

# Beam Squint Effects in THz Communications with UPA and ULA: Comparison and Hybrid Beamforming Design

Nhan Thanh Nguyen, Joonas Kokkonen, and Markku Juntti

Centre for Wireless Communications, University of Oulu, P.O.Box 4500, FI-90014, Finland

Email: {nhan.nguyen, Joonas.Kokkonen, markku.juntti}@oulu.fi

**Abstract**—Future 6G and beyond wireless systems aim at ultra-high data rates from very large bandwidths in millimeter-wave and terahertz bands. Large bandwidth signals in antenna arrays lead to beam squint or frequency-selective array response causing losses in the beamforming gain. Recently, several hybrid analog-digital beamforming (HBF) schemes, such as the delay-phase precoding employing time delay networks or wider-beam codebook design, have been proposed to tackle the beam squint. However, the severity of beam squint has not been thoroughly investigated in the literature. In this paper, by analyzing the impacts of the beam squint on the uniform planar array (UPA) and uniform linear array (ULA) structures, we analytically and numerically show that the effect of beam squint is less severe in the former. Based on this fact, we propose a simplified HBF design in which the analog beamformer is designed for the center frequency, while digital beamformers for all subcarrier frequencies are obtained based on water-filling and minimum mean square error solutions. Simulation results show that the proposed scheme for UPAs performs very close to the optimal digital beamforming scheme.

**Index Terms**—THz communications, multiple-input multiple-output, hybrid beamforming, OFDM.

## I. INTRODUCTION

The traditional sub-6 GHz wireless systems exploit rich scattering environments to increase spectral efficiency with multiple-input multiple-output (MIMO) channel access. Looking toward the future millimeter wave (mmWave, 30–300 GHz) and THz frequency (>300 GHz) systems, the spectral efficiency is no longer the key to large data rates because very large spectral resources are available therein [1]. At these frequencies, large antenna arrays, which are also known as mmWave band massive MIMO (mMIMO), are seen as one of the key enabling technologies to provide flexible beamforming along with enough gain to counter the large channel losses [2]–[5]. Therefore, the focus on mmWave and above is often more on the phased array operation in order to maximize the channel gain rather than enabling multi-stream operation as in the actual MIMO communications. Regardless of the channel sparsity, the high frequencies have been shown to support multiple propagation paths in most cases to theoretically enable multi-stream communications as well [6].

The mMIMO antennas or phased arrays tend to be uniform linear or planar arrays (ULA or UPA, respectively). Due to their simplicity, ULA antennas are often assumed in the theoretical beamforming-related literature. Regardless of ULA or

UPA, the maximum achievable gain of such antenna groups are equal to the number of the antenna elements in a fully coherent summation of energy at the desired direction/location. The UPA antennas, however, give further flexibility by allowing beamforming in all cardinal directions. Thus, UPAs have one extra degree of freedom compared to the ULA antennas. As the required antenna structures in mMIMO systems are large, many of the recent papers have been focusing on hybrid beamforming (HBF) architectures for more energy and cost-efficient solutions by maintaining high antenna number while reducing the number of radio frequency (RF) chains by combining analog and digital processing [4], [5], [7]–[13].

In hybrid beamforming schemes designed for the conventional narrowband systems, the beams generated by the analog beamformer are aligned with the physical directions of channel path components to achieve the full array gain [8], [9], [14], [15]. However, in wideband systems, the beams at different subcarrier frequencies point to different physical directions because of the frequency-independent phase shifters in the analog beamforming network, causing a loss in array gain. This phenomenon is called the *beam squint effect*, whose severity significantly depends on the array size and bandwidth [14], [16]. Overcoming this effect becomes one of the important goals in HBF design in wideband mmWave mMIMO systems [14], [16]–[21]. Specifically, to deal with the array gain loss incurred by this effect, while efficient solutions to hybrid precoders in wideband mmWave MIMO systems are proposed in [17] and [18], Cai *et al.* [19] focus on designing the codebook with more beams and improved minimum array gains. Furthermore, Liu *et al.* [20] present a space-time block coding-based HBF scheme to compensate for beam squint. This design maximizes the average beam gain within the bandwidth while minimizing the average beam gain outside the bandwidth. However, most of these methods are proposed for mmWave systems, wherein the beam squint effect is not very severe. It has become quite an important phenomenon lately due to the popularity of wideband high frequencies in THz communications where even small mobility applications require antenna array solutions to offer gain and beam steering. An efficient way to tackle the severe beam squint is to apply true-time-delays (TTDs) between the RF chains and phase shift network to counteract the physical propagation delay [21]. These structures are often referred to as TTD-based

array processing [21] or delay-phase precoding [14], [16], [22], which have been shown to mitigate the beam squint very efficiently. However, while effective, these solutions come with the cost of increased hardware and computational complexity as well as power consumption.

Although the beam squint effect has been analyzed in [14], [16], [19], all of these works consider ULAs, while UPAs are usually deployed in practical antenna arrays. This is even more so in the wideband mMIMO systems with thousands of antenna elements. With a different array structure, the impact of beam squint on UPAs may be different, even with the same number of antennas and bandwidth. In this regard, quantifying the severity of this effect is important for HBF design [21]. In this paper, we show that the beam squint has an impact on both UPA and ULA structures through a comprehensive analysis and comparison. However it is much less severe for the UPA. This implies that with a reasonable and acceptable loss in array gain due to beam squint in UPAs, conventional HBF architectures without TTDs can in some cases maintain the low hardware cost and power consumption while providing satisfactory performance. We then propose an efficient HBF scheme wherein the analog beamformer is designed based only on the center frequency, while the digital beamformers are obtained dedicated for each subcarrier. It is shown in the numerical results that the proposed HBF scheme performs close to the optimal digital beamformer. In contrast, with the same number of antennas and bandwidth, the ULA suffers from a significant performance loss. It should be noticed that we assume herein an uplink scenario where the mobile station (MS) employs fully digital beamforming architecture, whereas the base station (BS) is equipped with an HBF structure. Therefore, the beam squint analysis is performed for the receive array. However, since beam squint does not depend on the link direction, the analysis herein applies to any wideband system utilizing large arrays.

## II. SYSTEM MODEL

We consider the uplink of a point-to-point THz massive MIMO-OFDM system, where the MS and BS are equipped with  $N_t$  and  $N_r$  antennas, respectively. The MS sends signal vector  $\mathbf{s}[k] \in \mathbb{C}^{N_s \times 1}$  of  $N_s$  data streams to the BS via subcarrier  $k$ , with  $\mathbb{E}\{\mathbf{s}[k]\mathbf{s}[k]^H\} = \mathbf{I}_{N_s}$ ,  $k = 1, \dots, K$ . To process  $\mathbf{s}[k]$ , we assume that a fully digital baseband precoder  $\mathbf{F}[k] \in \mathbb{C}^{N_t \times N_s}$  is used at the MS due to its small number of antennas,  $\|\mathbf{F}[k]\|_F^2 \leq P_{\text{MS}}[k]$ . In contrast, at the BS, the analog combiner  $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_r \times N_{\text{RF}}}$  and digital baseband combiner  $\mathbf{W}_{\text{BB}}[k] \in \mathbb{C}^{N_{\text{RF}} \times N_s}$  are employed. Here,  $N_{\text{RF}}$  is the number of RF chains at the BS,  $N_s \leq N_{\text{RF}} \leq N_r$ . Note that  $\mathbf{W}_{\text{RF}}$  is the frequency-flat combining matrix whose entries have constant amplitude but variable phase shifts. The post-processed signal at the BS is expressed as

$$\mathbf{y}[k] = \mathbf{W}_{\text{BB}}^H[k] \mathbf{W}_{\text{RF}}^H \mathbf{H}[k] \mathbf{F}[k] \mathbf{s}[k] + \mathbf{W}_{\text{BB}}^H[k] \mathbf{W}_{\text{RF}}^H \mathbf{n}[k], \quad (1)$$

where  $\mathbf{n}[k] \in \mathbb{C}^{N_r \times 1}$  is an additive white Gaussian noise (AWGN) vector with elements distributed as  $\mathcal{CN}(0, \sigma_n^2)$ , and  $\mathbf{H}[k]$  is the channel matrix at subcarrier  $k$ .

In THz communications, the high free-space pathloss leads to limited spatial selectivity and scattering. At the same time, the channel is highly correlated due to the large tightly-packed antenna arrays [9]. Therefore, to accurately capture the mathematical structure of the channel, we adopt the widely used extended Saleh-Valenzuela model and express  $\mathbf{H}[k]$  at frequency  $f_k = f_c + \frac{B(2k-1-K)}{2K}$  as [14], [15]

$$\mathbf{H}[k] = \sum_{\ell=1}^L \alpha_{\ell} e^{-j2\pi f_k \tau_{\ell}} \mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_k) \mathbf{a}_t(\theta_{\ell}^t, \phi_{\ell}^t, f_k)^H. \quad (2)$$

Here,  $L$  is the number of propagation paths,  $\alpha_{\ell}$  and  $\tau_{\ell}$  are the complex gain and time-of-arrival (ToA) of the  $\ell$ th path. Note that in (2), we assume the deployments of UPAs at both sides. Specifically, the BS is equipped with a UPA of size  $(N_{r,h} \times N_{r,v})$ , where  $N_{r,h}$  and  $N_{r,v}$  are the number of antennas in each row and column of the UPA, respectively, and  $N_{r,h} N_{r,v} = N_r$ . Assuming half-wavelength antenna spacing at the BS,  $\mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_k) \in \mathbb{C}^{N_r \times 1}$  is then given by [14], [15]

$$\mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_k) = \frac{1}{\sqrt{N_r}} \begin{bmatrix} 1, \dots, e^{j\pi \frac{f_k}{f_c} (i_h \theta_{\ell}^r + i_v \phi_{\ell}^r)}, \\ \dots, e^{j\pi \frac{f_k}{f_c} ((N_{r,h}-1)\theta_{\ell}^r + (N_{r,v}-1)\phi_{\ell}^r)} \end{bmatrix}^T, \quad (3)$$

where  $\theta_{\ell}^r \triangleq \sin(\bar{\theta}_{\ell}^r) \sin(\tilde{\theta}_{\ell}^r)$  and  $\phi_{\ell}^r \triangleq \cos(\tilde{\theta}_{\ell}^r)$ . Here,  $\bar{\theta}_{\ell}^r(\tilde{\theta}_{\ell}^r)$  and  $\bar{\theta}_{\ell}^t(\tilde{\theta}_{\ell}^t)$  represent the azimuth (elevation) angles of arrival and departure (AoAs and AoDs) of the  $\ell$ th path, and  $0 \leq i_h < N_{r,h}$ ,  $0 \leq i_v < N_{r,v}$  are the antenna indices on the horizontal and vertical dimensions, respectively. The array response vector  $\mathbf{a}_t(\theta_{\ell}^t, \phi_{\ell}^t, f_k)$  at the MS can be modeled in a similar fashion.

## III. THE BEAM SQUINT EFFECT

Due to the ultra-high bandwidth employed in THz communications, the propagation delay across the array at the BS can exceed the sampling period. This causes the *beam squint effect*, in which the direction-of-arrival (DoA) varies across the OFDM subcarriers, and the array gain becomes frequency-selective [22]. This effect has been analyzed in several works for ULAs, such as [14], [16], [19]. However, those analyses may be invalid or inaccurate for UPAs in which the array structure is different. In this section, we first characterize the normalized array gain of the analog combiner, which allows us to evaluate the severity of the beam squint effect in UPAs and compare the effect in the two array structures.

To investigate this effect, we first consider the  $\ell$ th path component of a narrowband channel at the center frequency  $f_c$  represented by  $\alpha_{\ell} \mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_c) \mathbf{a}_t(\theta_{\ell}^t, \phi_{\ell}^t, f_c)^H$ . The normalized array gain at the BS achieved by analog combining vector  $\mathbf{w}_n$  (i.e., the  $n$ th column of  $\mathbf{W}_{\text{RF}}$ ) is given as

$$\eta(\mathbf{w}_n, \theta_{\ell}^r, \phi_{\ell}^r, f_c) = |\mathbf{w}_n^H \mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_c)|.$$

It has been shown in [9] that

$$\mathbf{w}_n = \mathbf{a}_r(\theta_{\ell}^r, \phi_{\ell}^r, f_c) \quad (4)$$

is the near-optimal solution minimizing the Euclidean distance between the optimal (unconstrained) beamformer and the

hybrid one, and it achieves the highest array gain at  $f_c$  as  $\eta^*(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_c) = 1$  [14]. Thus, the highest array gain can be attained over the entire band, if it is narrow enough, because  $f_k \approx f_c, \forall k$ .

In contrast, the difference between subcarriers and the center frequency can be significant and cannot be ignored in wideband systems, where the spatial directions varies across the path components and subcarriers, i.e.,  $(\theta_{\ell k}^r, \phi_{\ell k}^r) = \left(\frac{f_k}{f_c} \theta_\ell^r, \frac{f_k}{f_c} \phi_\ell^r\right), \forall \ell, k$ , as seen in (3). In this case, the analog beamformer in (4) at the BS cannot cover the MS with its mainlobe, causing beam squint effect.

**Lemma 1:** For the UPA, when physical direction  $(\theta_{\ell k}^r, \phi_{\ell k}^r)$  at frequency  $f_k$  satisfies

$$\left| \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right| \geq \frac{2}{N_{r,h}} \text{ and } \left| \frac{f_k}{f_c} \phi_{\ell k}^r - \phi_\ell^r \right| \geq \frac{2}{N_{r,v}}, \quad (5)$$

the beam generated at the BS with the analog combiner  $\mathbf{w}_n$  in (4) cannot cover the MS with its mainlobe. Under this condition, the maximum normalized array gain achieved by  $\mathbf{w}_n$  in DoA  $(\theta_\ell^r, \phi_\ell^r)$  is limited as

$$\eta(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k) \leq \frac{1}{N_r \sin \frac{3\pi}{2N_{r,h}} \sin \frac{3\pi}{2N_{r,v}}}. \quad (6)$$

*Proof:* The normalized array gain achieved by the analog beamformer  $\mathbf{w}_n = \mathbf{a}_r(\theta_\ell^r, \phi_\ell^r, f_c)$  in (4) for DoA  $(\theta_{\ell k}^r, \phi_{\ell k}^r)$  (associated with array response vector  $\mathbf{a}_r(\theta_{\ell k}^r, \phi_{\ell k}^r, f_k)$ ), can be computed as

$$\begin{aligned} \eta(\mathbf{w}_n, \theta_{\ell k}^r, \phi_{\ell k}^r, f_k) &= \left| \mathbf{w}_n^H \mathbf{a}_r(\theta_{\ell k}^r, \phi_{\ell k}^r, f_k) \right| \\ &= \left| \mathbf{a}_r^H(\theta_\ell^r, \phi_\ell^r, f_c) \mathbf{a}_r(\theta_{\ell k}^r, \phi_{\ell k}^r, f_k) \right|. \end{aligned} \quad (7)$$

Here,  $\mathbf{a}_r(\theta_{\ell k}^r, \phi_{\ell k}^r, f_k)$  can be decomposed as  $\mathbf{a}_r(\theta_{\ell k}^r, \phi_{\ell k}^r, f_k) = \mathbf{a}_{r,h}(\theta_{\ell k}^r, f_k) \otimes \mathbf{a}_{r,v}(\phi_{\ell k}^r, f_k)$ , where

$$\begin{aligned} \mathbf{a}_{r,h}(\theta_{\ell k}^r, f_k) &= \frac{1}{\sqrt{N_{r,h}}} \left[ 1, \dots, e^{j\pi \frac{f_k}{f_c} (N_{r,h}-1) \theta_{\ell k}^r} \right]^T, \\ \mathbf{a}_{r,v}(\phi_{\ell k}^r, f_k) &= \frac{1}{\sqrt{N_{r,v}}} \left[ 1, \dots, e^{j\pi \frac{f_k}{f_c} (N_{r,v}-1) \phi_{\ell k}^r} \right]^T. \end{aligned}$$

As a result, we can rewrite (7) as

$$\begin{aligned} \eta(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k) &\stackrel{(a)}{=} \left| (\mathbf{a}_{r,h}^H(\theta_\ell^r, f_c) \otimes \mathbf{a}_{r,v}^H(\phi_\ell^r, f_c)) (\mathbf{a}_{r,h}(\theta_{\ell k}^r, f_k) \otimes \mathbf{a}_{r,v}(\phi_{\ell k}^r, f_k)) \right| \\ &\stackrel{(b)}{=} \left| \mathbf{a}_{r,h}^H(\theta_\ell^r, f_c) \mathbf{a}_{r,h}(\theta_{\ell k}^r, f_k) \right| \left| \mathbf{a}_{r,v}^H(\phi_\ell^r, f_c) \mathbf{a}_{r,v}(\phi_{\ell k}^r, f_k) \right|, \end{aligned} \quad (8)$$

where equality (a) follows the property  $(\mathbf{a}^H \otimes \mathbf{b}^H)(\mathbf{c} \otimes \mathbf{d}) = (\mathbf{a}^H \mathbf{c}) \otimes (\mathbf{b}^H \mathbf{d})$ , and (b) is obtained by the fact that both  $\mathbf{a}^H \mathbf{c}$  and  $\mathbf{b}^H \mathbf{d}$  are scalars. We further have

$$\begin{aligned} \left| \mathbf{a}_{r,h}^H(\theta_\ell^r, f_c) \mathbf{a}_{r,h}(\theta_{\ell k}^r, f_k) \right| &= \frac{1}{N_{r,h}} \left| \sum_{i_h=0}^{N_{r,h}-1} e^{j\pi i_h \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right)} \right| \\ &\stackrel{(c)}{=} \left| \frac{\sin \frac{N_{r,h}\pi}{2} \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right)}{N_{r,h} \sin \frac{\pi}{2} \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right)} e^{-j\pi \frac{(N_{r,h}-1)}{2} \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right)} \right| \\ &\stackrel{(d)}{=} \frac{1}{N_{r,h}} \left| \text{D}_{N_{r,h}} \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right) \right|, \end{aligned}$$

where in equality (c), we have used property  $\sum_{i=0}^{N-1} e^{j\pi i x} = \frac{\sin \frac{N\pi}{2} x}{\sin \frac{\pi}{2} x} e^{-j\pi \frac{(N-1)}{2} x}$ , and in equality (d),  $\text{D}_N(x) = (\sin \frac{N\pi}{2} x) / (\sin \frac{\pi}{2} x)$  is the Dirichlet sinc function [14]. By a similar expansion of the second absolute value in (8) and by recalling that  $N_r = N_{r,h} N_{r,v}$ , we obtain

$$\begin{aligned} \eta(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k) &= \frac{1}{N_r} \left| \text{D}_{N_{r,h}} \left( \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right) \right| \left| \text{D}_{N_{r,v}} \left( \frac{f_k}{f_c} \phi_{\ell k}^r - \phi_\ell^r \right) \right|. \end{aligned} \quad (9)$$

From (9), the following observations are made:

- Because  $|\text{D}_N(x)| = 0$  at  $|x| = \{\frac{N}{2}, N, \frac{3N}{2}, 2N, \dots\}$ , we have  $\eta(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k) = 0$  when  $\left| \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right| = q \frac{N_{r,h}}{2}$  or  $\left| \frac{f_k}{f_c} \phi_{\ell k}^r - \phi_\ell^r \right| = q \frac{N_{r,v}}{2}$ , with  $q = 1, 2, \dots$
- The mainlobe of the normalized array gain  $\eta(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k)$  covers the direction  $(\theta_{\ell k}^r, \phi_{\ell k}^r)$  satisfying

$$\left| \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right| < \frac{2}{N_{r,h}} \text{ and } \left| \frac{f_k}{f_c} \phi_{\ell k}^r - \phi_\ell^r \right| < \frac{2}{N_{r,v}}, \quad (10)$$

and inside the mainlobe, the maximum normalized array gain is  $\eta^*(\mathbf{w}_n, \theta_\ell^r, \phi_\ell^r, f_k) = 1$  achieved at  $(\theta_\ell^r, \phi_\ell^r) = \left(\frac{f_k}{f_c} \theta_{\ell k}^r, \frac{f_k}{f_c} \phi_{\ell k}^r\right)$ .

- We have that  $|\text{D}_N(x)| \leq \frac{1}{\sin(\frac{3\pi}{2N})}$  for  $|x| \geq \frac{2}{N}$ , i.e.,  $\frac{1}{\sin(\frac{3\pi}{2N})}$  is the maximum value of  $|\text{D}_N(x)|$  locating out of the mainlobe of  $|\text{D}_N(x)|$ . Equivalently, when  $\left| \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right| \geq \frac{2}{N_{r,h}}$  and  $\left| \frac{f_k}{f_c} \phi_{\ell k}^r - \phi_\ell^r \right| \geq \frac{2}{N_{r,v}}$ , we obtain the equality in (6).  $\square$

**Remark 1:** It is observed from Lemma 1 that as the bandwidth  $B$  and/or the number of antennas  $N_r$  increases, the condition (5) is easier to be satisfied. Equivalently, the beam squint effect becomes more severe in a system employing a wider band/or and a larger number of antennas. Indeed, it is also seen from (10) that as  $N_r$  and/or  $B$  increase, the main lobe becomes sharper, making beam squint to occur. Furthermore, as  $N_r \rightarrow \infty$ , the normalized array gain outside the main lobe tends to zero. These are due to the increasing deviation of the physical derivation and the narrowing bandwidth as  $B$  and  $N_r$  increases [14].

We note that the observations in Remark 1 are also made for receivers equipped with ULAs. Specifically, by a similar derivation as in [14], one can show that with a ULA, a beam generated by an analog combiner cannot cover the MS with its mainlobe if

$$\left| \frac{f_k}{f_c} \theta_{\ell k}^r - \theta_\ell^r \right| \geq \frac{2}{N_r}, \quad (11)$$

where  $\theta_\ell^r$  is the AoA. By comparing the conditions (5) and (11), we remark the following observation.

**Remark 2:** Due to the fact that  $N_{r,h}, N_{r,v} < N_r$ , it is clear that the condition (11) is easier to be satisfied than (5). Equivalently, comparing the UPA and ULA with the same number of antenna elements and system bandwidth, the former is much more robust to beam squint effect compared to the latter. Therefore, UPAs are favorable in wideband systems for

a smaller loss in array gain.

The comparison in Remark 2 unveils important insights into the design of the HBF scheme in wideband systems. First, it shows that the beam squint in such systems may not be severe if UPAs of appropriate sizes are deployed, which are more practical than ULAs of the same size due to smaller size. In this case, the delay-phase precoding scheme may not be required because the performance improvement offered by the time delay network is small while its additional power consumption and hardware cost are considerable. Second, the challenge of designing the frequency-independent analog beamformer can be relaxed. More specifically, when the beam squint effect is acceptable, conventional well-developed HBF algorithms for narrowband systems can be applied to solve the analog beamformer by choosing the center frequency. In the next section, we present an HBF design based on this fact.

#### IV. PROPOSED HBF DESIGN

##### A. Problem Formulation

Based on (1), the achievable spectral efficiency (in bits/s/Hz) when the transmitted symbols follow a Gaussian distribution is given by

$$R[k] = \log_2 \det \left( \mathbf{I}_{N_s} + \mathbf{R}_n[k]^{-1} \mathbf{W}_{\text{BB}}[k]^H \mathbf{W}_{\text{RF}}^H \mathbf{H}[k] \mathbf{F}[k] \right. \\ \left. \times \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}[k] \right), \quad (12)$$

where  $\mathbf{R}_n[k] = \sigma_n^2 \mathbf{W}_{\text{BB}}[k]^H \mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}[k]$  is the noise covariance matrix. We aim at maximizing the total spectral efficiencies over all subcarriers, i.e.,  $\sum_{k=1}^K R[k]$ . This problem can be formulated as

$$\begin{aligned} & \underset{\{\mathbf{W}_{\text{BB}}[k], \mathbf{F}[k]\}_{k=1}^K, \mathbf{W}_{\text{RF}}}{\text{maximize}} && \sum_{k=1}^K R[k] \\ & \text{subject to} && \|\mathbf{F}[k]\|_{\mathcal{F}}^2 \leq P_{\text{MS}}[k], \end{aligned} \quad (13a)$$

$$\text{subject to} \quad \|\mathbf{F}[k]\|_{\mathcal{F}}^2 \leq P_{\text{MS}}[k], \quad (13b)$$

$$|w_{mn}| = \frac{1}{\sqrt{N_r}}, \quad \forall m, n. \quad (13c)$$

where  $w_{mn}$  is the  $(m, n)$ th entry of  $\mathbf{W}_{\text{RF}}$ . Constraint (13b) is to ensure the per-subcarrier power constraint, while (13c) is the constant modulus constraint of analog beamforming coefficients. It is challenging to solve problem (13) since constraint (13c) is non-convex. Furthermore, beamforming matrices  $\{\mathbf{W}_{\text{BB}}[k], \mathbf{W}_{\text{RF}}, \mathbf{F}[k]\}$  are highly coupled in the objective function, which is neither convex nor concave. Thus, problem (13) is intractable to solve. To tackle these challenges, we decouple the problem (13) to the designs of digital precoder  $\{\mathbf{F}[k]\}$ , analog combiner  $\mathbf{W}_{\text{RF}}$ , and digital combiner  $\mathbf{W}_{\text{BB}}[k]$ , which are respectively elaborated next.

##### B. Proposed Design

1) *Transmit Precoding Design:* We first aim at designing the transmit digital precoding matrices  $\{\mathbf{F}[k]\}$ . The problem can be formulated as

$$\begin{aligned} & \underset{\{\mathbf{F}[k]\}}{\text{maximize}} && \sum_{k=1}^K \log_2 \det \left( \mathbf{I}_{N_r} + \frac{1}{\sigma^2} \mathbf{H}[k] \mathbf{F}[k] \mathbf{F}[k]^H \mathbf{H}[k]^H \right) \\ & \text{subject to} && (13b). \end{aligned}$$

The optimal precoder  $\mathbf{F}[k]$  at subcarrier  $k$  admits a well-known water-filling solution as [5]

$$\mathbf{F}[k] = \mathbf{U}[k] \mathbf{\Sigma}[k], \quad \forall k, \quad (15)$$

where  $\mathbf{U}[k]$  is the matrix whose columns are the  $N_s$  right-singular vectors corresponding to the  $N_s$  largest singular values  $\{\lambda_1, \dots, \lambda_{N_s}\}$  of  $\mathbf{H}[k]$ , and  $\mathbf{\Sigma}[k] = \text{diag}\{\sqrt{p_1[k]}, \dots, \sqrt{p_{N_s}[k]}\}$  with  $p_i[k]$  being the power amount allocated to the  $i$ th data stream. Here,  $p_i[k] = \max\left\{\frac{1}{\bar{p}[k]} - \frac{\sigma^2}{\lambda_i}, 0\right\}$ , where  $\bar{p}[k]$  satisfy  $\sum_{i=1}^{N_s} p_i[k] = P_{\text{MS}}[k]$ .

2) *Hybrid Combining Design:* The design of analog and digital combiners, i.e.,  $\mathbf{W}_{\text{RF}}$  and  $\{\mathbf{W}_{\text{BB}}[k]\}$ , respectively, are decoupled by first solving  $\mathbf{W}_{\text{RF}}$  with given  $\{\mathbf{W}_{\text{BB}}[k]\}$ , and then,  $\{\mathbf{W}_{\text{BB}}[k]\}$  are obtained by the well-known optimal minimum mean square error (MMSE) solution [8]. Specifically, the analog combiner  $\mathbf{W}_{\text{RF}}$  can be solved in the following problem

$$\begin{aligned} & \underset{\mathbf{W}_{\text{RF}}}{\text{maximize}} && \sum_{k=1}^K \log_2 \det \left( \mathbf{I}_{N_r} + \frac{1}{\sigma^2} (\mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}})^{-1} \right. \\ & && \left. \times \mathbf{W}_{\text{RF}}^H \tilde{\mathbf{H}}[k] \mathbf{W}_{\text{RF}} \right) \end{aligned}$$

subject to (13c),

where  $\tilde{\mathbf{H}}[k] = \mathbf{H}[k] \mathbf{F}[k] \mathbf{F}[k]^H \mathbf{H}[k]^H \in \mathbb{C}^{N_r \times N_r}$ . In THz massive MIMO system, because  $N_r$  is very large, we have  $\mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}} \approx \mathbf{I}_{N_{\text{RF}}}$ . Thus, we can rewrite the above problem as

$$\begin{aligned} & \underset{\mathbf{W}_{\text{RF}}}{\text{maximize}} && \bar{R} \triangleq \sum_{k=1}^K \log_2 \det \left( \mathbf{I}_{N_r} + \frac{1}{\sigma^2} \mathbf{W}_{\text{RF}}^H \tilde{\mathbf{H}}[k] \mathbf{W}_{\text{RF}} \right) \\ & \text{subject to} && (13c). \end{aligned} \quad (16)$$

$$\text{subject to (13c).} \quad (17)$$

The objective function in (16) can be re-expressed as [8]

$$\begin{aligned} \bar{R} &= \sum_{k=1}^K \log_2 \det (\mathbf{C}_n[k]) \\ &+ \log_2 (2\Re(w_{mn}^* \nu_{mn}[k]) + \zeta_{mn}[k] + 1), \end{aligned} \quad (18)$$

where

$$\mathbf{C}_n[k] = \mathbf{I}_{N_{\text{RF}}-1} + \frac{1}{\sigma^2} \left( \mathbf{W}_{\text{RF}}^{(n)} \right)^H \tilde{\mathbf{H}}[k] \mathbf{W}_{\text{RF}}^{(n)}, \quad (19)$$

with  $\mathbf{W}_{\text{RF}}^{(n)} \in \mathbb{C}^{N_r \times (N_{\text{RF}}-1)}$  being the sub-matrix  $\mathbf{W}_{\text{RF}}$  with the  $n$ th column removed, and

$$\nu_{mn}[k] = \sum_{\ell \neq m}^{N_r} g_{m\ell}^{(n)}[k] w_{\ell n}, \quad (20)$$

$$\zeta_{mn}[k] = g_{mm}^{(n)}[k] + 2\Re \left( \sum_{i,j \neq m}^{N_r} w_{in}^* g_{ij}^{(n)}[k] w_{nj} \right), \quad (21)$$

with  $g_{m\ell}^{(n)}[k]$  being the  $(m, \ell)$ th element of matrix

$$\mathbf{G}_n[k] = \frac{1}{\sigma^2} \tilde{\mathbf{H}}[k] - \frac{1}{\sigma^4} \tilde{\mathbf{H}}[k] \mathbf{W}_{\text{RF}}^{(n)} \mathbf{C}_n[k]^{-1} (\mathbf{W}_{\text{RF}}^{(n)})^H \tilde{\mathbf{H}}[k], \quad (22)$$

and  $\Re(\cdot)$  represents the real part of a complex number. It is observed from (19)–(22) that in (18),  $\mathbf{C}_n[k]$ ,  $\nu_{mn}[k]$ , and  $\zeta_{mn}[k]$  are independent of  $w_{mn}$ . Therefore, if all the coefficients  $\{w_{ij}\}_{i \neq m, j \neq n}$  are fixed, the optimal solution to  $w_{mn}$  at the center frequency can be given as

$$w_{mn} = \begin{cases} \frac{1}{\sqrt{N_r}}, & \text{if } \nu_{mn}[\bar{k}] = 0, \\ \frac{\nu_{mn}[\bar{k}]}{\sqrt{N_r}|\nu_{mn}[\bar{k}]|}, & \text{otherwise.} \end{cases} \quad (23)$$

where  $\bar{k} = \lfloor \frac{K}{2} \rfloor$ . Finally, the optimal digital combiner can be obtained as the MMSE solution, i.e.,

$$\mathbf{W}_{\text{BB}}[k] = \mathbf{Q}[k]^{-1} \mathbf{W}_{\text{RF}}^H \mathbf{H}[k] \mathbf{F}[k], \quad (24)$$

with  $\mathbf{Q}[k] = \mathbf{W}_{\text{RF}}^H \mathbf{H}[k] \mathbf{F}[k] \mathbf{F}[k]^H \mathbf{H}[k]^H \mathbf{W}_{\text{RF}} + \sigma^2 \mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}}$  [8].

We summarize the overall algorithm for solving problem (13) in Algorithm 1. Specifically, we first find the digital transmit precoders at all  $K$  subcarrier frequencies, i.e.,  $\{\mathbf{F}[k]\}_{k=1}^K$ , in step 1 using the water-filling solution in (15). Then, the coefficients of the analog combiner are alternatively solved in steps 3–11, and finally, the digital combiner is obtained as the MMSE solution in step 12. We note here that the analog combiner is obtained at the center frequency associated with  $\bar{k} = \lfloor \frac{K}{2} \rfloor$  based on the fact that the beam squint effect is less severe in UPAs compared to ULAs. This HBF scheme is suboptimal when the impact of beam squint is severe. However, its performance loss is generally small when UPAs are deployed, and especially, it avoids the increased power consumption and hardware cost of the delay-phase beamforming scheme in [14], [16].

---

#### Algorithm 1 HBF Design to Solve Problem (13)

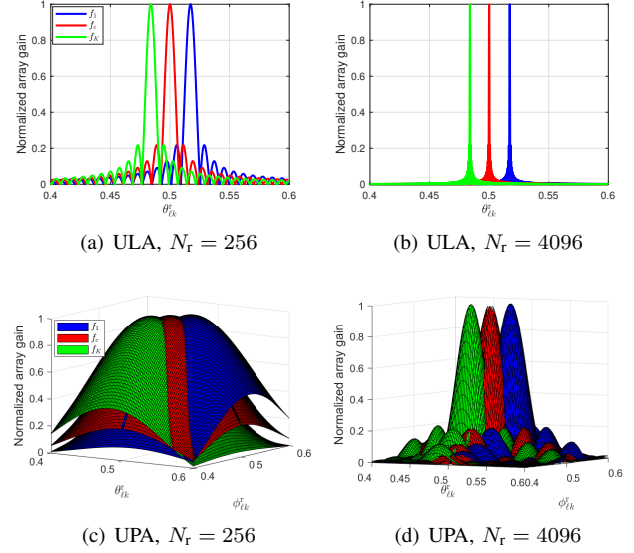
---

- 1: Obtain digital transmit precoders  $\{\mathbf{F}[k]\}_{k=1}^K$  based on (15).
  - 2: **Initialization:** Set all the analog combining coefficients to ones, i.e.,  $w_{mn} = 1, m = 1, \dots, N_r, n = 1, \dots, N_{\text{RF}}$ . Set  $\bar{k} = \lfloor \frac{K}{2} \rfloor$ .
  - 3: **repeat**
  - 4:   **for**  $n = 1 \rightarrow N_{\text{RF}}$  **do**
  - 5:     Compute  $\mathbf{C}_n[\bar{k}]$  and  $\mathbf{G}_n[\bar{k}]$  based on (19) and (22), respectively.
  - 6:   **for**  $m = 1 \rightarrow N_r$  **do**
  - 7:     Compute  $\nu_{mn}[\bar{k}]$  based on (20).
  - 8:     Obtain  $w_{mn}$  based on (23).
  - 9:   **end for**
  - 10: **end for**
  - 11: **until** convergence.
  - 12: Obtain digital combiners  $\{\mathbf{W}_{\text{BB}}[k]\}_{k=1}^K$  based on (24).
- 

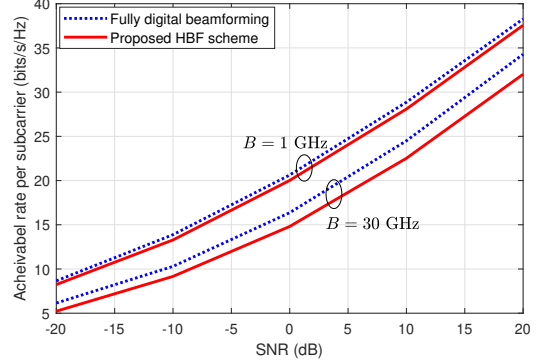
## V. SIMULATION RESULTS

We herein provide numerical results to verify the analysis. In the simulations, we set  $N_t = N_{\text{RF}} = N_s = \{4, 16\}$ , and  $N_{r,h} = N_{r,v} = \sqrt{N_r}$  is assumed for UPAs. When generating the channels, we assume  $L = 4$ ,  $\theta_\ell^r, \theta_\ell^t \sim \mathcal{U}(-\pi, \pi)$ , and  $\bar{\theta}_\ell^r, \bar{\theta}_\ell^t \sim \mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$ . The signal-to-noise ratio (SNR) is defined as  $\text{SNR} = \frac{\rho}{\sigma_n^2}$ , where  $P_{\text{MS}}[1] = \dots = P_{\text{MS}}[K] = \rho$ . The other parameters are detailed in each figure.

We first compare the beam squint effects that occur in wideband systems employing ULAs and ULAs in Fig. 1. It is



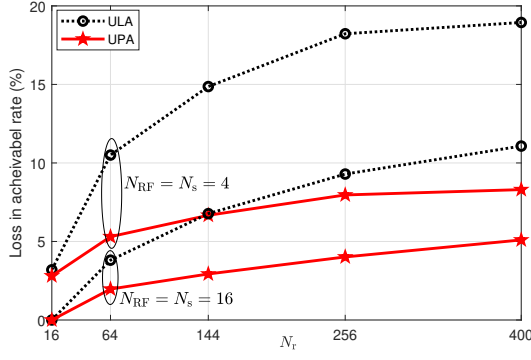
**Fig. 1.** Normalized array gain achieved by  $\mathbf{w}_n$  in ULAs/UPAs with  $f_c = 300$  GHz,  $B = 30$  GHz,  $K = 3$ ,  $\theta_\ell^r = 0.5$  for ULAs, and  $(\theta_\ell^r, \phi_\ell^r) = (0.5, 0.5)$  for UPAs.



**Fig. 2.** Achievable rate per subcarrier of wideband mMIMO systems employing UPAs with  $N_r = 256$ ,  $N_t = N_{\text{RF}} = N_s = 4$ ,  $K = 128$ , and  $B = \{1, 30\}$  GHz.

observed that as  $N_r$  and  $B$  are large, both the ULAs and UPAs suffer from the beam squint effect, but the severity is different between these two. Specifically, with the same number of antennas and bandwidth, the effect in the former is much more severe than that in the latter. For example, it is observed in Fig. 1(a) for the ULA that the main lobes of the normalized array gains at  $\{f_1, f_c, f_K\}$  are separated. In contrast, in Fig. 1(c) for the UPA, the main lobe at center frequency  $f_c$  can almost cover those at  $f_1$  and  $f_K$ . With  $N_r = 4096$ , the main lobes of ULAs become very sharp and totally separated, showing a very severe beam squint in the system. However, with the UPA, as seen in Fig. 1(d), the main lobe at  $f_c$  can still cover some parts of those at  $f_1$  and  $f_K$ . The observations in Fig. 1 imply that with a reasonably large array and bandwidth, the analog beamformer designed for the UPA at  $f_c$  can still perform well at other frequencies, as will be further demonstrated next.

In Fig. 2, we show the average achievable rates of the wideband systems employing UPAs for  $B = \{1, 30\}$  GHz. The benchmark scheme for comparison is the optimal fully digital beamforming scheme. Specifically, the digital combiner



**Fig. 3.** Performance loss due to beam squint effect of the proposed HBF scheme with respect to the fully digital beamforming with  $N_t \in [16, 400]$ ,  $K = 128$ , and  $B = 30$  GHz.

at frequency  $k$  is the matrix whose columns are the  $N_s$  left-singular vectors corresponding to the  $N_s$  largest singular value of  $\mathbf{H}[k]$ . It is seen that the proposed HBF scheme performs very close to the optimal fully digital beamformer although there exists a small performance loss at high  $B$ , i.e.,  $B = 30$  GHz, due to the beam squint effect.

The loss also becomes more significant as  $N_t$  increases, as seen in Fig. 3. This figure shows the performance loss of the proposed HBF scheme with respect to the optimal digital beamformer in both UPA and ULA systems. We consider  $N_{RF} = \{4, 16\}$  and  $N_t = \{16, 64, 144, 256, 400\}$ . Obviously, with  $N_{RF} = N_t = 16$ , i.e., the number of RF chains is equal to the number of antennas, both UPA and ULA systems have no performance loss. As  $N_t$  increases, both the UPA and ULA suffer from more significant performance degradation. However, the loss in the system employing UPAs is much smaller than that with ULAs. For example, with  $N_t = 400$  and  $N_t = N_{RF} = N_s = 4$ , the former has a loss of only 8% in the average achievable rate, while that of the latter is nearly 20%. The observations in Fig. 3 does not only justifies our findings in Remarks 1 and 2 but also shows that in some scenarios, the performance loss in beam squint can be small so that conventional HBF architecture without TTDs can be used to maintain low power consumption and hardware cost.

## VI. CONCLUSION

We have investigated the beam squint effect and HBF design in wideband THz massive MIMO-OFDM systems employing UPAs. Through the analytical results on the normalized array gains of the analog combiners in UPA and ULA systems, we show that UPAs still suffers from the beam squint effect; however, it is much less severe than that in ULAs. Based on this, we proposed a simple HBF scheme in which the analog combiner is designed only for the center frequency. The numerical results verified our analysis and show that the proposed HBF scheme performs very close to the optimal digital beamforming.

## ACKNOWLEDGEMENT

This research has been supported in part by the Academy of Finland, 6G Flagship program under Grant 346208, EERA

Project under grant 332362, Infotech Program funded by University of Oulu Graduate School, and Horizon 2020, European Union's Framework Programme for Research and Innovation, under grant 871464 (ARIADNE).

## REFERENCES

- [1] M. Latva-Aho and K. Leppänen, Eds., *Key drivers and research challenges for 6G ubiquitous wireless intelligence*, ser. 6G research visions. University of Oulu, Sep. 2019, no. 1.
- [2] A. L. Swindlehurst *et al.*, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, 2014.
- [3] E. G. Larsson *et al.*, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.
- [4] N. T. Nguyen and K. Lee, "Unequally sub-connected architecture for hybrid beamforming in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1127–1140, 2019.
- [5] N. T. Nguyen, K. Lee, and H. Dai, "Hybrid beamforming and adaptive RF chain activation for uplink cell-free millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, 2022.
- [6] P. Kyösti, M. F. De Guzman, K. Haneda, N. Tervo, and A. Pärssinen, "How many beams does sub-thz channel support?" *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 1, pp. 74–78, 2022.
- [7] S. H. Hong, J. Park, S.-J. Kim, and J. Choi, "Hybrid beamforming for intelligent reflecting surface aided millimeter wave mimo systems," *IEEE Trans. Commun.*, pp. 1–1, 2022.
- [8] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, 2016.
- [9] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.
- [10] M. Ma, N. T. Nguyen, and M. Juntti, "Closed-form hybrid beamforming solution for spectral efficiency upper bound maximization in mmwave mimo-ofdm systems," in *Proc. IEEE Veh. Technol. Conf.*, 2021, pp. 1–5.
- [11] —, "Switch-based hybrid beamforming for wideband multi-carrier communications," in *25th Int. ITG Workshop Smart Antennas*, 2021, pp. 1–6.
- [12] Q. *et al.*, "Intelligent radio signal processing: A survey," *IEEE Access*, vol. 9, pp. 83 818–83 850, 2021.
- [13] G. M. Gadiel, N. T. Nguyen, and K. Lee, "Dynamic unequally sub-connected hybrid beamforming architecture for massive mimo systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3469–3478, 2021.
- [14] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband thz massive mimo," *IEEE Transactions on Wireless Communications*, 2022.
- [15] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, 2016.
- [16] J. Tan and L. Dai, "Delay-phase precoding for THz massive MIMO with beam split," in *Proc. IEEE Global Commun. Conf.* IEEE, 2019, pp. 1–6.
- [17] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, 2017.
- [18] L. Kong, S. Han, and C. Yang, "Hybrid precoding with rate and coverage constraints for wideband massive mimo systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4634–4647, 2018.
- [19] M. *et al.*, "Effect of wideband beam squint on codebook design in phased-array wireless systems," in *Proc. IEEE Global Commun. Conf.*, 2016, pp. 1–6.
- [20] X. Liu and D. Qiao, "Space-time block coding-based beamforming for beam squint compensation," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 241–244, 2018.
- [21] H. Hashemi, T.-S. Chu, and J. Roderick, "Integrated true-time-delay-based ultra-wideband array processing," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 162–172, 2008.
- [22] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, 2021.