



## **Concurrent Multi-Beam Digital Predistortion Using FFT Beamforming and Virtual Arrays**

Downloaded from: <https://research.chalmers.se>, 2026-05-05 19:11 UTC

Citation for the original published paper (version of record):

Langborn, B., Fager, C., Hou, R. et al (2026). Concurrent Multi-Beam Digital Predistortion Using FFT Beamforming and Virtual Arrays. *Sensors*, 26(8). <http://dx.doi.org/10.3390/s26082400>

N.B. When citing this work, cite the original published paper.

Article

# Concurrent Multi-Beam Digital Predistortion Using FFT Beamforming and Virtual Arrays

Björn Langborn <sup>1</sup>, Christian Fager <sup>2</sup>, Rui Hou <sup>2,3</sup> and Thomas Eriksson <sup>1,\*</sup>

<sup>1</sup> Department of Electrical Engineering, Chalmers University of Technology, 412 96 Göteborg, Sweden; langborn@chalmers.se

<sup>2</sup> Department of Microtechnology and Nanoscience, Chalmers University of Technology, 412 96 Göteborg, Sweden; christian.fager@chalmers.se (C.F.); rui.hou@ericsson.com (R.H.)

<sup>3</sup> Ericsson AB, 164 83 Stockholm, Sweden

\* Correspondence: thomase@chalmers.se

## Abstract

A digital predistortion (DPD) scheme for concurrent multi-beam transmission in fully digital multiple-input, multiple-output (MIMO) systems, using Fast Fourier Transform (FFT) beamforming and so-called virtual-array processing, is proposed. In a MIMO array with nonlinear power amplifiers (PAs), transmitting multiple beams concurrently yields intermodulation products that end up in both user and non-user directions. In the setting with few users in a large array, the array dimension will typically be much larger than the number of generated intermodulation products. At the same time, linearization per PA is excessively costly for large arrays. This work shows that it is instead possible to linearize the system by producing predistorted user beams, and non-user intermodulation products, through DPD processing in a virtual array of a much smaller dimension than the physical array. Theoretical derivations and simulation examples show how this approach can lead to manyfold reductions in DPD complexity.

**Keywords:** digital predistortion; virtual array; power amplifier; MIMO; multi-beam; beam domain; precoding; FFT beamforming

## 1. Introduction

Multi-antenna transmitters offer the possibility to spatially direct multiple beams (The concept of a beam in this paper will sometimes be used interchangeably with a precoded data stream, generalizing statements from line-of-sight to other channel conditions.) concurrently, which is of interest for many applications. For communication, it allows base stations to serve multiple users in different locations concurrently. For localization and sensing, utilizing multi-beam schemes can improve localization accuracy [1]. To utilize these possibilities fully in future radio systems, the number of transmit branches is expected to increase dramatically. Simultaneously, operating PAs with as high efficiency as possible always comes with a compromise in their linearity. However, generating distortion due to nonlinear operation is problematic for multiple reasons. Firstly, distortion will be emitted in-band in the directions of the users, deteriorating communication link quality. Secondly, distortion will be generated both in-band and out-of-band into non-user directions and can interfere with other systems. As such, nonlinear distortion must be considered to fulfill transmission requirements. To this end, DPD is widely adopted.

Numerous approaches have been proposed for linearizing multi-antenna transmitters, as reviewed in Section 1.1. In this work, we focus on systems employing fully digital



Academic Editor: Guanding Yu

Received: 6 February 2026

Revised: 23 March 2026

Accepted: 8 April 2026

Published: 14 April 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

beamforming. Prior contributions have targeted either distortion cancellation specifically in the user directions [2,3] or full-angle distortion suppression when a single dominant beam is transmitted [4]. Other recent state-of-the-art research has extended these ideas to scenarios with multiple simultaneous beams through so-called beam-domain DPD (BD-DPD) [5,6].

BD-DPD for multi-beam linearization is similar to multi-dimensional DPDs from the concurrent multi-band setting, such as 2D and 3D-DPD [7–10]. These linearization approaches take multiple user data streams as input to a DPD, and combine the streams to generate a set of multi-dimensional basis functions. As such, they suffer the drawback of a drastically increasing number of multi-dimensional basis functions as the number of user streams increases. One alternative approach to this in the multi-band setting is frequency relocation [11], which involves relocating bands to lie closer in frequency, and generating all the necessary basis functions for linearization through a single nonlinearity before inverse frequency relocation. No similar technique to this exists in the multi-beam setting, to the authors' best knowledge.

### 1.1. Previous Work

Handling distortion emitted from multi-input, multi-output (MIMO) transmitters can be done differently depending on system requirements, available resources, and number of users to be served simultaneously. This subsection provides a brief overview of two main categories of prior work: distortion suppression through beamforming strategies and distortion mitigation through linearization.

A first line of work addresses nonlinear distortion by shaping the beamforming pattern such that distortion is nulled in specific directions, as demonstrated in [12]. In [12], a single-user scenario is studied, with reference to [13] that intermodulation will tend towards being radiated isotropically once multiple beams, or many channel taps, are considered. This glosses over the case when, e.g., there are multiple, but few, beams in relation to the array size. Additionally, a nullforming approach can prevent distortion from being transmitted in specific directions, but it will not be able to prevent distortion from being generated at the PA outputs and subsequently transmitted in many directions. For that, distortion compensation is needed.

A second approach to handling nonlinear distortion is thereby through linearization. In a fully digital beamforming MIMO system, the most straightforward approach to linearization is adopting one DPD per PA in the transmitter chain. While effective, this solution becomes costly in hardware, energy consumption, and computational complexity as the number of transmit branches grows. Consequently, several alternative linearization strategies have been proposed.

For a MIMO transmitter with a single beam, so-called full-angle DPD [4] has been proposed. Here, instead of linearizing  $L$  PAs with  $L$  equal-size DPDs, it is proposed to use a common DPD prior to the precoder, and to perform only minor tuning in the branches to consider branch-wise variations. It is shown to perform well, and to be able to effectively cancel distortion from the transmitter. It is worth noting that this approach still requires tuning for all the branches but one, which is used as reference. From their results, it can be seen that the tuning boxes are not insignificant in size, as their common DPD had 21 coefficients whilst the tuning boxes each had 3–10 coefficients. Furthermore, the DPD scheme considers only a single beam, while in practice it is of great interest to transmit multiple beams concurrently.

Several works have addressed multi-beam linearization. In [3], the goal is to cancel inter-user distortion arising from concurrent multi-beam transmission, and the approach is validated for two users. A related technique for three concurrent users is presented in [14]. However, neither of these studies considers distortion radiated in non-user directions.

In [5], BD-DPD was adopted, applying linearization prior to the beamformer and operating on user data streams rather than on per-antenna signals. With this approach, the DPD complexity scales in relation to the number of users,  $K$ , and user stream intermodulation products, rather than the number of antennas,  $L$ . For systems where  $K \ll L$ , this can yield large complexity savings. However, as in 2D- and 3D-DPD for multi-band signals, the number of intermodulation products grows rapidly with  $K$ , leading to multi-dimensional basis functions such as  $x_1 x_2 x_3^*$ . This is significant for practical implementations that rely on lookup tables (LUTs) to reduce arithmetic computations: multi-dimensional basis functions require multi-dimensional LUTs, increasing both memory usage and interpolation/extrapolation complexity, as discussed in [15]. Lastly, BD-DPD assumes equal PAs across branches. In the presence of PA variations, the performance of BD-DPD is expected to deteriorate.

Other recent contributions include neural network-based DPD architectures applied prior to the precoder [16], as well as DPD strategies for hybrid beamforming systems [2]. Although the neural network approach in [16] obtains good performance and is supported by experimental validation, it does not offer new theoretical or architectural insights. Furthermore, it incurs high memory requirements and substantial floating-point computational costs. As for hybrid beamforming architectures, they are highly relevant to study. However, subarray-based approaches suffer reduced beamforming gain, whereas fully connected architectures entail significant complexity due to crosstalk effects that must be incorporated into the DPD unless model-order reductions are applied [2,17].

### 1.2. Key Contributions and Paper Organization

In this paper, a novel approach to DPD processing using a so-called virtual array is proposed for concurrent multi-beam linearization. The core idea is to process the user data streams through a low-dimensional virtual transmit array, and a corresponding virtual channel representation, whose operation reflects that of the actual array. In the physical transmitter, the  $L$  branches with nonlinear PAs generate intermodulation (IM) distortion that is radiated over the air (OTA) in both user and non-user directions. In the virtual array, by contrast,  $L_v$  nonlinear DPD branches seek to generate the IM terms required to cancel the distortion from the physical array—while using far fewer branches, with  $L_v \ll L$ . This concept has similarities to how previous work performed dimensionality reduction for the multi-band setting, utilizing frequency domain sparseness, via frequency relocation [11] or non-uniform downsampling [18]. However, this work performs dimensionality reduction for the multi-beam setting by utilizing so-called beam-domain sparseness.

The specific contributions of this paper are as follows:

- A novel approach to multi-beam DPD for MIMO transmitters performing DPD in a small virtual antenna space is proposed and exemplified. It is shown how such an approach can lead to manifold reductions in complexity.
- Practical considerations in implementing the proposed scheme are accounted for through simulation and discussion, considering multi-user cases, different channels, and lookup table (LUT) realizations.
- Simulation results verify the theoretical findings, show the applicability of the proposed scheme, and demonstrate how the DPD scales in computational complexity in relation to the number of users.

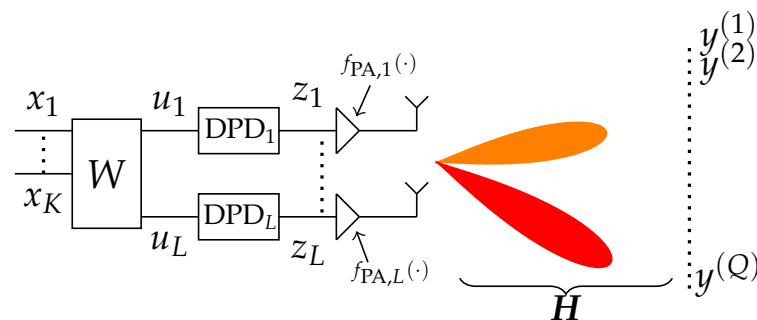
This paper is organized as follows: The assumed system model description is expanded upon in Section 2. Firstly, in Section 2.1, the model description is expanded upon in order to comprehend how distortion is spatially distributed in the beam domain during concurrent multi-beam operation. After that, in Section 2.2, Fast Fourier Transform (FFT) beamforming is described. This is followed in Section 3 by an example where the orthogo-

nality of FFT beamforming highlights the sparsity of the beam domain after over-the-air (OTA) transmission. Section 4 goes on to describe how processing in a virtual array, smaller than the actual physical one, can produce the same linearization signals as per-PA (PP) DPD but at a lower computational cost. Lastly, numerical simulation results are presented in Section 5 to verify the proposed approach, alongside a conclusion in Section 6.

## 2. System Modeling

The system modeling of this work focuses on a discrete time-domain, baseband description of concurrent multi-beam transmission in a fully digital multi-branch transmitter setting, firstly by detailing how intermodulation products arising from nonlinear PAs can be understood in such a description, and then relating them to FFT beamforming to demonstrate the suitability of the proposed approach.

Two key concepts used throughout this work are the beam domain and the antenna space. The beam domain refers to the domain in which signals have already propagated through a channel and formed beams as a result of beamforming. In Figure 1, this is illustrated by the colored beams (red and orange), which are observed over the air in various directions as signals  $y^{(1)}, \dots, y^{(Q)}$ . Some beams carry the user streams  $x_1, \dots, x_K$ , while others contain only IM distortion. In contrast, the antenna space refers to the signals on the different transmit branches, prior to channel propagation. In Figure 1, the antenna space encompasses the signals  $u_1, \dots, u_L, z_1, \dots, z_L$  and also  $f_{PA,1}(z_1), \dots, f_{PA,L}(z_L)$ . The beam domain and antenna space are related through multiplication with a channel or beamforming matrix. For a channel  $\mathbf{H} \in \mathcal{C}^{Q \times L}$ , a signal  $\mathbf{z} \in \mathcal{C}^{L \times N}$  in the antenna space, where  $N$  is the number of baseband data samples, the corresponding beam-domain signal is  $\mathbf{Y} = \mathbf{H}\mathbf{z} \in \mathcal{C}^{Q \times N}$ . This relation is of particular interest when  $Q = L$ , and  $\mathbf{H}$  is an orthogonal matrix, as will be later seen in the subsection on Fast Fourier Transform beamforming. Additionally, whilst this paper initially expands upon the relation between the antenna space and beam domain in the context of line-of-sight (LOS) channels, it is later addressed for non-LOS conditions as well.



**Figure 1.** Concurrent multi-beam transmitter schematic of  $K$  user data streams  $x_1, \dots, x_K$ , beamformed by  $\mathbf{W}$  to constitute signals  $u_1, \dots, u_L$  that are passed through DPDs, transmitted through PAs on  $L$  antenna branches and observed in  $Q$  directions. Signals  $z_1, \dots, z_L$  are predistorted streams,  $f_{PA,l}$  denote a nonlinear PA on branch  $l$ ,  $\mathbf{H}$  represents the wireless channel, and  $y^{(q)}$  refers to the observed signal in direction  $(q)$ .

To limit the scope of the paper, all signals  $x_1, \dots, x_K$  listed in this work are considered to be single-band signals. This can however be generalized to multi-band signals as well, with appropriate modification to the involved DPD units.

As a final comment, since this work considers fully digital beamforming, the numbers of antennas, transmit branches, and PAs are all equal and denoted by  $L$ .

### 2.1. Concurrent Transmission of Multiple Beams

Consider the schematic setup of Figure 1, but start by letting the DPDs be bypassed such that  $z_l = u_l$ ,  $l = 1, \dots, L$ . In this setup, the beamformer  $W$  is formed to direct user streams  $x_1, \dots, x_K$  to directions  $1, \dots, Q$  in matrix form as

$$W = [w^{(1)} \dots w^{(K)}], \quad (1)$$

with  $\|w^{(k)}\|_2 = 1$ , yielding PA input signals

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_L \end{bmatrix} = [w^{(1)}, \dots, w^{(K)}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix}. \quad (2)$$

Alternatively, the signal  $z_l$  passed to PA  $l$  can be written as

$$z_l = w_l^{(1)} x_1 + \dots + w_l^{(K)} x_K. \quad (3)$$

Now, it is known that Volterra series are well-suited to model a broad range of nonlinear functions, and as such, it is commonly applied to model saturation, soft clipping, and other effects in PAs. For practical application, different finite truncations focused on the lower-order nonlinear terms of the Volterra series are frequently used. This can be justified by the fact that lower-order nonlinear terms typically dominate the induced distortion. For qualitative analysis here, consider simply a third-order polynomial baseband PA model as

$$f_{PA,l}(z_l) = c_{1,l} z_l + c_{3,l} z_l |z_l|^2. \quad (4)$$

where  $f_{PA,l}$  refers to the nonlinear function of the PA on branch  $l$ , and  $c_{1,l}, c_{3,l}$  are modeling coefficients. Then, inserting (3) into (4) yields

$$\begin{aligned} f_{PA,l}(z_l) &= c_{1,l} (w_l^{(1)} x_1 + \dots + w_l^{(K)} x_K) + \\ &\quad c_{3,l} (w_l^{(1)} x_1 + \dots + w_l^{(K)} x_K) |w_l^{(1)} x_1 + \dots + w_l^{(K)} x_K|^2 \\ &= c_{1,l} (w_l^{(1)} x_1 + \dots + w_l^{(K)} x_K) + \\ &\quad c_{3,l} (w_l^{(1)} w_l^{(1)} w_l^{(1)*} x_1 |x_1|^2 + \\ &\quad w_l^{(1)} w_l^{(2)} w_l^{(2)*} x_1 |x_2|^2 + \\ &\quad w_l^{(1)} w_l^{(1)} w_l^{(2)*} x_1^2 x_2^* + \dots). \end{aligned} \quad (5)$$

First, assume that all PAs are equal, such that  $c_{1,l}$  and  $c_{3,l}$  are the same for each branch  $l$ . Then, let each PA output propagate through a channel vector  $h_l = [h_l^{(1)} \dots h_l^{(Q)}]$  from branch  $l$  to OTA location  $q$ , as

$$y^{(q)} = [h_1^{(q)} \dots h_L^{(q)}] [f_{PA,1}(z_1) \dots f_{PA,L}(z_L)]^T. \quad (6)$$

Consider the first-order linear terms

$$\begin{aligned} y_{1st}^{(q)} &= h_1^{(q)} c_1 (w_1^{(1)} x_1 + \dots + w_1^{(K)} x_K) + \dots \\ &\quad + h_L^{(q)} c_1 (w_L^{(1)} x_1 + \dots + w_L^{(K)} x_K). \end{aligned} \quad (7)$$

Then it is known from beamforming theory that if  $x_1$  is to go in direction  $(q)$ , then beamforming weights can be chosen such that  $w_l^{(1)} = (h_l^{(q)})^*$ . This yields constructive interference

$c_1 \sum h_l^{(q)} w_l^{(1)} = c_1$ . Then, by studying the third-order terms in (5), it can be seen that since  $w_l^{(1)} w_l^{(1)*}, w_l^{(2)} w_l^{(2)*}$  will be real constants,  $x_1 |x_1|^2$  and  $x_1 |x_2|^2$  will also go in direction  $(q)$ . However, the summation of the IM products last listed in (5),  $c_3 \sum_l h_l^{(q)} w_l^{(1)} w_l^{(1)} w_l^{(2)*}$ , typically has constructive interference in a direction other than  $(q)$ . Generally, it can be seen that the terms of the form  $w^{(i)} w^{(j)} w^{(k)*}$ , where  $k \neq i, j$ , will give rise to new beam directions of constructive interference, which are referred to as off-beams.

Two interesting observations can be made from (5), still assuming  $c_{3,l} = c_3 \forall l$ . Firstly, the generated off-beam terms and directions can be deterministically computed from the intended transmit directions. To provide an example, off-beam terms for a static nonlinearity of order  $P = 3$  and for  $K = 2, 3, 4$  concurrent users are listed in Table 1. Secondly, the nonlinear products of the input streams are of a different form in the off-beams than for the in-beam. Thus, one needs to be aware of all beam directions and manage to generate the correct IM terms in the respective directions in order to linearize the system. This can readily be understood to grow more complex with more beams.

Simplified assumptions have been made in this subsection for qualitative analysis. The simple beamformer in (1) will later be further elaborated on in Section 4.3. Furthermore, the validity of the proposed method when PAs vary on different branches is investigated by simulation in Section 5.5, addressing the assumption of equal PAs. Lastly, it is well-known that PAs will often exhibit higher-order nonlinear behavior than order 3, as in (4), and that memory effects will be present unless the signal bandwidth is very narrow. Section 4.4 comments on introducing higher-order terms in the scheme, and whilst the analysis here was done for static PAs, memory can be added without any loss of generality.

**Table 1.** Off-beam terms for static nonlinearity of order  $P = 3$ , and with  $K = 2, 3, 4$  concurrent users.

| Number of Users | Off-Beam Terms  | Number of Off-Beam Terms |
|-----------------|---|--------------------------|
| $K = 2$         | $x_1 x_1 x_2^*, x_2 x_2 x_1^*$  | 2                        |
| $K = 3$         | $x_1 x_1 x_2^*, x_2 x_2 x_1^*, x_1 x_1 x_3^*, x_3 x_3 x_1^*$<br>$x_2 x_2 x_3^*, x_3 x_3 x_2^*, x_1 x_2 x_3^*, x_1 x_3 x_2^*, x_2 x_3 x_1^*$   | 9                        |
| $K = 4$         | $x_1 x_1 x_2^*, x_2 x_2 x_1^*, x_1 x_1 x_3^*, x_3 x_3 x_1^*$<br>$x_2 x_2 x_3^*, x_3 x_3 x_2^*, x_1 x_2 x_3^*, x_1 x_3 x_2^*,$<br>$x_2 x_3 x_1^*, x_1^2 x_4^*, x_4^2 x_1^*, x_2^2 x_4^*, x_4^2 x_2^*, x_3^2 x_4^*,$<br>$x_4^2 x_3^*, x_1 x_2 x_4^*, x_1 x_4 x_2^*, x_2 x_4 x_1^*, x_1 x_3 x_4^*,$<br>$x_1 x_4 x_3^*, x_3 x_4 x_1^*, x_2 x_3 x_4^*, x_2 x_4 x_3^*, x_3 x_4 x_2^*$ | 24                       |

## 2.2. Fast Fourier Transform Beamforming

FFT beamforming is one common way to efficiently limit the beamforming matrix to an orthogonal set of beam directions, by first defining

$$\omega = e^{-j2\pi/L}. \quad (8)$$

Then, a one-dimensional discrete Fourier transform (DFT) beamformer matrix can be formulated as

$$\mathbf{W}_{\text{DFT}} = \frac{1}{\sqrt{L}} \begin{bmatrix} 1 & 1 & 1 & \dots \\ 1 & \omega & \omega^2 & \dots \\ 1 & \omega^2 & \omega^4 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \frac{1}{\sqrt{L}} [\omega^{(b_1)} \ \omega^{(b_2)} \ \omega^{(b_3)} \ \dots], \quad (9)$$

where  $b_i \in [1, L]$  are beam indices indexing the beams on the DFT beamgrid. In implementation though, no matrix multiplication of  $\mathbf{W}_{\text{DFT}}$  with data is needed, as would otherwise be indicated by (2). Instead, FFT algorithms can perform this computation at lower complexity.

It is worth noting that when the main beams are limited to a DFT beamgrid,  $b_i \in [1, L]$ , the off-beam directions will be restricted to the same beamgrid as well, given equal PAs. Furthermore, extension to two-dimensional beamforming is straightforward, by implementing FFT along both horizontal and vertical directions, but it is omitted from the derivations for brevity. All the analysis following in this paper can then be applied to each dimension, horizontal or vertical, separately.

With the system model established, the following section details how per-PA (PP) DPD produces predistorted streams and how intermodulation beam placement is affected by aliasing due to the antenna array being finite and discrete.

### 3. Per-PA DPD

This section details the processing steps for transmitting  $K$  user streams using per-PA (PP) DPD, as illustrated in Figure 1. A three-user example is used to demonstrate two important aspects: (i) the inherent sparsity of the resulting beam-domain representation for large  $L$  and (ii) an effect referred to as array aliasing, which appears due to the finite and discrete nature of the antenna array.

#### 3.1. Per-PA DPD Example for Three Concurrent Users

Let the antenna array be a linear, one-dimensional array in space, and restrict the beamformer to a DFT beamgrid. As such, this subsection follows the notation of Figure 1, in which  $x_1, \dots, x_K$  represent user streams and  $L$  represents the number of transmit branches, whilst the beamformer  $\mathbf{W}$  is represented by  $\mathbf{W}_{\text{DFT}}$  from (9).

Consider that there are  $L = 16$  transmit branches (While  $L = 16$  is chosen here for clarity of exposition, practical systems would typically employ significantly larger arrays), transmitting  $x_1, x_2, x_3$  to  $K = 3$  users in different LOS directions. Being restricted to a DFT beamgrid, the users can be located at beam indices  $\{b_1, b_2, b_3\} \in [1, L]$ . To exemplify, let  $b_1 = 1, b_2 = 2, b_3 = 5$  such that

$$\begin{aligned} \mathbf{z} = \mathbf{W}_{\text{DFT}}\mathbf{x} &= \frac{1}{\sqrt{L}} \begin{bmatrix} 1 & 1 & 1 & \dots \\ 1 & \omega & \omega^2 & \dots \\ 1 & \omega^2 & \omega^4 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ 0 \\ x_3 \\ 0 \\ \vdots \end{bmatrix} \\ &= \frac{1}{\sqrt{L}} \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + \omega x_2 + \omega^4 x_3 \\ x_1 + \omega^2 x_2 + \omega^8 x_3 \\ \vdots \end{bmatrix} \\ &= \omega^{(b_1)} x_1 + \omega^{(b_2)} x_2 + \omega^{(b_3)} x_3. \end{aligned} \quad (10)$$

Assume for analytical brevity that each DPD is represented by the same nonlinear function  $f(\cdot)$ . Applying the nonlinear DPDs to each of these beamformed data streams, with a third-order nonlinearity  $f(x) = \theta_1 x + \theta_3 x|x|^2$  as the DPD function, then

$$\begin{aligned} f(\mathbf{W}_{\text{DFT}}\mathbf{x}) &= \\ \frac{1}{\sqrt{L}} &\underbrace{\begin{bmatrix} \theta_1(x_1 + x_2 + x_3) + \frac{1}{L}\theta_3(x_1 + x_2 + x_3)|x_1 + x_2 + x_3|^2 \\ \theta_1(x_1 + \omega x_2 + \omega^4 x_3) + \frac{1}{L}\theta_3(x_1 + \omega x_2 + \omega^4 x_3)|x_1 + \omega x_2 + \omega^4 x_3|^2 \\ \theta_1(x_1 + \omega^2 x_2 + \omega^8 x_3) + \frac{1}{L}\theta_3(x_1 + \omega^2 x_2 + \omega^8 x_3)|x_1 + \omega^2 x_2 + \omega^8 x_3|^2 \\ \vdots \end{bmatrix}}_{L \times 1}, \end{aligned} \quad (11)$$

is obtained. How these signals will be spatially distributed OTA can be seen by transformation to the beam domain, achieved through an inverse Fourier transform. The user streams will then be clearly separated in the beam domain, like the beams in Figure 1. With the

matrix  $W_{\text{IDFT}} = LW_{\text{DFT}}^*$ , where  $*$  denotes the Hermitian transpose, factors  $L$  outside the matrix cancel out, resulting in data streams

$$LW_{\text{DFT}}^* f(W_{\text{DFT}} \mathbf{x}) = \begin{bmatrix} \theta_1 x_1 + \frac{1}{L} \theta_3 x_1 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \theta_1 x_2 + \frac{1}{L} \theta_3 x_2 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \frac{1}{L} \theta_3 x_2^2 x_1^* \\ \frac{1}{L} \theta_3 x_1 x_3 x_2^* \\ \theta_1 x_3 + \frac{1}{L} \theta_3 x_3 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \frac{1}{L} \theta_3 x_2 x_3 x_1^* \\ 0 \\ \frac{1}{L} \theta_3 x_3^2 x_2^* \\ \frac{1}{L} \theta_3 x_3^2 x_1^* \\ 0 \\ 0 \\ \frac{1}{L} \theta_3 x_1^2 x_3^* \\ \frac{1}{L} \theta_3 x_1 x_2 x_3^* \\ \frac{1}{L} \theta_3 x_2^2 x_3^* \\ \frac{1}{L} \theta_3 x_1^2 x_2^* \end{bmatrix}. \quad (12)$$

Each entry at a separate index here represents a signal going in a unique beam direction. Two observations are worth noting. The first observation is that only 12 out of the 16 entries are non-zero. This is unsurprising, but important, as Table 1 shows that there can only be up to nine off-beam terms going in potentially different directions, in addition to the  $K = 3$  user directions, for a third-order nonlinearity. This is also true when  $L$  is a very large value. Thus, for large  $L$ , there are very few low-order signal components in relation to the array size. As for the second observation, concerning where each signal component ends up on the beamgrid, it is detailed in the following subsection.

### 3.2. Array Aliasing

In (12), each intermodulation (IM) term maps to a beam index determined by adding or subtracting the corresponding precoder beam indices. For example, the intermodulation term  $(\omega^{(b_2)} x_2)^2 (\omega^{(b_1)} x_1)^* = (\omega^{(b_2)} \omega^{(b_2)} \omega^{(b_1)*}) x_2^2 x_1^*$  indeed produces the intermodulation term  $x_2^2 x_1^*$  up on beam index  $b_2 + b_2 - b_1 = 2 + 2 - 1 = 3$ . However, this simple index addition does not always yield a valid beam index in the range  $[1, L]$ . For instance, the IM product  $(\omega^{(b_1)} x_1)^2 (\omega^{(b_2)} x_2)^*$  corresponds to the index  $2b_1 - b_2$ , which would fall outside  $[1, L]$ . Instead, this term is aliased back in to the beamgrid range  $[1, L]$  according to

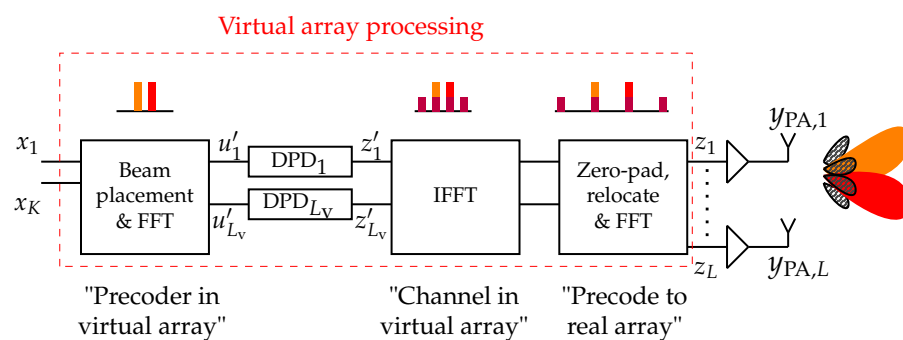
$$\overline{b_{i_1} + b_{i_2} - b_{i_3}} = b_{i_1} + b_{i_2} - b_{i_3} - L \left\lfloor \frac{b_{i_1} + b_{i_2} - b_{i_3} - 1/2}{L} \right\rfloor, \quad (13)$$

for  $\{b_{i_1}, b_{i_2}, b_{i_3}\} \in [1, L]$ ,  $\{i_1, i_2, i_3\} \in [1, \dots, K]$ , and where the bar notation represents the aliasing operation. With this in mind, it can be noted that whilst the example of Section 3.1 produced 12 uniquely placed data streams of main beam and intermodulation beam data, aliasing can, for a different beam configuration  $\{b_1, b_2, b_3\}$ , end up overlapping intermodulation terms to go in the same beam directions. This will lead to distortion being concentrated in fewer beam directions.

## 4. Virtual-Array DPD

In this section, virtual-array DPD (VA-DPD) processing, as shown in Figure 2, is proposed to reduce the complexity compared to applying a DPD per PA. As previously described in Section 1.2, the VA-DPD aims to produce predistorted user streams and off-beam IM terms in a low-dimensional virtual array, and use these to linearize the larger, real transmit array. This processing conceptually follows four main steps. First, user streams are

precoded to the antenna space of a virtual array, by placement on a beamgrid and through an FFT. The beam placement in the virtual array depends on the precoding to be done in the real array, which will be detailed later. Second, nonlinear DPDs are applied to the data on each of the  $L_v$  virtual antenna space branches. Third, an IFFT is performed, acting as a virtual channel and to separate user streams and off-beam IM terms in the virtual beam space. Fourth, and last, the predistorted streams that have been generated are precoded to the real array. The mapping in the virtual domain is chosen such that the predistorted terms generated by the virtual array match those needed to cancel the distortion produced in the physical array. In the following subsections, the derivations of Section 3.1 will be used as a starting point to give mathematical motivation as to how a smaller virtual array, in place of the original array, can be used for performing DPD. This is presented first with an example comparison to the per-PA case, followed by presenting special beam configuration cases in which the complexity can be reduced even further. Finally, generalizations to more general channels and higher-order modeling are discussed.



**Figure 2.** Schematic of DPD using virtual-array processing. The bar plots exemplify the beam space representation of beamforming the data streams to beams and induced IM beams. Red and orange correspond to desired user signals, while purple corresponds to the injected distortion by the DPDs to cancel out the unwanted distortion, shown in striped dark gray. Here, the number of virtual branches  $L_v$  can always be made less than or equal to the number of antenna branches  $L$ .

#### 4.1. Virtual-Array DPD Example for Three Concurrent Users

Revisiting the example of Section 3.1, consider again the beam indices  $\{b_1, b_2, b_3\} = \{1, 2, 5\}$ , but now replace the physical array size  $L = 16$  with a virtual-array size of  $L_v = 13$ . The processing steps then follow directly from the VA-DPD structure of Figure 2: precoding into the virtual antenna space, nonlinear DPD processing, and transformation into the virtual beam domain via an IFFT. This yields

$$L_v \mathbf{W}_{\text{DFT}}^* f(\mathbf{W}_{\text{DFT}} \mathbf{x}) = \begin{bmatrix} \theta_1 x_1 + \frac{1}{L_v} \theta_3 x_1 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \theta_1 x_2 + \frac{1}{L_v} \theta_3 x_2 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \frac{1}{L_v} \theta_3 x_2^2 x_1^* \\ \frac{1}{L_v} \theta_3 x_1 x_3 x_2^* \\ \theta_1 x_3 + \frac{1}{L_v} \theta_3 x_3 (|x_1|^2 + |x_2|^2 + |x_3|^2) \\ \frac{1}{L_v} \theta_3 x_2 x_3 x_1^* \\ 0 \\ \frac{1}{L_v} \theta_3 x_3^2 x_2^* \\ \frac{1}{L_v} \theta_3 x_3^2 x_1^* \\ \frac{1}{L_v} \theta_3 x_1^2 x_3^* \\ \frac{1}{L_v} \theta_3 x_1 x_2 x_3^* \\ \frac{1}{L_v} \theta_3 x_2^2 x_3^* \\ \frac{1}{L_v} \theta_3 x_1^2 x_2^* \end{bmatrix}. \quad (14)$$

instead of (12). Since  $L_v < L$ , fewer DPD computations have been performed whilst generating the same predistorted data streams as in (12). And since these data streams are

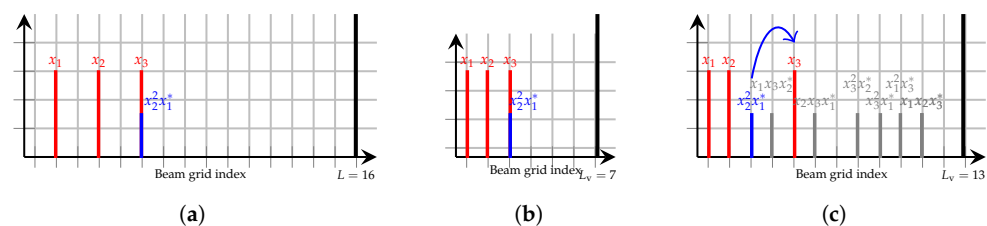
separated, they can then easily be combined or precoded to an arbitrarily large physical array of size  $L \geq L_v$  by merely adding zeros and rearranging the terms.

By combinatorial trials, i.e., computing (13) for all possible combinations of  $\{b_1, b_2, b_3\}$ ,  $L_v = 13$  is found to be the smallest virtual domain that, for any  $L$ , produces all main and intermodulation streams separated as in (14) for  $K = 3$ ,  $P = 3$ . Similarly,  $L_v = 4$  is the smallest such virtual domain for  $K = 2$  users, and  $L_v = 30$  for  $K = 4$ . However, there are cases when the virtual domain can be made with an even smaller  $L_v$ , without loss of information. This will be explained further in the following subsection.

#### 4.2. Special Beam Configuration Cases

Let  $\{b'_1, b'_2, b'_3\}$  denote the set of main beam indices in the virtual domain, with  $\{b_1, b_2, b_3\}$  representing the corresponding user beam indices in the real array. Depending on the real beam configuration,  $\{b_1, b_2, b_3\}$ , and the physical array size,  $L$ , different choices of  $\{b'_1, b'_2, b'_3\}$  and  $L_v$  may be used to obtain an optimally dimensioned virtual array. More exact details on the problem of finding a virtual subspace are found in Appendix A, and are detailed statistically in the results of Section 5.4. There, it is noted how all sets of  $\{b_1, b_2, b_3\}$  can be represented with varying sets of  $L_v$  and  $\{b'_1, b'_2, b'_3\}$ .

The key phenomenon explaining why different sizes of  $L_v$  can occur is array aliasing, as observed in Section 3.2, which can cause intermodulation terms to overlap. When this happens, the number of distinct IM directions is reduced, and consequently, the virtual domain required to represent them need not be as large. For example, evenly spaced beams in an even length array produce IM terms only at even beam indices. Concretely, say  $L = 16$  and  $\{b_1, b_2, b_3\} = \{2, 4, 6\}$ . Then, by (13), the beam indices for a third-order nonlinearity are  $[2, 4, 6, 8, 10, 14, 16]$ , a total of seven unique beam directions. However, computing (13) with  $L_v = 7$ , it can be noted that with  $\{b'_1, b'_2, b'_3\} = \{1, 2, 3\}$ , the same IM terms will overlap with one another, but at indices  $[1, 2, 3, 4, 5, 6, 7]$ . This is illustrated in Figure 3a–c. In Figure 3a, placement of main beams on indices  $\{b_1, b_2, b_3\} = \{2, 4, 6\}$  is shown, in the real array of size  $L = 16$ . It is highlighted how this particular configuration results in the IM-term  $x_2^2 x_1^*$  overlapping with  $x_3$  in the beam domain. Figure 3b then showcases how the same overlap of  $x_2^2 x_1^*$  and  $x_3$  is obtained in the virtual beam domain, for a virtual array of size  $L_v = 7$ . Should it hold that IM overlap is the same for all beams in the real array and the virtual array, then the virtual array can perfectly represent the nonlinear system.



**Figure 3.** Beam occurrences for  $\{b_1, b_2, b_3\} = \{2, 4, 6\}$  in real and virtual arrays, with placement of  $x_2^2 x_1^*$  highlighted. (a) Beam occurrences in real array. (b) Beam occurrences when  $L_v = 7$ ,  $\{b'_1, b'_2, b'_3\} = \{1, 2, 3\}$ . (c) Beam occurrences when  $L_v = 13$ ,  $\{b'_1, b'_2, b'_3\} = \{1, 2, 5\}$ .

In Figure 3c, it is shown that a larger virtual array (here  $L_v = 13$ ) can also be used to first generate all IM terms on distinct beam indices and then shift them to match the pattern of IM overlap in the real array.

In all, this subsection showcases that: (1) the virtual domain size can be made smaller in special cases and (2) the resulting virtual domain sizes in the proposed approach are, at worst, similar in size to the cardinality of unique main beam and IM beam directions.

#### 4.3. Further Notes on Beamforming

For analytical clarity, this work has focused on LOS channels, wherein the columns of the DFT precoder matrix of (9) describe the inverse channel for transmission in those orthogonal LOS directions. However, more generally, if the received signal can be related to the wireless channel model as

$$Y = Hf_{\text{PA}}(z), \quad (15)$$

then the method is still applicable with some modifications. Looking at the bar plots in Figure 2, and exemplified in (14), it is seen that the main and IM beam components are separated after the IFFT. Thus, the final block can be replaced with a different precoding matrix  $W$ . Presume for example that some general beamforming vector  $w^{(1)}$  should be applied to spatially direct  $x_1$ , and  $w^{(2)}$  to spatially direct  $x_2$ . Seen exemplified in (5), the third-order off-beam beamforming vectors will then be  $(w^{(1)})^2(w^{(2)})^*$  and  $(w^{(2)})^2(w^{(1)})^*$ , for  $x_1^2x_2^*$  and  $x_2^2x_1^*$ , respectively. Thus, the precoder in the final block could be computed with  $W = [w^{(1)} \ w^{(2)} \ (w^{(1)})^2(w^{(2)})^* \ (w^{(2)})^2(w^{(1)})^*]$  instead of through an FFT. A simple simulation case is shown in Section 5.6, with more complex cases left for future work.

Realistic channels would in practice also include phenomena such as fading, interference, and multi-path effects. This can require more sophisticated approaches to channel estimation and coding, and can be addressed in many different ways, e.g., as found in [19,20]. Practical precoding strategies, such as the Enhanced Type II codebook in 5G NR in 3GPP, will in many complex channel scenarios often rely on a grid of beams, but in addition consider frequency domain information in the processing [21,22]. This is based on the observation that any realistic channel can be decomposed into the sum of orthogonal beams. Additionally, such codebooks feature oversampling of the DFT beamgrid, allowing flexible beam selection for users who would otherwise fall in between beams on a regular beamgrid. Depending on the practical scenario, the spatial sparsity of the channel can take one of three forms for the precoder. Consider this for a tri-user scenario. Let  $B$  denote the unique number of beam directions present due to nonlinearity. In the first form, the precoding scheme decides that the three users can be perfectly served with three main beams, and  $B - 3$  off-beams, on the DFT beamgrid. This form is optimally sparse and can yield the largest possible reduction in complexity for VA-DPD, since IM beams may overlap and thereby reduce  $B$ . In the second form, the precoder decides that three main beams are dominant in serving the users, although performance can be improved by injecting weak signals on other beam indices as well. In this case, the VA-DPD can decide to only consider the dominant beams for linearization, whilst IM from low-magnitude beams is neglected. This gives form one, but by approximation. In the third form, the precoder decides that power needs to be distributed evenly across several beams for each respective user. In this case, the VA-DPD can treat the sum of these beamgrid vectors as one effective precoding vector. For example, if the first user needs signal injection on beam indices  $\{b_1, b_3, b_7\}$  with respective magnitudes of 0.27, 0.33, 0.4, then the joint precoding vector can analytically be considered as  $w^{(1)} = 0.27\omega^{(b_1)} + 0.33\omega^{(b_3)} + 0.4\omega^{(b_7)}$ . As such, the earlier analysis of this subsection can be used.

Finally, advances in MIMO beamforming research include concepts such as cooperative networks [23], as well as content-centric beamforming rather than user-centric beamforming [24]. For the latter, the VA-DPD framework can be applied without any loss of generality by interpreting  $x_1, \dots, x_K$  as  $K$  content streams rather than user streams.

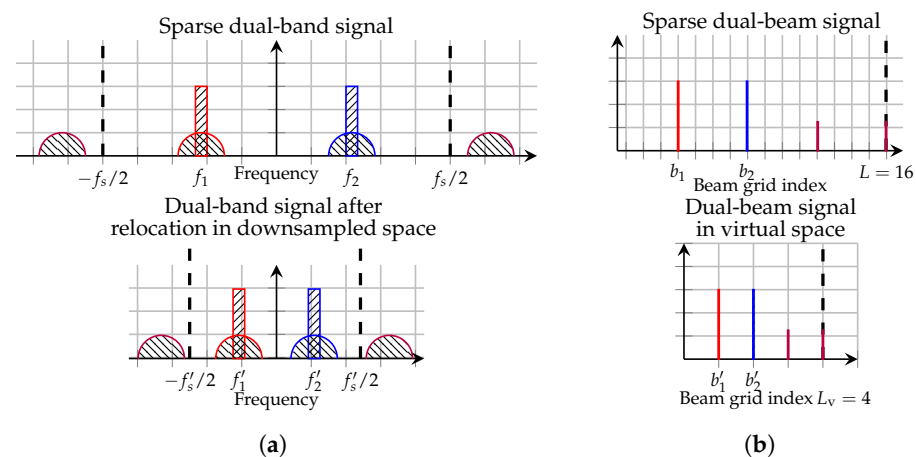
#### 4.4. Generalization to Higher-Order Terms

The analysis so far has focused on third-order IM terms, but extension to higher-order nonlinearities is straightforward. By augmenting the PA model in (4) with higher-order

terms and repeating the derivations, additional IM products appear. For instance, including a fifth-order term in the PA model gives rise to IM components such as  $x_3^3 x_1^* x_2^*$ , which, following the same indexing rules as before, appear at beam index  $3b_3 - b_1 - b_2$ . As for the third-order IM terms, some of these higher-order IM terms will always be directed towards the users, and some will be off-beams. Considering all possible high-order IM beams, the beam domain is only sparse for larger  $L$ , and for VA-DPD to guarantee complete equivalence to PP-DPD for any beam configuration,  $L_v$  must be increased accordingly. However, the fifth-order IM terms will typically have significantly lower power than the first- and third-order terms. Thus, they can often be neglected in off-beam directions. At the end, it is also a trade-off between DPD complexity costs against PA efficiency gains that dictates how nonlinear the PAs should be driven.

#### 4.5. Analogy with Frequency Relocation

Whilst not a direct counterpart, the proposed VA-DPD approach bears conceptual similarities to frequency relocation as in [11]. Namely, a signal that is sparse in one of its domains is densified in that domain prior to DPD processing. In the upper part of Figure 4a, a sparse dual-band signal, with the respective bands centered on frequencies  $f_1$  and  $f_2$  and sampled with rate  $f_s$ , is shown. IM terms here appear outside the first Nyquist zone. It is then, in the lower part, illustrated how the frequency domain can be densified by relocating the bands to frequencies  $f'_1$  and  $f'_2$ , thus enabling a sample rate reduction to  $f'_s$ . The DPD processing is then carried out at this lower rate before the frequency shift is reversed.



**Figure 4.** Comparison of a dual-band frequency relocation approach for densifying the frequency spectrum to the virtual-array approach for densifying the beam domain. (a) Power spectral density of a dual-band signal passed through a nonlinearity, with induced IM, in the frequency domain. (b) Power of a multi-beam signal passed through a nonlinearity, with induced IM, in the beam domain.

Figure 4b illustrates the corresponding spatial interpretation. The upper part shows the DFT beam-domain representation of an  $L$ -element array transmitting two beams at indices  $b_1$  and  $b_2$ , along with their associated IM components. In the lower part, these beams (and their IM products) are mapped onto a smaller virtual beamgrid of size  $L_v$ , creating a densified representation of the same spatial content. Instead of allowing a sample rate reduction, the corresponding gain is that it allows fewer DPDs to linearize a larger array.

A challenge in multi-band frequency relocation is when bands are unevenly spaced, giving rise to IM products that end up at new frequency locations. In such cases, in [11], uniform downsampling whilst maintaining the relative spacing of bands is performed. This, however, will often not densify the frequency spectrum to only contain the relevant bands and IM products. In contrast, VA-DPD can in the multi-beam setting achieve a

greater degree of dimensionality reduction by careful choice of the relocated beam indices in the virtual domain.

## 5. Numerical Setup and Results

For the numerical simulations, MATLAB (v2025b) is used. A setup with  $L = 28$  or 64 antennas, depending on the subsection, in a uniform linear array serving  $K = 2$  or 3 users, is considered. Each PA is modeled with a baseband Saleh model [25], equal on each PA branch, with added Gaussian noise  $\mathbf{n}_l$  at the output as

$$f_{\text{PA},l}(z_l) = \frac{az_l}{1 + b|z_l|^2} + \mathbf{n}_l, \quad (16)$$

with  $a = 1, b = 0.47$ . The desired signal is set to

$$\mathbf{Y}_d = a_d \mathbf{H} \mathbf{z}, \quad (17)$$

with  $a_d = 0.95$ . The noise power of  $\mathbf{n}_l$  is chosen to give signal-to-noise ratios between 50 and 60 dB.  $\mathbf{H}$  is chosen as an LOS channel, from  $L$  branches to  $Q = 181$  evenly spaced observation directions in the azimuth plane. Each channel element is  $h_{i_1, i_2} = \exp(j2\pi i_1 d \cos(\phi_{i_2})/\lambda)$ , where  $d$  denotes the array antenna spacing and  $\lambda$  the wavelength, with  $d/\lambda = 0.5$ ,  $\phi_{i_2}$  is the azimuth angle, and  $i_1 \in [-L/2, -L/2 + 1, \dots, L/2 - 1] + 1/2$ ,  $i_2 \in [1, Q]$ . The baseband user data streams are generated through bandlimiting complex Gaussian noise to a bandwidth of 30 MHz, and normalized to a root mean square value of 0.13, with a peak-to-average power ratio of about 8.3 dB. For each user stream,  $N = 655,36$  baseband samples are generated.

Four linearization cases are simulated per transmit configuration: not using DPD (No DPD), using per-PA (PP) DPD, using beam-domain (BD) DPD based on [5,6], and using virtual-array (VA) DPD.

### 5.1. DPD Identification

With this setup, conventional iterative learning control (ILC) [26] is modified with a simple extension and used to find the optimal input signals  $z_l$  to each PA in the array. Let

$$\mathbf{z}^{[i]} = \begin{bmatrix} z_1^{[i]} \\ \vdots \\ z_L^{[i]} \end{bmatrix}. \quad (18)$$

denote the DPD input at iteration  $[i]$ . Then, the DPD input at iteration  $[i + 1]$ ,  $\mathbf{z}^{[i+1]}$ , based on the OTA measured  $\mathbf{Y}^{[i]}$  and desired signal  $\mathbf{Y}_d$ , is found as

$$\mathbf{z}^{[i+1]} \leftarrow \mathbf{z}^{[i]} + \eta \mathbf{H}^\dagger (\mathbf{Y}_d - \mathbf{Y}^{[i]}), \quad (19)$$

Here,  $\eta$  is a step parameter for the iterations,  $\mathbf{H}^\dagger = (\mathbf{H}^* \mathbf{H})^{-1} \mathbf{H}^*$  is the pseudoinverse of the channel matrix  $\mathbf{H}$ , and  $*$  denotes the Hermitian transpose. This ILC procedure is suitable for performing offline as an initial identification, whilst online adaptation of DPD coefficients can be done with other known methods. The signals  $z_l$  are, after ILC convergence, shifted to the virtual domain, yielding target signals  $z'_l$ , as in Figure 2. The shifting is performed by doing the VA processing steps in reverse. That is, an IFFT is performed on  $\mathbf{z}$  along the branches, followed by relocating the streams of different considered beams from  $\mathbf{z}$  to their VA placement  $\{b'_1, \dots, b'_B\}$ , whilst removing streams that were zero-padded in the forward path. Lastly, an FFT is done in the virtual domain to go from the virtual beam domain to the virtual antenna domain and obtain  $z'_l$ . Forming a nonlinear model from  $u'_l$  to  $z'_l$  in

the virtual domain through polynomial DPD models yields a predistorted virtual domain signal and basis

$$z'_l = \mathbf{F}_{u'_l} \boldsymbol{\theta}_l, \mathbf{F}_{u'_l} = [u'_l, u'_l |u'_l|^2, \dots, u'_l |u'_l|^{P-1}] \quad (20)$$

where DPD coefficients are obtained as

$$\boldsymbol{\theta}_l = \mathbf{F}_{u'_l}^\dagger z'_l. \quad (21)$$

For conventional PP-DPD, the same relations (20) and (21) are used, but with target signals  $z_l$  and inputs  $u_l$  in the original antenna domain (as in Figure 1). For BD-DPD, each unique basis function expressed in terms of  $x_1, \dots, x_K$ , obtainable from the expansion of  $\sum_{p=0,2,\dots}^{P-1} z_l |z_l|^p$ , is used. This includes off-beam terms as in Table 1, fifth-order terms as in Table 2, and main beam terms such as  $x_1 |x_1|^2, x_2 |x_3|^2$ .

**Table 2.** Fifth-order IM terms for static nonlinearity of order  $P = 5$ , with  $K = 2, 3$  concurrent users.

| Number of Users | Fifth-Order Terms  | Total Number of Fifth-Order Terms |
|-----------------|--|-----------------------------------|
| $K = 2$         | $x_1  x_1 ^4, x_1  x_2 ^4, x_1  x_1 ^2  x_2 ^2$<br>$x_2  x_2 ^4, x_2  x_1 ^4, x_2  x_1 ^2  x_2 ^2$<br>$x_1^3 (x_2^*)^2, x_2^3 (x_1^*)^2$<br>$x_1  x_1 ^2 x_2^2, x_2  x_2 ^2 x_1^2$<br>$x_1^2  x_1 ^2 x_2^*, x_2^2  x_2 ^2 x_1^*$ | 12                                |
| $K = 3$         | $x_1  x_1 ^4, \dots, x_1  x_2 ^2  x_3 ^2$  | 60                                |

## 5.2. Complexity Analysis

To quantitatively compare BD-DPD and VA-DPD, this subsection provides a brief complexity analysis. It limits its scope to the filtering complexity, as the initial identification can be done offline, and as the filtering complexity typically dominates over the adaptation complexity [5,27].

Consider first the VA-DPD scheme in Figure 2. The computational complexity of the scheme consists of four parts: the VA FFT precoding, VA DPDs, the IFFT, and precoding back to the real array. The first step maps  $K$  user streams to  $L_v$  virtual branches using an  $L_v$ -point FFT. The complexity  $F_{\text{FFT}}(L_v)$  depends on the FFT size and implementation:

- $L_v = 4$ : A radix-4 FFT requires 16 real additions  $\Rightarrow F_{\text{FFT}}(4) = 16$  FLOPs.
- $L_v = 13$  (prime): A Winograd 13-point FFT requires 42 real multiplications and 188 real additions,  $\Rightarrow F_{\text{FFT}}(13) = 230$  FLOPs.
- $L_v = 30$ : A Good–Thomas FFT exploiting  $2 \times 3 \times 5 = 30$  using a two-point FFT + Winograd-3 and Winograd-5 kernels: 50 additions + 22 multiplications,  $\Rightarrow F_{\text{FFT}}(30) = 576$  FLOPs [28].

These costs are reported separately from the DPD complexity, as FFT implementation in hardware may differ significantly from DPD processing. In cases where  $F_{\text{FFT}}(L_v)$  is high, it may even be beneficial to increase  $L_v$  to reduce the FFT cost at the expense of a slight increase in DPD complexity. The final IFFT has the same cost and is therefore included as a factor of 2 in Table 3.

As for the VA-DPD units, complexity will scale corresponding to the number of basis functions and virtual branches as  $B_{\text{basis}} L_v$ . For an odd-order memory polynomial (MP) of nonlinearity order  $P$  and memory  $M$ ,  $B_{\text{basis}} = 4(P + 1)(M + 1)$  [27]. Finally, the precoder goes from  $L_v$  virtual branches to  $L$  real branches. However, this complexity is omitted from comparison, as precoding to the real array must be performed for all DPD methods, to approximately the same computational cost.

**Table 3.** DPD filtering complexity for the different schemes, assuming implementation of memory-polynomial bases, given in the number of floating-point operations (FLOPs) per baseband sample. The number of lookup (LUT) table entries is also listed, if DPDs are implemented as LUTs instead.  $T$  is the LUT size. The virtual-array FFT/IFFT complexity is listed as  $F_{\text{FFT}}(L_v)$ .

| Operation       | Per-PA DPD<br>(FLOPs)/(# LUT Entries) | Beam-Domain DPD<br>(FLOPs)/(# LUT Entries)   | Virtual-Array DPD<br>(FLOPs)/(# LUT Entries) |
|-----------------|---------------------------------------|--|--|
| FFT/IFFT        | NA/NA                                 | NA/NA  | $2F_{\text{FFT}}(L_v)/\text{NA}$             |
| DPD computation | $4(P+1)(M+1)L/(M+1)TL$                | $8(M+1) \sum_{p=1,3,\dots}^P  \kappa_p  /K(M+1)T^K +\sum_{i=0}^{K-2} B_i(M+1)^{K-i}$ | $4(P+1)(M+1)L_v /(M+1)TL_v$                  |

The filtering complexity for BD-DPD implementing an MP model is  $8(M+1) \sum_{p=1,3,\dots}^P |\kappa_p|$ , where  $\kappa_p$  denotes the cardinality of the set  $\kappa_p$ , i.e., the number of bases for nonlinearity terms of order  $p$ . Only odd-order polynomials are considered. For example, for a two-user case,  $\kappa_5 = 12$ , as seen from Table 2. After the multi-dimensional DPD, a precoder of dimension  $L$  follows, as in the VA-DPD case.

The filtering complexity for an MP PP-DPD follows as  $4(P+1)(M+1)L$ . Table 3 summarizes the comparison. In Table 3, the number of coefficients necessary for simple lookup table implementations is provided as well, based on [29]. That is, whilst each VA-DPD and PP-DPD branch can adapt a 1D LUT per branch, as

$$z_l[n] = \sum_{m=0}^{M-1} u_l[n-m] f_{\text{LUT}}(|u_l[n-m]|), \quad (22)$$

beam-domain approaches in direct extension need to consider multi-dimensional LUTs. For example, for the first main beam, they are of the form

$$u_1[n-m] = \sum_{m=0}^{M-1} x_1[n-m] f_{\text{LUT}}(|x_1[n-m]|, \dots, |x_K[n-m]|). \quad (23)$$

Correspondingly, LUTs for off-beam IM terms also need to be adapted, exemplified as

$$u_i[n-m] = \sum_{m=0}^{M-1} x_1[n-m] x_2[n-m] x_3^*[n-m] \cdot f_{\text{LUT}}(|x_1[n-m]|, |x_2[n-m]|, |x_3[n-m]|), \quad (24)$$

where  $i \in [1, B]$ . For PP-DPD and VA-DPD, each branch requires  $(M+1)$  one-dimensional LUTs. Denote the LUT depth as  $T$ . Then, VA-DPD requires  $(M+1)TL_v$  coefficients, whilst PP-DPD requires  $(M+1)TL$ . For BD-DPD, a multi-dimensional LUT per unique beam is necessary prior to the precoder. Compared to Table 1, it is noted that off-beams will not always need  $K$ -dimensional LUTs. Instead, the necessary number of LUT coefficients can be computed as  $K(M+1)T^K + \sum_{i=0}^{K-2} B_i(M+1)^{K-i}$ . Here,  $B_i$  is the number of off-beam terms requiring  $(K-i)$ -dimensional LUTs.

As a final note, it is worth mentioning that the introduction of FFT and IFFT blocks in the VA-DPD is expected to introduce only minor latency, as the FFT lengths are small.

### 5.3. Tri-Beam Scenario

In Figure 5, a concurrent tri-beam scenario is shown. For this scenario, there are  $L = 64$  transmit branches, with beams in beam directions  $\{b_1, b_2, b_3\} = \{7, 23, 49\}$ . In Figure 5a,b, the power of the desired signal  $Y_d$ , over different angles, is visualized in contrast to the distortion power remaining after applying no DPD, PP-DPD, BD-DPD, and the proposed VA-DPD approach with either odd-order  $P = 3$  or  $P = 5$  DPDs. For VA-DPD,

the virtual array uses  $L_v = 13$  branches, meaning it can capture third-order beam direction information without loss. Consequently, Figure 5a highlights how VA-DPD is equivalent in performance to PP-DPD and BD-DPD for a DPD of order  $P = 3$ . This is seen from the fact that the yellow line, corresponding to PP-DPD distortion, is precisely covered by the purple and green ones, of VA-DPD and BD-DPD, respectively. On the other hand, compared to Figure 5a, Figure 5b shows that applying fifth-order DPDs whilst only considering up to third-order terms in the virtual-array dimensioning reduces the distortion power in the user directions equivalently to PP-DPD and BD-DPD, by an additional 5 dB compared to  $P = 3$ . However, the off-beam linearization performance is not improved correspondingly, as seen by the residual distortion, e.g., around angles 28–35°. The power spectral density of the desired signal in the user directions and the residual error after linearization are shown for the respective methods in Figure 5c. The normalized mean square error (NMSE) between the desired and received user signal is about  $-50$  dB for each method in all user directions, for  $P = 5$ .

With  $L_v = 13$  branches, no memory, and  $P = 5$ , Table 3 gives a DPD unit complexity of 312 FLOPs, and  $2F_{\text{FFT}}(13) = 460$  FLOPs. Notably, for such simple DPD units, the FFT complexity is large. However, for more complex DPDs, the impact is relatively diminished. Alternatively, for the DPD units,  $L_v = 13$  1D-LUT lookups are done to find the appropriate predistorted values per sample. Assuming a LUT depth of  $T = 10$ , this gives 130 LUT entries. In contrast, PP-DPD requires 1536 FLOPs or 640 LUT entries. Thus, the DPD unit complexity is reduced by  $100 \cdot |312 - 1536| / 1536 \approx 80\%$  in VA-DPD compared to PP-DPD, though the FFTs also need to be accounted for. Lastly, as seen from Tables 1 and 2, BD-DPD needs to account for 3 linear, 12 third-order and 60 fifth-order terms. Table 3 shows that with  $B = 12$  unique data streams, BD-DPD requires 648 FLOPs or 6600 LUT entries.

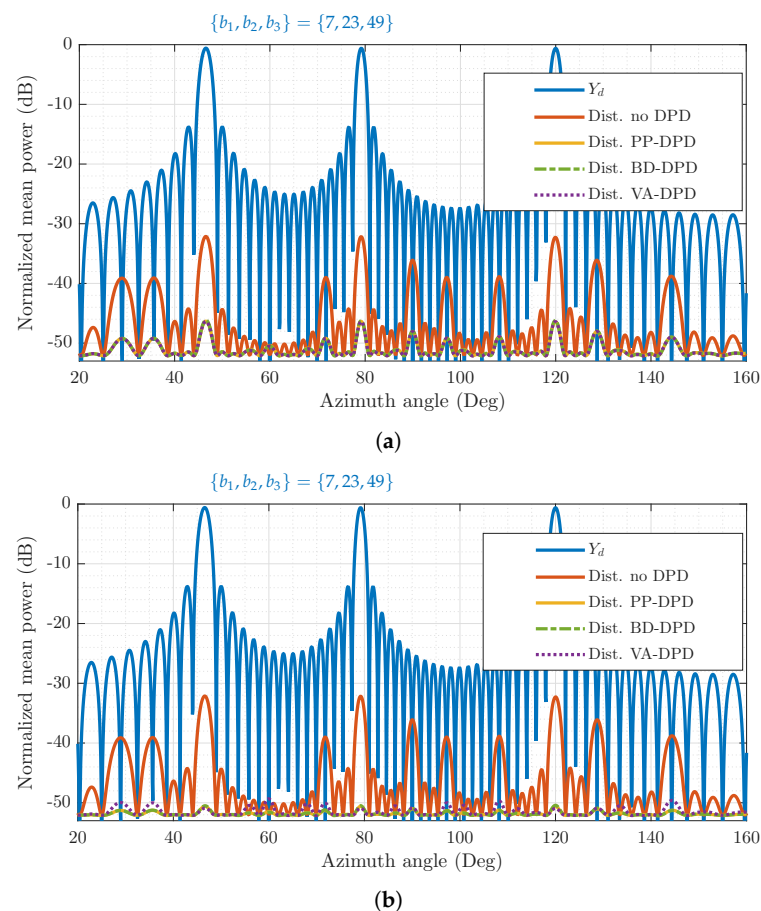
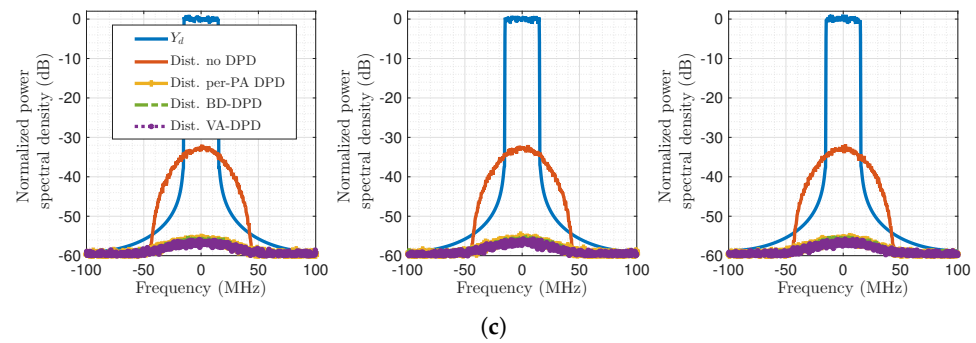


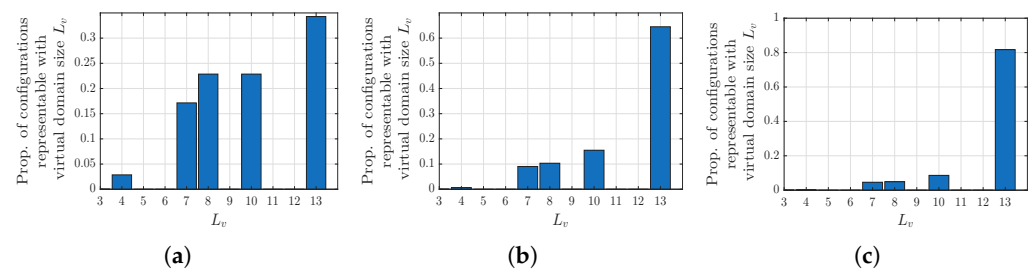
Figure 5. Cont.



**Figure 5.** Normalized power of the desired signal  $Y_d$  shown in comparison to the distortion power remaining after no DPD, per-PA (PP) DPD, the proposed virtual-array (VA) DPD, and beam-domain (BD) DPD, for concurrent tri-beam transmission. DPDs include odd-order terms up to, and including, order  $P$ . (a) Mean power in varying spatial directions,  $P = 3$ . (b) Mean power in varying spatial directions,  $P = 5$ . (c) Power spectral density in user directions,  $P = 5$ .

#### 5.4. Virtual-Array Dimensioning

The results of this section describe how small the virtual array can be dimensioned in terms of  $L_v$ , to account for all IM beams, with respect to special cases as described in Section 4.2. In Figure 6, the proportion of beam configuration  $\{b_1, b_2, b_3\}$  that can be represented with a virtual array of length  $L_v$  is shown for varying array sizes  $L = 16, 32, 64$ , and in Tables 4 and 5, it is shown for various  $L$  and  $P = 3, 5$ . Most noticeably, the proportion of cases requiring a larger  $L_v$  increases with  $L$ . This correlates to the fact that in a larger beam space, the IM products end up in unique beam directions more often. The results highlight that there are cases when the beam domain is especially sparse, such that the VA-DPD can be performed at an even lower computational complexity.



**Figure 6.** Proportion of main beam configuration cases in the original domain of size  $L$  that can be represented by a virtual domain of size  $L_v$ , considering IM terms of order  $P = 3$ . (a) Array size  $L = 16$ . (b) Array size  $L = 32$ . (c) Array size  $L = 64$ .

**Table 4.** Percentage (%) of main beam configurations in the original domain of size  $L$  that can be represented by a virtual domain of size  $L_v$ , for IM terms of order  $P = 3$ .

| $L \backslash L_v$ | 3  | 4       | 5       | 6 | 7  | 8  | 9 | 10 | 12 | 13 |
|--------------------|----|---------|---------|---|----|----|---|----|----|----|
| 20                 | 0  | 2       | 4       | 0 | 11 | 14 | 0 | 21 | 0  | 49 |
| 24                 | <1 | 1       | 0       | 4 | 10 | 12 | 0 | 14 | 21 | 38 |
| 30                 | <1 | 0       | 2       | 2 | 7  | 9  | 0 | 15 | 18 | 47 |
| 36                 | <1 | 1       | 0       | 2 | 7  | 7  | 3 | 11 | 12 | 58 |
| 130                | 0  | 0       | $\ll 1$ | 0 | 2  | 2  | 0 | 5  | 0  | 91 |
| 256                | 0  | $\ll 1$ | 0       | 0 | 1  | 1  | 0 | 2  | 0  | 95 |

**Table 5.** Percentage (%) of main beam configurations in the original domain of size  $L$  that can be represented by a virtual domain of size  $L_v$ , for IM terms of order  $P = 5$ .

| $L \backslash L_v$ | 4       | 8  | 11 | 12 | 16 | 21 | 22 | 24 | 26 | 30 |
|--------------------|---------|----|----|----|----|----|----|----|----|----|
| 32                 | 1       | 4  | 8  | 8  | 23 | 10 | 5  | 21 | 10 | 10 |
| 64                 | <1      | 1  | 4  | 4  | 10 | 7  | 4  | 15 | 7  | 47 |
| 128                | $\ll 1$ | <1 | 2  | 2  | 5  | 4  | 2  | 8  | 4  | 71 |

### 5.5. Power Amplifier Variations

It was noted earlier that BD-DPD assumes equal PAs across branches. It is this assumption that mathematically allows for a common DPD to be used prior to the precoder in BD-DPD, without losing any ability to linearize the system [6]. This assumption is not quite as rigidly held in the proposed VA-DPD, as DPD variations across the virtual-array branches can be adapted. As such, to investigate whether or not VA-DPD might handle PA variations better in comparison to BD-DPD, the PA model is changed to

$$f_{PA,l}(x) = c_{1,l}x + c_{3,l}x|x|^2, \quad (25)$$

to more easily represent gain and phase variations in the nonlinear PA response relative to the linear response, variations which cannot be compensated by calibration. The first-order coefficients are set to  $c_{1,l} = 1$  for all branches, whilst  $c_{3,l}$  are varied from a nominal value of  $-0.55$  with magnitude standard deviation  $0.1$  and random phases from a uniform distribution in the range of  $[-20, 20]^\circ$ . Figure 7 showcases a tri-beam case with  $\{b_1, b_2, b_3\} = \{1, 3, 10\}$ ,  $L = 16$ ,  $L_v = 7$ . In Figure 7a,c, the PAs exhibit no variations, whilst in Figure 7b,d, they do. Notably, both VA-DPD and BD-DPD deteriorate in performance relative to PP-DPD with PA variations. However, VA-DPD reduces distortion more than BD-DPD in all user directions. This is because, whilst BD-DPD adapts one set of coefficients prior to the precoder to hold for linearization across all branches, VA-DPD can have some DPD coefficient variations with  $l = 1, \dots, L_v$ . For VA-DPD, the NMSE is about  $-42$  dB in each user direction when no PA variations are present, and about  $-40$  dB with PA variations. The corresponding values for PP-DPD are  $-42$  dB and  $-43$  dB, and for BD-DPD are  $-42$  dB and  $-36$  dB.

For both BD-DPD and VA-DPD, additional consideration of PA variations across branches can be necessary, depending on the severity of the PA variations. One approach of mitigation is to account for variations through minor DPD computations at each branch, whilst the main part of predistortion is still performed as per the respective method, similar to [4].

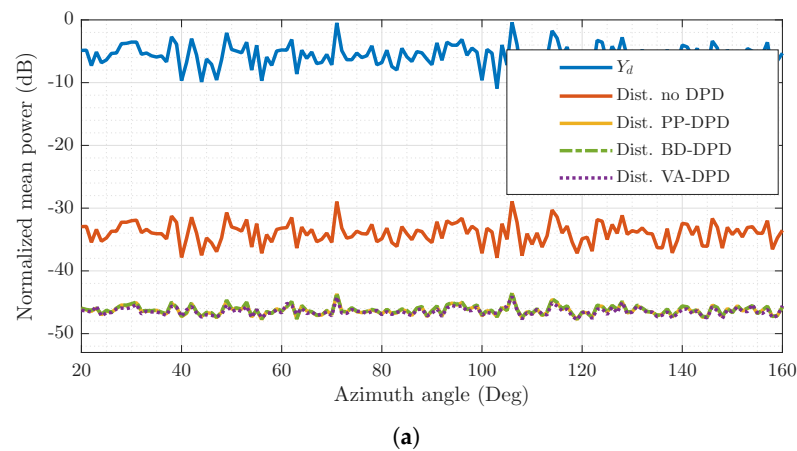
### 5.6. Non-FFT Precoder

In Figure 8, a two-user scenario for a random channel  $\mathbf{H}$  of dimension  $L \times Q$  consisting of complex Gaussian numbers distributed as  $\mathcal{CN}(0, 1)$  is shown, with users in directions  $[71, 106]^\circ$ . The array is of size  $L = 28$ , and virtual array of size  $L_v = 4$ . Precoding vectors are chosen as the pseudoinverse of the respective channel vectors, i.e.,  $\mathbf{w}_1 = \mathbf{H}_{71}^\dagger$ ,  $\mathbf{w}_2 = \mathbf{H}_{106}^\dagger$ . First, it is important to note that all linearization methods achieve about 20 dB improvement in distortion reduction. Secondly, neither the linear power nor the residual distortion forms clear beam directions but is instead spread across the spatial range due to the randomness of the channel. From Table 3, VA-DPD has a DPD unit filtering complexity of 128 FLOPs for  $P = 7$ ,  $M = 0$ ,  $2F_{\text{FFT}}(4) = 32$  FLOPs, and requires  $L_v T = 4 \cdot 10 = 40$  coefficients for LUTs of size 10. Correspondingly, BD-DPD requires 256 FLOPs and 220 LUT coefficients, whilst PP-DPD costs 896 FLOPs and would require

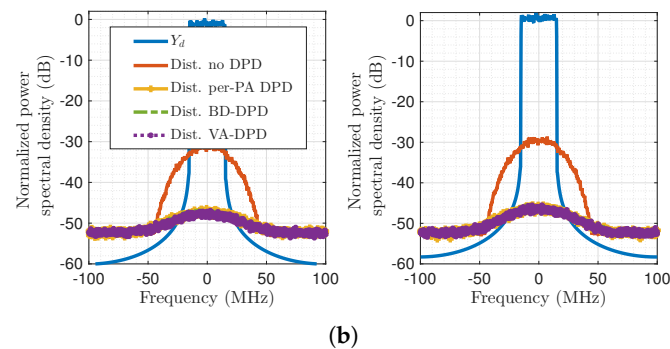
280 LUT coefficients. Thus, a  $100 \cdot |156 - 896| / 896 \approx 83\%$  reduction in DPD unit complexity is achieved for VA-DPD compared to PP-DPD. For all DPD methods, an NMSE of about  $-40$  dB is obtained in each user direction.



**Figure 7.** Normalized power of the desired signal  $Y_d$  shown in comparison to the distortion power remaining after no DPD, per-PA (PP) DPD, the proposed virtual-array (VA) DPD, and beam-domain (BD) DPD, for concurrent tri-beam transmission, with or without PA variations. DPDs include first- and third-order terms. (a) Mean power in varying spatial directions. No PA variations. (b) Mean power in varying spatial directions. PA variations. (c) Power spectral density in user directions. No PA variations. (d) Power spectral density in user directions. PA variations.



**Figure 8.** Cont.



**Figure 8.** Normalized power of the desired signal  $Y_d$  shown in comparison to the distortion power remaining after no DPD, per-PA (PP) DPD, the proposed virtual-array (VA) DPD, and beam-domain (BD) DPD, for concurrent dual-beam transmission in a random channel.  $P = 7$  is used for the DPDs. (a) Mean power in varying spatial directions. (b) Power spectral density in user directions.

## 6. Conclusions

The proposed DPD schematic, using FFT beamforming and virtual-array DPD processing, is shown to be capable of reducing DPD complexity manyfold in contrast to conventional per-PA DPD, when many transmit branches serve few users. By theoretical derivation and simulation, it is shown how an arbitrarily large array can be linearized by no more than 4, 13, or 30 virtual-array branches when only 2, 3, or 4 users, respectively, are served. The virtual-array DPD scheme is shown to be operable for varying degrees of PA nonlinearities, and for different channel conditions and multi-user scenarios. In contrast to previous approaches utilizing a beam-domain framework, the proposed approach utilizes single-input, single-output DPD units that are computationally cheap to implement. Additionally, it is shown how VA-DPD can adapt to PA variations better than beam-domain-oriented solutions. Overall, the approach offers a flexible and powerful alternative to multi-user linearization for large MIMO arrays with digital beamforming.

Future work will include experimental validation and studying practical concerns such as FPGA implementation feasibility and other real-time constraints. Additionally, the power consumption trade-off between PA back-off and DPD complexity warrants further study.

**Author Contributions:** Conceptualization, B.L. and R.H.; implementation, B.L.; technical refinement and generalization, all authors; writing, B.L.; review and editing, all authors; funding acquisition, T.E., C.F. and R.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Vinnova, Formas, and the Swedish Energy Agency within the strategic innovation program ‘Smartare Elektroniksystem’, and by Chalmers University of Technology and Ericsson AB.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data generated in this work, described in Section 5, can be reproduced with the MATLAB function `randn`.

**Conflicts of Interest:** Björn Langborn, Christian Fager, and Thomas Eriksson declare no conflicts of interest. Rui Hou is employed by Ericsson AB. The company played no role in the design of the study, the collection, analysis, or interpretation of data, the writing of the manuscript, or the decision to publish the article.

## Appendix A. Finding a Virtual DPD Space

Finding a virtual-array representation for a given array of size  $L$  transmitting beams  $b_1, \dots, b_K$  involves finding an appropriate virtual-array size  $L_v$  and virtual beam placement  $b'_1, \dots, b'_K$ . In the original antenna domain, for a tri-beam case, main and intermodulation beams will end up on integer indices  $c_1, \dots, c_{12}$ , as in

$$\begin{aligned}
 b_1 &= c_1 \\
 b_2 &= c_2 \\
 b_3 &= c_3 \\
 (2b_1 - b_2) - L \left\lfloor \frac{2b_1 - b_2 - 1/2}{L} \right\rfloor &= c_4 \\
 &\vdots \\
 (b_2 + b_3 - b_1) - L \left\lfloor \frac{b_2 + b_3 - b_1 - 1/2}{L} \right\rfloor &= c_{12},
 \end{aligned} \tag{A1}$$

where  $\{c_1, \dots, c_{12}\} \in [1, L]$ . The set  $\{c_1, \dots, c_{12}\}$  might not be a unique set of 12 integers. This can be exemplified with beams  $\{b_1, b_2, b_3\} = \{1, 2, 4\}$ ,  $L = 16$ , where the off-beam indices are  $\{16, 3, 14, 7, 16, 6, 15, 3, 5\}$  such that (A1) becomes

$$\begin{aligned}
 b_1 &= c_1 = 1 \\
 b_2 &= c_2 = 2 \\
 b_3 &= c_3 = 4 \\
 (2b_1 - b_2) - L \left\lfloor \frac{2b_1 - b_2 - 1/2}{L} \right\rfloor &= c_4 = c_8 = 16 \\
 (2b_2 - b_1) - L \left\lfloor \frac{2b_2 - b_1 - 1/2}{L} \right\rfloor &= c_5 = c_{11} = 3 \\
 &\vdots \\
 (b_1 + b_3 - b_2) - L \left\lfloor \frac{b_1 + b_3 - b_2 - 1/2}{L} \right\rfloor &= c_{11} = c_5 = 3 \\
 (b_2 + b_3 - b_1) - L \left\lfloor \frac{b_2 + b_3 - b_1 - 1/2}{L} \right\rfloor &= c_{12} = 5,
 \end{aligned} \tag{A2}$$

The objective is then to find a virtual subspace with  $L_v, b'_1, b'_2, b'_3$  such that

$$\begin{aligned}
 b'_1 &= c'_1 \\
 b'_2 &= c'_2 \\
 b'_3 &= c'_3 \\
 (2b'_1 - b'_2) - L_v \left\lfloor \frac{2b'_1 - b'_2}{L_v} \right\rfloor &= c'_4 = c'_8 \\
 (2b'_2 - b'_1) - L_v \left\lfloor \frac{2b'_2 - b'_1}{L_v} \right\rfloor &= c'_5 = c'_{11} \\
 &\vdots \\
 (b'_1 + b'_3 - b'_2) - L_v \left\lfloor \frac{b'_1 + b'_3 - b'_2}{L_v} \right\rfloor &= c'_{11} = c'_5 \\
 (b'_2 + b'_3 - b'_1) - L_v \left\lfloor \frac{b'_2 + b'_3 - b'_1}{L_v} \right\rfloor &= c'_{12},
 \end{aligned} \tag{A3}$$

for any  $\{c'_1, \dots, c'_{12}\} \in [1, L_v]$ . This is a Diophantine system of equations and is generally not trivial to solve. However, in practice, it can be solved offline, and the indices can be obtained during runtime using a lookup table. In this work, the pseudo-algorithm of Algorithm A1 describes one approach to solving the systems of equations, for a general case with an arbitrary number of beams  $K$  and branches  $L$ . Depending on  $K$ , heuristic knowledge can limit the search space  $L_v$ s and  $m_b$ s in Algorithm A1. One general constraint is  $L_v \geq |\{b'_1, b'_2, \dots, b'_2 + b'_3 - b'_1\}|$ ; i.e., the virtual domain size can never be reduced below the cardinality of the set it is supposed to represent.

---

**Algorithm A1** Algorithm to find virtual space representation.

---

```

Require:  $m_b, L$  // Main beam indices  $m_b$  and array size  $L$ 
// Get off-beams, and cardinality of full index set
ob = get_offbeams( $m_b, L$ )
ab = [ $m_b$  ob]
c = length(unique(ab))
// Try virtual subspace size  $L_v$  out of candidates  $L_v$ s
for  $L_v=L_v$ s do
// Try virtual main beams  $m_{bv}$  out of candidates  $m_{bv}$ s
for  $m_{bv}=m_{bv}$ s do
// Get off-beam indices for  $m_{bv}$ 
obv = get_offbeams( $m_{bv}, L_v$ )
abv = [ $m_{bv}$  obv]
// Relabel ab and abv such that the first element, and
all other elements in the vector with the same value,
is labelled 1. Do correspondingly for the remaining elements.
Obtain ab_relabel, abv_relabel
if ab_relabel==abv_relabel then
// Virtual array representation  $\{m_{bv}, L_v\}$  found.
end if
end for
end for

```

---

## References

1. Kakkavas, A.; Seco-Granados, G.; Wymeersch, H.; Garcia, M.H.C.; Stirling-Gallacher, R.A.; Nossek, J.A. 5G Downlink Multi-Beam Signal Design for LOS Positioning. In *Proceedings of the IEEE Global Communications Conference*; IEEE: New York, NY, USA, 2019. [CrossRef]
2. Liu, X.; Chen, W.; Chu, J.; Ghannouchi, F.M.; Feng, Z. Multi-Stream Spatial Digital Predistortion for Fully-Connected Hybrid Beamforming Massive MIMO Transmitters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2021**, *68*, 2998–3011. [CrossRef]
3. Jing, J.; Yu, C. Multibeam Digital Predistortion for Millimeter-Wave Analog Beamforming Transmitters. *IEEE Microw. Wirel. Components Lett.* **2020**, *30*, 209–212. [CrossRef]
4. Yu, C.; Jing, J.; Shao, H.; Jiang, Z.H.; Yan, P.; Zhu, X.W.; Hong, W.; Zhu, A. Full-Angle Digital Predistortion of 5G Millimeter-Wave Massive MIMO Transmitters. *IEEE Trans. Microw. Theory Tech.* **2019**, *67*, 2847–2860. [CrossRef]
5. Brihuega, A.; Anttila, L.; Valkama, M. Beam-Level Frequency-Domain Digital Predistortion for OFDM Massive MIMO Transmitters. *IEEE Trans. Microw. Theory Tech.* **2023**, *71*, 1412–1427. [CrossRef]
6. Braithwaite, R.N. Amplifier Nonlinearities in an Antenna Array During Spatially-Multiplexed Transmissions. In *Proceedings of the IEEE Radio and Wireless Symposium*; IEEE: New York, NY, USA, 2020; pp. 235–238. [CrossRef]
7. Bassam, S.A.; Helaoui, M.H.; Ghannouchi, F.M. 2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters. *IEEE Trans. Microw. Theory Tech.* **2011**, *59*, 2547–2553. [CrossRef]
8. Liu, Y.J.; Chen, W.; Zhou, J.; Zhou, B.H.; Ghannouchi, F.M. Digital predistortion for concurrent dual-band transmitters using 2-D modified memory polynomials. *IEEE Trans. Microw. Theory Tech.* **2013**, *61*, 281–290. [CrossRef]
9. Eriksson, T.; Fager, C. Digital predistortion of concurrent multiband communication systems. In *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; IEEE: New York, NY, USA, 2014; pp. 3918–3922. [CrossRef]

10. Younes, M.; Kwan, A.; Rawat, M.; Ghannouchi, F.M. Three-Dimensional digital predistorter for concurrent tri-band power amplifier linearization. In *Proceedings of the IEEE MTT-S International Microwave Symposium Digest (MTT)*; IEEE: New York, NY, USA, 2013; pp. 1–4. [[CrossRef](#)]
11. Yu, C.; Xia, J.; Zhu, X.W.; Zhu, A. Single-Model Single-Feedback Digital Predistortion for Concurrent Multi-Band Wireless Transmitters. *IEEE Trans. Microw. Theory Tech.* **2015**, *63*, 2211–2224. [[CrossRef](#)]
12. Kolomvakis, N.; Bavand, M.; Bahceci, I.; Gustavsson, U. A Distortion Nullforming Precoder in Massive MIMO Systems with Nonlinear Hardware. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 1775–1779. [[CrossRef](#)]
13. Mollen, C.; Gustavsson, U.; Eriksson, T.; Larsson, E.G. Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6663–6679. [[CrossRef](#)]
14. Jing, J.; Shao, H.; Yan, P.; Jiang, Z.H.; Yu, C. Digital Predistortion of Millimeter-Wave Multi-beam Transmitters with Digital Beam-forming Network. In *Proceedings of the 2019 IEEE MTT-S International Wireless Symposium (IWS)*, Guangzhou, China, 19–22 May 2019; pp. 1–3. [[CrossRef](#)]
15. Molina, A.; Rajamani, K.; Azadet, K. Concurrent Dual-Band Digital Predistortion Using 2-D Lookup Tables with Bilinear Interpolation and Extrapolation: Direct Least Squares Coefficient Adaptation. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 1381–1393. [[CrossRef](#)]
16. Zhang, Y.; Chen, Q.; Peng, H.; Chen, W.; Feng, Z.; Ghannouchi, F.M. Stream-Level Digital Predistortion for Extremely Large-Scale MIMO Transmitters. *IEEE Trans. Microw. Theory Tech.* **2026**, *74*, 3503–3519. [[CrossRef](#)]
17. Wang, J.; Gao, K.; Liu, X.; Zhang, Y.; Chen, W.; Feng, H.; Feng, Z. Digital Predistortion Based on Clustering for Fully-Connected Hybrid Beamforming MIMO Transmitters with Dynamic Beam Steering. In *Proceedings of the 2024 IEEE MTT-S International Wireless Symposium (IWS)*; IEEE: New York, NY, USA, 2024; pp. 1–3. [[CrossRef](#)]
18. Wang, S.; Cao, W.; Hou, R.; Eriksson, T. A Digital Predistortion for Concurrent Dual-Band Power Amplifier Linearization Based on Periodically Nonuniform Sampling Theory. *IEEE Trans. Microw. Theory Tech.* **2022**, *70*, 466–475. [[CrossRef](#)]
19. Zhou, Y.; Ng, T.S. MIMO-OFCDM systems with joint iterative detection and optimal power allocation. *IEEE Trans. Wirel. Commun.* **2008**, *7*, 5504–5516. [[CrossRef](#)]
20. Sun, B.; Zhou, Y.; Yuan, J.; Shi, J. Interference Cancellation Based Channel Estimation for Massive MIMO Systems with Time Shifted Pilots. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 6826–6843. [[CrossRef](#)]
21. Ziao, Q.; Haifan, Y. A review of codebooks for CSI feedback in 5G new radio and beyond. *China Commun.* **2025**, *22*, 112–127. [[CrossRef](#)]
22. Ning, B.; Yin, H.; Liu, S.; Deng, H.; Yang, S.; Zhang, Y.; Mei, W.; Gesbert, D.; Park, J.; Heath, R.W.; et al. Precoding Matrix Indicator in the 5G NR Protocol: A Tutorial on 3GPP Beamforming Codebooks. *IEEE Commun. Surv. Tutor.* **2026**, *28*, 4581–4623. [[CrossRef](#)]
23. Tan, F.; Wu, P.; Wu, Y.C.; Xia, M. Cooperative Beamforming for Wireless Fronthaul and Access Links in Ultra-Dense C-RANs with SWIPT: A First-Order Approach. *IEEE J. Sel. Top. Signal Process.* **2021**, *15*, 1242–1257. [[CrossRef](#)]
24. Li, Y.; Xia, M.; Wu, Y.C. First-Order Algorithm for Content-Centric Sparse Multicast Beamforming in Large-Scale C-RAN. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 5959–5974. [[CrossRef](#)]
25. Saleh, A. Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers. *IEEE Trans. Commun.* **1981**, *29*, 1715–1720. [[CrossRef](#)]
26. Chani-Cahuana, J.; Landin, P.N.; Fager, C.; Eriksson, T. Iterative Learning Control for RF Power Amplifier Linearization. *IEEE Trans. Microw. Theory Tech.* **2016**, *64*, 2778–2789. [[CrossRef](#)]
27. Tehrani, A.S.; Cao, H.; Afsardoost, S.; Eriksson, T.; Isaksson, M.; Fager, C. A Comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models. *IEEE Trans. Microw. Theory Tech.* **2010**, *58*, 1510–1520. [[CrossRef](#)]
28. Blahut, R.E. *Fast Algorithms for Signal Processing*; Cambridge University Press: Cambridge, UK, 2010; p. 75.
29. Ding, L.; Yang, Z.; Gandhi, H. Concurrent dual-band digital predistortion. In *Proceedings of the 2012 IEEE/MTT-S International Microwave Symposium Digest*; IEEE: New York, NY, USA, 2012; pp. 1–3. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.