw. Theory Tech.* **2019**, *67*, 1601–1614. [CrossRef]
6. Jalili, F.; Tafuri, F.F.; Jensen, O.K.; Li, Y.; Shen, M.; Pedersen, G.F. Linearization Trade-Offs in a 5G mmWave Active Phased Array OTA Setup. *IEEE Access* **2020**, *8*, 110669–110677. [CrossRef]
7. Pedro, J.C.; Nunes, L.C.; Cabral, P.M. Soft compression and the origins of nonlinear behavior of GaN HEMTs. In Proceedings of the 2014 44th European Microwave Conference, Rome, Italy, 6–9 October 2014; pp. 1297–1300.
8. Tome, P.M.; Barradas, F.M.; Cunha, T.R.; Pedro, J.C. A Multiple-Time-Scale Analog Circuit for the Compensation of Long-Term Memory Effects in GaN HEMT-Based Power Amplifiers. *IEEE Trans. Microw. Theory Tech.* **2020**, *68*, 3709–3723. [CrossRef]
9. Angelotti, A.M.; Gibiino, G.P.; Florian, C.; Santarelli, A. Broadband Error Vector Magnitude Characterization of a GaN Power Amplifier using a Vector Network Analyzer. In Proceedings of the 2020 IEEE /MTT-S International Microwave Symposium (IMS), Los Angeles, CA, USA, 4–6 August 2020; pp. 747–75.
10. Rawat, K.; Roblin, P.; Koul, S.K. *Bandwidth and Efficiency Enhancement in Radio Frequency Power Amplifiers for Wireless Transmitters*; Springer Nature: Berlin, Germany, 2020.
11. Gibiino, G.P.; Santarelli, A.; Schreurs, D.; Filicori, F. Two-input nonlinear dynamic model inversion for the linearization of envelope-tracking RF PAs. *IEEE Microw. Wirel. Compon. Lett.* **2016**, *27*, 79–81. [CrossRef]
12. Dawar, N.; Sharma, T.; Darraji, R.; Ghannouchi, F.M. Linearisation of radio frequency power amplifiers exhibiting memory effects using direct learning-based adaptive digital predistoriton. *IET Commun.* **2016**, *10*, 950–954. [CrossRef]
13. Hausmair, K.; Landin, P.N.; Gustavsson, U.; Fager, C.; Eriksson, T. Digital predistortion for multi-antenna transmitters affected by antenna crosstalk. *IEEE Trans. Microw. Theory Techn.* **2017**, *66*, 1524–1535. [CrossRef]
14. Gibiino, G.P.; Łukasik, K.; Barmuta, P.; Santarelli, A.; Schreurs, D.; Filicori, F. A two-port nonlinear dynamic behavioral model of RF PAs subject to wideband load modulation. *IEEE Trans. Microw. Theory Techn.* **2017**, *66*, 831–844. [CrossRef]
15. Tervo, N.; Jokinen, M.; Leinonen, M.E.; Aikio, J.; Kursu, O.; Rahkonen, T.; Pärssinen, A. Digital Predistortion Concepts for Linearization of mmW Phased Array Transmitters. In Proceedings of the 16th International Symposium on Wireless Communication Systems (ISWCS), Oulu, Finland, 27–30 August 2019; pp. 325–329.
16. Eun, C.; Powers, E.J. A new Volterra predistorter based on the indirect learning architecture. *IEEE Trans. Signal Process.* **1997**, *45*, 223–227. [CrossRef]
17. Lei Ding.; Zhou, G.T.; Morgan, D.R.; Zhengxiang Ma.; Kenney, J.S.; Kim, J.; Giardina, C.R. A robust digital baseband predistorter constructed using memory polynomials. *IEEE Trans. Commun.* **2004**, *52*, 159–165. [CrossRef]

18. Zhou, D.; DeBrunner, V.E. Novel adaptive nonlinear predistorters based on the direct learning algorithm. *IEEE Trans. Signal Process.* **2006**, *55*, 120–133. [CrossRef]
19. Schetzen, M. Theory of pth-order inverses of nonlinear systems. *IEEE Trans. Circuits Syst.* **1976**, *23*, 285–291. [CrossRef]
20. Paaso, H.; Mammela, A. Comparison of direct learning and indirect learning predistortion architectures. In Proceedings of the IEEE International Symposium on Wireless Communication Systems, Reykjavik, Iceland, 21–24 October 2008; pp. 309–313. [CrossRef]
21. Magesacher, T.; Singerl, P. Benchmarking of learning architectures for digital predistortion. In Proceedings of the 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; pp. 648–651. [CrossRef]
22. Chani-Cahuana, J.; Landin, P.N.; Fager, C.; Eriksson, T. Iterative learning control for RF power amplifier linearization. *IEEE Trans. Microw. Theory Techn.* **2016**, *64*, 2778–2789. [CrossRef]
23. Schoukens, M.; Hammenecker, J.; Cooman, A. Obtaining the preinverse of a power amplifier using iterative learning control. *IEEE Trans. Microw. Theory Techn.* **2017**, *65*, 4266–4273. [CrossRef]
24. Freiberger, K.; Wolkerstorfer, M.; Enzinger, H.; Vogel, C. Digital predistorter identification based on constrained multi-objective optimization of WLAN standard performance metrics. In Proceedings of the 2015 IEEE International Symposium on Circuits and Systems, Lisbon, Portugal, 24–27 May 2015; pp. 862–865.
25. Li, L.; Ghazi, A.; Boutellier, J.; Anttila, L.; Valkama, M.; Bhattacharyya, S.S. Design space exploration and constrained multiobjective optimization for digital predistortion systems. In Proceedings of the 2016 IEEE 27th International Conference on Application-specific Systems, Architectures and Processors (ASAP), London, UK, 6–8 July 2016, pp. 182–185.
26. Ma, R.; Benosman, M.; Manjunatha, K.A.; Komatsuzaki, Y.; Shinjo, S.; Teo, K.H.; Orlik, P.V. Machine-Learning Based Digital Doherty Power Amplifier. In Proceedings of the IEEE International Symposium on radio–frequency Integration Technology (RFIT), Melbourne, Australia, 15–17 August 2018; pp. 1–3. [CrossRef]
27. Kaisa, M. *Nonlinear Multiobjective Optimization*; International Series in Operations Research & Management Science; Kluwer Academic Publishers: Boston, MA, USA, 1999; Volume 12.
28. Sun, W.; Yuan, Y.x. *Optimization Theory and Methods: Nonlinear Programming*; Springer: Berlin/Heidelberg, Germany, 2006.
29. Landin, P.N.; Gustafsson, S.; Fager, C.; Eriksson, T. WebLab: A Web-based setup for PA digital predistortion and characterization [application notes]. *IEEE Microw. Mag.* **2015**, *16*, 138–140. [CrossRef]
30. Mengozzi, M.; Gibiino, G.P.; Angelotti, A.M.; Florian, C.; Santarelli, A. Supply-Modulated PA Performance Enhancement by Joint Optimization of RF Input and Supply Control. In Proceedings of the Asia-Pacific Microwave Conference (APMC), Hong Kong, China, 8–11 December 2020.
31. Barmuta, P.; Ferranti, F.; Gibiino, G.P.; Lewandowski, A.; Schreurs, D.M.P. Compact behavioral models of nonlinear active devices using response surface methodology. *IEEE Trans. Microw. Theory Tech.* **2014**, *63*, 56–64. [CrossRef]