Distortion shaping in space and frequency using digital predistortion in MIMO transmitters

Citation for the original published paper (version of record):
Langborn, B., Wang, S., Eriksson, T. (2026). Distortion shaping in space and frequency using digital predistortion in MIMO transmitters. IEEE Communications Letters, 30: 2770-2774.
http://dx.doi.org/10.1109/LCOMM.2026.3719126

Distortion Shaping in Space and Frequency Using Digital Predistortion in MIMO Transmitters

Björn Langborn, Siqi Wang, *Member, IEEE*, and Thomas Eriksson, *Member, IEEE*

***Abstract*—A novel approach to multiple-input, multiple-output transmitter distortion shaping that shapes emitted distortion at different frequencies, and in different spatial directions, differently is proposed. In this approach, a new criterion of optimality using weighted over-the-air observations, and a digital predistortion (DPD) scheme particularly adapted to this criterion, are formulated. Through simulation, it is shown that emitted in-band and out-of-band distortion can be shaped spatially in the proposed approach, with up to 8 dB in-band distortion reduction in a user direction, compared to conventional per-PA DPD.**

***Index Terms*—Digital predistortion, distortion, power amplifier.**

I. Introduction

POWER amplifiers (PAs) in wireless transmitters are often designed or driven to operate nonlinearly to improve energy efficiency, capacity, and coverage, at the cost of increased signal distortion. To avoid signal degradation for intended users and interference leakage to other spectrum users, this distortion must be properly managed. One mitigation strategy is spatial nullforming of distortion toward selected directions [1], [2], which protects specific users but leaves emissions in other directions uncompensated. Alternatively, distortion can be suppressed through linearization, most commonly via digital predistortion (DPD), which injects compensatory nonlinearities to approximate a desired linear output.

In a conventional single-input, single-output (SISO) setting, the problem of identifying a DPD is well-studied and often relatively straight-forward. The input signal is adjusted until it yields a desired, linear output signal. In a multiple-input, multiple-output (MIMO) transmitter setting, the problem is increased in dimensionality. The simplest approach to tackle this would be to treat each PA individually when linearizing, thus reducing the MIMO problem to a set of SISO optimization problems. However, this does not consider the filtering and recombination of signals when passing through antennae and a wireless channel.

Several MIMO DPD approaches explicitly exploit spatial degrees of freedom. In [3], distortion toward a specific user direction is minimized, whereas [4] combines PA biasing and DPD to achieve sidelobe and omnidirectional linearization, albeit without an explicit optimization criterion. Beam-oriented DPD combined with post-weighting is proposed in [5] and [6], extending linearization over a wider angular region while preserving user-directed power. These works demonstrate effective spatial shaping of distortion, but offer limited control over where residual distortion is emitted and do not allow separate treatment of in-band (IB) and out-of-band (OOB) distortion.

Frequency-domain (FD) DPD, as in [8], has been proposed to achieve frequency selective linearization. Other works, e.g. with a time-domain approach and bandpass filtering [9], or a Pareto-optimal multimetric DPD approach [10], have also studied this problem. However, the approach of [8] lends itself very flexibly to allowing different sets of DPD coefficients to be applied on different parts of the frequency spectrum. This, in conjunction with the idea of spatial distortion shaping, lays the groundwork of this work.

II. Motivation and Contribution

Spatial and frequency distortion shaping enables the use of antenna, channel, and scenario-specific information to control how distortion is radiated. Spatial shaping allows distortion to be reduced in selected directions at the expense of increased distortion elsewhere, while frequency-selective DPD enables separate handling of IB and OOB emissions across spatial directions. This is particularly relevant in MIMO systems, where additional degrees of freedom can be exploited.

Improving the signal-to-noise-and-distortion ratio (SNDR) in a user direction enables higher-order modulation formats and increased data rates, whereas any IB power radiated in non-user directions is generally undesirable. Consequently, increasing IB distortion in non-user directions, while keeping the total IB power constant, does not degrade performance. This because non-users do not distinguish between clean and distorted interference. Furthermore, permitting higher distortion levels in non-user directions may enable reduced DPD complexity or improved performance in user directions. Alternatively, when a user employs a low-order modulation format with relaxed SNDR requirements, it can be advantageous to tolerate increased IB distortion in exchange for reduced OOB emissions.

Following this motivation, this work contributes as follows:

- An optimization criterion that incorporates spatial and frequency weighting of radiated distortion, enabling spatial control of IB and OOB distortion separately, is proposed.
- Details on how FD-DPD identification, particularly suited for the proposed criterion, can be performed.

Received 2 May 2026; revised 20 June 2026; accepted 25 July 2026. Date of publication 3 August 2026; date of current version 6 August 2026. This work was supported by Vinnova, Formas and the Swedish Energy Agency within the strategic innovation programme ‘Smartare Elektroniksystem’, and by Chalmers University of Technology and Ericsson AB. The associate editor coordinating the review of this letter and approving it for publication was S. Naser. *(Corresponding author: Björn Langborn.)*

Björn Langborn and Thomas Eriksson are with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: langborn@chalmers.se; thomase@chalmers.se).

Siqi Wang is with CNRS, Laboratoire de Genie Electrique et Electronique de Paris, Sorbonne Universite, 75252 Paris, France (e-mail: wangsiqi0807@gmail.com).

Digital Object Identifier 10.1109/LCOMM.2026.3719126

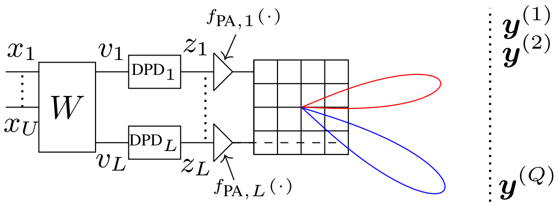


Fig. 1. Transmitter with U user data streams precoded through $\boldsymbol{W}$ to L transmit branches with DPDs, PAs and propagated through antennae OTA to Q observation locations.

- Simulation results to show the feasibility of this approach.

III. System Modeling

Consider a fully-digital beamforming transmitter serving U users with data streams $x_1, \ldots, x_U$, precoded by $\boldsymbol{W}$ to L transmit branches, each equipped with a DPD and a PA, as shown in Fig. 1. Denote the precoded signals by $v_1, \ldots, v_L$ and the DPD outputs by $z_1, \ldots, z_L$. A received baseband signal sample in the qth observation direction, at time index n, is modeled as

$$y^{(q)}[n] = (\boldsymbol{h}^{(q)})^T f_{\text{PA}}(\boldsymbol{z}[n])\,, \tag{1}$$

where it has propagated through the channel $\boldsymbol{h}^{(q)}$ as

$$\boldsymbol{h}^{(q)} = \left[h_1^{(q)} \ldots h_L^{(q)}\right]^T , \tag{2a}$$

$$f_{\text{PA}}(\boldsymbol{z}[n]) = \left[f_{\text{PA},1}(z_1[n]) \ldots f_{\text{PA},L}(z_L[n])\right]^T . \tag{2b}$$

The PA output $y_{\text{PA},l}[n] = f_{\text{PA},l}(z_l[n])$ is given by passing a predistorted signal $z_l[n]$ through some nonlinear PA function $f_{\text{PA},l}(\cdot)$. In this work, as defined later in (16), a static PA model is used simulations, for clarity. However, the theory in this work also applies to models with memory, without any loss of generality. Moreover, $\boldsymbol{H}$ is here considered as a line-of-sight (LOS) channel. Further, to consider multiple directions simultaneously, over-the-air (OTA) observations

$$\boldsymbol{y}[n] = \boldsymbol{H} f_{\text{PA}}(\boldsymbol{z}[n])\,, \tag{3}$$

can be stacked into a vector, where

$$\boldsymbol{y}[n] = \left[y^{(1)}[n] \ldots y^{(Q)}[n]\right]^T , \boldsymbol{H} = \begin{bmatrix} (\boldsymbol{h}^{(1)})^T \\ \vdots \\ (\boldsymbol{h}^{(Q)})^T \end{bmatrix} . \tag{4}$$

Additionally, let each predistorted signal be generated by a DPD modelled as

$$z_l[n] = f_{\text{DPD},l}(v_l[n]) = \boldsymbol{f}_l[n]\boldsymbol{\theta}_l\,, \tag{5}$$

where $\boldsymbol{f}_l[n]$ are DPD bases for the sample $v_l[n]$ and $\boldsymbol{\theta}_l$ modeling coefficients. The formulation is easily extended to consider multiple baseband samples in (1)-(5) by stacking samples column-wise, notationally changing $z_l[n] \rightarrow \boldsymbol{z}_l$, $y^{(q)}[n] \rightarrow \boldsymbol{y}^{(q)}$, $v_l[n] \rightarrow \boldsymbol{v}_l$ and $\boldsymbol{f}_l[n] \rightarrow \boldsymbol{F}_l$. This is necessary for identification, and extends the DPD bases from $\boldsymbol{f}_l$ with dimension $1 \times B$ to $\boldsymbol{F}_l$ with dimension $K \times B$, where B denotes the number of DPD bases and K the number of baseband samples.

IV. Spatial and Frequency Weighted DPD

DPD parameter identification requires an appropriate optimality criterion. This section first reviews the conventional per-PA SISO criterion and then introduces the proposed spatially and frequency weighted (SFW) OTA criterion.

A. Conventional SISO DPD

Conventionally, each transmit branch in Fig. 1 is linearized independently by identifying DPD coefficients $\boldsymbol{\theta}_l$ that solve

$$\arg\min_{\boldsymbol{\theta}_l} ||y_{\text{PA,d},l} - f_{\text{PA},l}(\boldsymbol{F}_l\boldsymbol{\theta}_l)||^2 . \tag{6}$$

The desired PA output is typically chosen as a linear scaling of the precoded input, $y_{\text{PA,d},l} = gv_l$. This approach, denoted here as per-PA DPD (PP-DPD), yields a unique optimal solution for each amplifier but neglects antenna-array and channel effects, as the optimization is performed separately for each branch.

B. Criteria for DPD Optimization

When DPDs are optimized jointly based on OTA observations, a natural formulation is

$$\arg\min_{\boldsymbol{\theta}} ||\boldsymbol{y}_{\text{d}} - \boldsymbol{H} f_{\text{PA}}(\boldsymbol{F}\boldsymbol{\theta})||^2 , \tag{7}$$

where $\boldsymbol{y}_{\text{d}}$ denotes the desired signals in Q spatial directions, of the same form as (3), and

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{F}_L \end{bmatrix} , \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \vdots \\ \boldsymbol{\theta}_L \end{bmatrix} . \tag{8}$$

In contrast to (6), (7) optimizes the linear OTA combination of PA outputs and inherently incorporates spatial weighting via the channel matrix $\boldsymbol{H}$. As a result, distortion in directions with low channel gain has limited impact on the optimal solution. This can be further exploited by applying different weights in different spatial directions, corresponding to where an operator can tolerate distortion.

To enable explicit spatial and frequency control, consider K processed baseband samples per user stream and subsequently observed OTA. Define $\hat{\boldsymbol{F}}$ and $\hat{\boldsymbol{y}}_{\text{d}}$ as the frequency-domain representations of $\boldsymbol{F}$ and $\boldsymbol{y}_{\text{d}}$, respectively. Introducing spatial–frequency weights $w_{q,k}$, the optimization becomes

$$\arg\min_{\boldsymbol{\Theta}} \sum_{q=1}^{Q}\sum_{k=1}^{K} \boldsymbol{w}_{q,k} \left|\left|\hat{\boldsymbol{y}}_{\text{d}}^{(q)}[k] - \boldsymbol{h}^{(q)} f_{\text{PA}}(\hat{\boldsymbol{F}}[k]\boldsymbol{\Theta}^{[k]})\right|\right|^2 . \tag{9}$$

Here, $\boldsymbol{\Theta} = \left[\boldsymbol{\Theta}^{[1]} \ldots \boldsymbol{\Theta}^{[K]}\right]$ is a set of DPD coefficients to be applied at different frequencies k, where each $\boldsymbol{\Theta}^{[k]}$ is a vector as in (8). In contrast to (6), (9) can be underdetermined and yield nonunique solutions, unless $|w_{q,k}| > 0$ holds in at least L spatial directions. This would result in a spatial null space, directions in which distortion does not affect the optimality of the solution, that must be treated with care as to not violate spectrum regulations. With $w_{q,k} = 1$ everywhere, and $\boldsymbol{\Theta}^{[k]} = \boldsymbol{\theta}\,, \forall k$, (9) is equivalent to (7).

Notably, (9) allows distortion to be shaped in both frequency and space by giving relative weight to where the distortion is most harmful to the system. Whilst (6) minimizes the

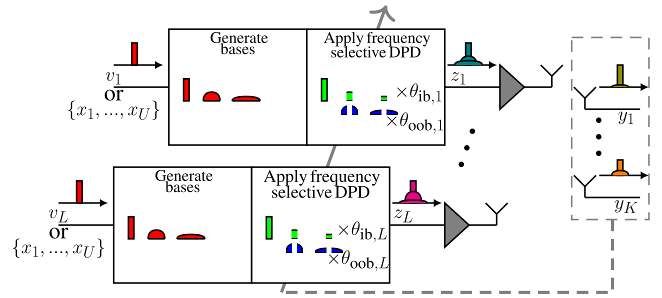


Fig. 2. Spatially and frequency weighted DPD with two frequency regions, in-band (ib) and out-of-band (oob).

distortion at each PA output, (9) allows for altering the DPDs to a solution better in some spatial and frequency regions, at the cost of degradation in other regions. This allows for flexibility in adhering to emission requirements, whilst obtaining improved performance in user directions.

Conventional DPD identification, i.e. solving (6), is a well-studied and understood problem, requiring no further details here. However, (9) being a novel criterion, the following section will detail approaches to identification.

V. Identification of the Spatially and Frequency Weighted DPD

Solving (9) for a unique set of DPD coefficients at every frequency bin is generally unnecessary. Instead, a small number of frequency regions $\boldsymbol{\kappa}_1, \ldots, \boldsymbol{\kappa}_S$ can be defined, such that

$$\boldsymbol{\Theta}^{[k]} = \boldsymbol{\theta}^{[i]} \quad \text{if } k \in \boldsymbol{\kappa}_i \, , i = 1, \ldots, S \, . \tag{10}$$

The case $S = 2$ can correspond to IB and OOB separation, if the sets $\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2$ are chosen correspondingly.

As in conventional DPD, (9) may be solved using direct or indirect learning, typically in an iterative manner. Independent of the learning approach, the weighted OTA error at iteration $[i]$ can be formed in the frequency domain as

$$\hat{e}_{q,k}^{[i]} = \boldsymbol{w}_{q,k} \left|\left| \hat{\boldsymbol{y}}_{\mathrm{d}}^{(q)}[k] - \boldsymbol{h}^{(q)} f_{\mathrm{PA}}(\hat{\boldsymbol{F}}[k]\boldsymbol{\Theta}^{[k]}) \right|\right|^2 . \tag{11}$$

Then, the errors are combined into a matrix $\hat{\boldsymbol{E}}^{[i]}$ of size $Q \times K$ with elements $\hat{e}_{q,k}^{[i]}$. In this work, direct learning is then used by forming a DPD basis

$$\hat{\boldsymbol{F}} = \begin{bmatrix} \hat{\boldsymbol{F}}_1[\kappa_1] & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \hat{\boldsymbol{F}}_L[\kappa_S] \end{bmatrix} , \tag{12}$$

yielding iterative updates

$$\begin{bmatrix} \boldsymbol{\theta}_{\kappa_1} \\ \vdots \\ \boldsymbol{\theta}_{\kappa_S} \end{bmatrix}^{[i+1]} = \begin{bmatrix} \boldsymbol{\theta}_{\kappa_1} \\ \vdots \\ \boldsymbol{\theta}_{\kappa_S} \end{bmatrix}^{[i]} + \eta \hat{\boldsymbol{F}}^{\dagger} \boldsymbol{H}^{\dagger} \hat{\boldsymbol{E}}^{[i]} . \tag{13}$$

Here, η is a step parameter and $\boldsymbol{H}^{\dagger} = (\boldsymbol{H}^{*}\boldsymbol{H})^{-1}\boldsymbol{H}^{*}$ is the pseudoinverse. The DPD using a basis as in (12), optimized for (9), constitutes a SFW-DPD. This scheme is visualized in Fig. 2, and shows a case with $S = 2$ in relation to (10). The scheme generalizes straightforwardly to larger S, that can be of interest for e.g. multi-band MIMO DPDs, or where different frequency regions have different regulation requirements.

The DPD unit implementation follows the frequency-domain DPD approach of [8]: DPD bases are generated in the time domain, transformed via FFT, and applied selectively over frequency bins corresponding to different nonlinear orders before inverse FFT reconstruction.

Lastly, whilst $\boldsymbol{y}^{(q)} = \boldsymbol{h}^{(q)} f_{\mathrm{PA}}(\hat{\boldsymbol{F}}[k]\boldsymbol{\Theta}^{[k]})$ denotes direct OTA observations, it can be approximately obtained using channel state information and PA estimates. Such PA estimates can be obtained with e.g transmit-receive observation antennae such as in [15] and [17], or using output couplers.

A. Single Beam

For a single user beam, nonlinear distortion prior to linearization is radiated primarily in the user direction. Then, the SFW-DPD objective would most often be to reduce and shape distortion away from that direction. This will result in that DPD coefficients are tapered along the branches, in relation to the conventional PP-DPD, seen later in Fig. 3. Assuming a static PA model, a per-branch DPD basis can be built as

$$\boldsymbol{F}_l = \begin{bmatrix} \boldsymbol{v}_l & \boldsymbol{v}_l|\boldsymbol{v}_l| & \ldots & \boldsymbol{v}_l|\boldsymbol{v}_l|^{P-1} \end{bmatrix} , \boldsymbol{\theta}_l = \begin{bmatrix} \theta_1 & \ldots & \theta_P \end{bmatrix}^T . \tag{14}$$

For more intricate PA model, memory and other nonlinear terms can easily be added to (14).

B. Multiple Beams

With beams to $U > 1$ users, there will be distortion going in off-beam directions, in addition to the user beam directions. An alternative is then to work with a beam-domain basis as in [12] for (12). For a static PA model, all the bases in terms of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_U$, attainable from expanding $\boldsymbol{v} = \boldsymbol{x}_1 + \ldots + \boldsymbol{x}_U$ as $\boldsymbol{z} = \sum_{p=1,3,\ldots} \boldsymbol{v}|\boldsymbol{v}|^{p-1}$ are used. Exemplified, for two users, nonlinearity order $P = 3$, and defining $\boldsymbol{\gamma}_{1,2} = |\boldsymbol{x}_1|^2 + |\boldsymbol{x}_2|^2$, this equates to

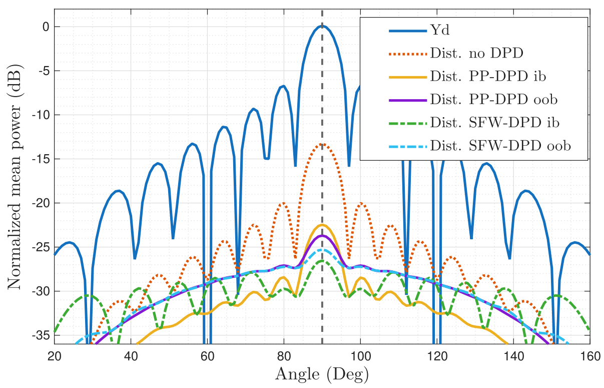
$$\boldsymbol{F}_l = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_1\boldsymbol{\gamma}_{1,2} & \boldsymbol{x}_2\boldsymbol{\gamma}_{1,2} & \boldsymbol{x}_1^2\boldsymbol{x}_2^* & \boldsymbol{x}_2^2\boldsymbol{x}_1^* \end{bmatrix} . \tag{15}$$

This gives greater flexibility in shaping the distortion based on each stream. The additional cost is that of multi-dimensional, rather than SISO, DPDs for such an implementation. The schematics of Fig. 1 and Fig. 2 are then altered such that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_U$ are input directly to the DPDs, with $\boldsymbol{z}_l$ as a single combined multi-user output signal, resulting in multi-input, single-output DPDs.

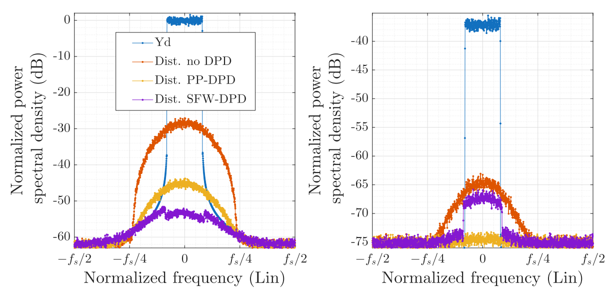
VI. Simulation Setup and Results

For simulation, the channel $\boldsymbol{H}$ is chosen as a LOS channel with antenna array pattern gain varying with angle ϕ as $\sin^3(\phi)$,[1] and with $L = 16$ branches and $Q = 179$ observation directions. No mutual coupling is considered. This assumption holds approximately if cost and size constraints allow the transmitters to be equipped with circulators, or if the coupling is estimated and mitigated through other methods

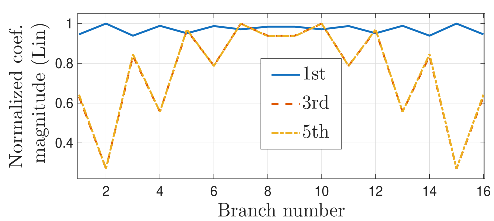
[1] The $\sin^3(\phi)$ function is a modeling choice, loosely based on experimental observations, to exemplify an array radiating worse towards the end-fire direction than towards broadside.



(a) Signal and distortion power, plotted separately for in-band (ib) and out-of-band (oob), in varying spatial directions. The user direction is indicated by the dashed vertical line.



(b) Power spectral density in user and non-user directions, 90° and 62°.



(c) Normalized, relative magnitude of the first, third and fifth order in-band SFW-DPD coefficients. Coefficient tapering along the branches is seen, shaping distortion.

Fig. 3. Normalized power in contrast to desired signal, for no DPD, conventional per-PA (PP) DPD or with spatially and frequency weighted (SFW) DPD, for a single beam scenario.

such as in [16]. The latter approach comes at an additional computational cost in the digital processing. The PA function is implemented as a baseband Saleh model [13], equal on each PA branch, with added gaussian noise $\boldsymbol{n}_l$ at the output as

$$f_{\mathrm{PA},l}(\boldsymbol{z}_l) = \frac{a\boldsymbol{z}_l}{1 + b|\boldsymbol{z}_l|^2} + \boldsymbol{n}_l\,, \tag{16}$$

with $a = 1, b = 1$ and a standard deviation $\boldsymbol{\sigma}_{v_l} \approx 0.2$ of the precoded inputs $\boldsymbol{v}_l$. The desired signal is

$$\boldsymbol{Y}_{\mathrm{d}} = a_d \boldsymbol{H}\boldsymbol{v}\,, \tag{17}$$

with $a_d = 0.88$ chosen as a desired linear gain of the PAs. Baseband data of $N = 65536$ samples is generated as bandlimited Gaussian noise with normalized bandwidth 0.16, allowing spectral regrowth.

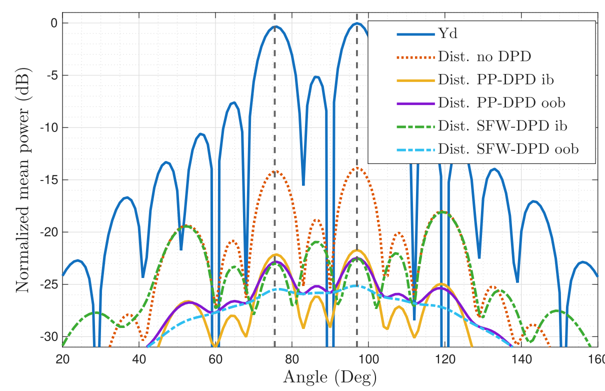
In Fig. 3, a single-beam case with $U = 1$ user is shown. The proposed criterion (9) is optimized using direct DPD learning as in (13), with weights $w_{q,k}$ as listed in Table I. In Fig. 3a and Fig. 3b it can be seen that the error towards the user is decreased up to 8 dB by shaping IB distortion in other spatial directions, for SFW-DPD compared to conventional PP-DPD. Fig. 3c illustrates how this SFW-DPD solution involves tapering of the DPD coefficients to achieve the distortion shaping, similar to how beamformers can amplitude taper coefficients to affect sidelobes and nulls. The weights of Table I are chosen, somewhat arbitrarily, to reflect higher importance of especially in-band user-oriented distortion. Larger weight magnitudes corresponds to that distortion there should be penalized more. The DPDs, modelled based on (14), include odd and even order static polynomial terms up to nonlinearity order $P = 7$.

In Fig. 4, a multi-beam case with $U = 2$ users is explored, with $w_{q,k}$ as in Table II. From Fig. 4a and Fig. 4b it is seen that OOB distortion is reduced everywhere in space for the proposed SFW-DPD compared to PP-DPD, at the expense of increased IB distortion in non-user directions. The total emitted OOB power is reduced by approximately 2.5 dB,

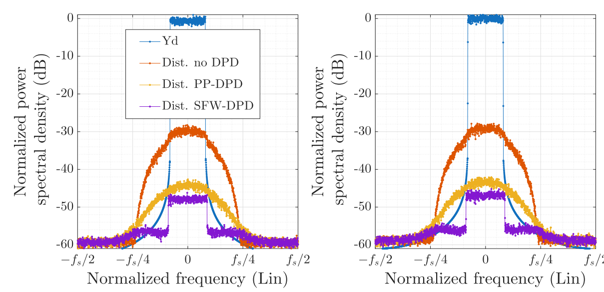
TABLE I

Distortion Weights Corresponding to Fig. 3, a Single-Beam Case Focused on User-Oriented Distortion Reduction

Distortion weights $w_{q,k}$ (Lin)	In-band	Out-of-band
Towards users $\{90 \pm 10\}°$	5	4.25
Towards nonusers	0.01	0.5



(a) Signal and distortion power, plotted separately for in-band (ib) and out-of-band (oob), in varying spatial directions. User directions are indicated by the dashed vertical lines.



(b) Power spectral density in user directions, 76° and 97°.

Fig. 4. Normalized power in contrast to desired signal, for no DPD, conventional per-PA (PP) DPD or with spatially and frequency weighted (SFW) DPD, for a dual-beam scenario.

TABLE II

Distortion Weights Corresponding to Fig. 4, a Dual-Beam Case Focused on Out-of-Band Distortion Reduction Everywhere

Distortion weights $w_{q,k}$ (Lin)	In-band	Out-of-band
Towards users $\{76 \pm 5, 97 \pm 5\}°$	1	5
Towards nonusers	0.01	5

TABLE III
IN-BAND NORMALIZED MEAN SQUARED ERROR (NMSE) AT USERS, AND OUT-OF-BAND (OOB) POWER EMITTED, FOR THE DIFFERENT SCENARIOS

Simulation scenario	DPD scheme	User-oriented in-band NMSE (dB)	Total OOB power (dBc)
Single-beam, single-user	No DPD	−28.6	−32.2
	PP-DPD	−45.3	−43.8
	SFW-DPD	−53.3	−44.9
Dual-beam, dual-user	No DPD	−29.1/−29.1	−33.2
	PP-DPD	−43.6/−43.4	−42.2
	SFW-DPD	−46.9/−46.6	−44.8

whilst user-directed IB distortion is simultaneously lowered. The weights of Table II are, in contrast to Table I, chosen to reflect higher importance of out-of-band distortion emitted overall. The DPDs, based on (15), include odd and even order static polynomial terms up to nonlinearity order $P = 5$.

Table III summarizes the performance for the two scenarios, using no DPD in contrast to PP-DPD and SFW-DPD.

VII. DISCUSSION

Implementing DPDs per branch is costly in large MIMO transmitters, which has motivated other work in e.g. beam-domain DPD [12]. Whilst the proposed work is not mainly focused on complexity reduction, complexity concerns are worth further comment. Firstly, one way to regulate complexity in the proposed DPD is to only linearize a subset of the frequency indices, showcased in [12]. Secondly, DPDs linearizing different frequency regions can include different number of coefficients, as the strictness of linearization requirements can vary. Lastly, an alternative approach to complexity reduction is to combine the proposed SFW-DPD with other complexity reduction methods such as [14], where the number of DPDs L is reduced through a so-called virtual array DPD.

These combined strategies suggests that the SFW-DPD can be adapted to meet various implementation constraints.

This work was evaluated for a simple line-of-sight channel setting. Extension to more complex channel scenarios is of interest for future work. Then, based on channel sounding and state information, a key challenge is to identify directions in which distortion can be allowed to radiate more or less. Lastly, the distortion weights of $w_{q,k}$ were chosen somewhat arbitrarily for illustrative purposes of the DPD concept. In practice, (9) could be extended to include optimizing the weights themselves. One proposed reformulation is to minimize user oriented in-band distortion, constrained to limit both in-band and out-of-band distortion below certain thresholds across all spatial directions.

VIII. CONCLUSION

Distortion shaping in both frequency and space is shown capable of improving linearization performance in specified frequency and spatial regions at the cost of shaping distortion elsewhere. The proposed error metric allows flexibility in how residual distortion is to be shaped, and shown implementable by the proposed spatial and frequency weighted (SFW) DPD approach. Specifically, simulation results for single-beam and dual-beam scenarios illustrate the ability of the SFW-DPD to reduce user-oriented distortion, selectively suppress out-of-band emissions, and exploit spatial and spectral degrees of freedom to meet diverse emission requirements. Overall, the approach offers a flexible framework for distortion management in digital MIMO transmitters.

REFERENCES

[1] N. Kolomvakis, M. Bavand, I. Bahceci, and U. Gustavsson, “A distortion nullforming precoder in massive MIMO systems with nonlinear hardware,” *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1775–1779, Sep. 2022, doi: 10.1109/LWC.2022.3175147.

[2] S. R. Aghdam, S. Jacobsson, U. Gustavsson, G. Durisi, C. Studer, and T. Eriksson, “Distortion-aware linear precoding for massive MIMO downlink systems with nonlinear power amplifiers,” 2020, *arXiv:2012.13337*.

[3] N. Tervo, J. Aikio, T. Tuovinen, T. Rahkonen, and A. Parssinen, “Digital predistortion of amplitude varying phased array utilising over-the-air combining,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, Honolulu, HI, USA, Jun. 2017, pp. 1165–1168, doi: 10.1109/MWSYM.2017.8058809.

[4] N. Tervo et al., “Combined sidelobe reduction and omnidirectional linearization of phased array by using tapered power amplifier biasing and digital predistortion,” *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 9, pp. 4284–4299, Sep. 2021, doi: 10.1109/TMTT.2021.3092357.

[5] J. Yan, H. Wang, and J. Shen, “Novel post-weighting digital predistortion structures for hybrid beamforming systems,” *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3980–3984, Dec. 2021, doi: 10.1109/LCOMM.2021.3120648.

[6] Z. Zhang, Y. Li, S. Yang, N. Wei, and C. Yuen, “Linearization angle widened predistortion for hybrid beamforming array utilizing iterative post-weighting,” in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, London, U.K., May 2025, pp. 1–5, doi: 10.1109/ISCAS56072.2025.11043215.

[7] 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, doi: 10.1109/TMTT.2018.2830772.

[8] A. Brihuega, L. Anttila, and M. Valkama, “Frequency-domain digital predistortion for OFDM,” *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 6, pp. 816–818, Jun. 2021, doi: 10.1109/LMWC.2021.3062982.

[9] C. Yu, L. Guan, E. Zhu, and A. Zhu, “Band-limited Volterra series-based digital predistortion for wideband RF power amplifiers,” *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 12, pp. 4198–4208, Dec. 2012, doi: 10.1109/TMTT.2012.2222658.

[10] H. Yin and A. Zhu, “Pareto-optimal multimetric model extraction for digital predistortion of RF power amplifiers for error spectrum redistribution,” *IEEE Trans. Microw. Theory Techn.*, vol. 73, no. 4, pp. 2230–2241, Apr. 2025, doi: 10.1109/TMTT.2024.3467250.

[11] 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, doi: 10.1109/TMTT.2016.2588483.

[12] A. Brihuega, L. Anttila, and M. Valkama, “Beam-level frequency-domain digital predistortion for OFDM massive MIMO transmitters,” *IEEE Trans. Microw. Theory Techn.*, vol. 71, no. 4, pp. 1412–1427, Apr. 2023, doi: 10.1109/TMTT.2022.3222320.

[13] A. A. M. Saleh, “Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers,” *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1715–1720, Nov. 1981, doi: 10.1109/TCOM.1981.1094911.

[14] B. Langborn, C. Fager, R. Hou, and T. Eriksson, “Concurrent multi-beam digital predistortion using FFT beamforming and virtual arrays,” *Sensors*, vol. 26, no. 8, p. 2400, Apr. 2026, doi: 10.3390/s26082400.

[15] X. Wang, Y. Li, C. Yu, W. Hong, and A. Zhu, “Digital predistortion of 5G massive MIMO wireless transmitters based on indirect identification of power amplifier behavior with OTA tests,” *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 1, pp. 316–328, Jan. 2020, doi: 10.1109/TMTT.2019.2944828.

[16] S. Hesami, S. R. Aghdam, C. Fager, T. Eriksson, R. Farrell, and J. Dooley, “Intra-array coupling estimation for MIMO transceivers utilizing blind over-the-air measurements,” in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2019, pp. 468–471, doi: 10.1109/MWSYM.2019.8700777.

[17] A. Ben Ayed, Y. Cao, P. Mitran, and S. Boumaiza, “Digital predistortion of millimeter-wave arrays using near-field based transmitter observation receivers,” *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 7, pp. 3713–3723, Jul. 2022, doi: 10.1109/TMTT.2022.3174857.