Joint Delay-Phase Precoding Under True-Time Delay Constraints in Wideband Sub-THz Hybrid Massive MIMO Systems

Dang Qua Nguyen, *Student Member, IEEE* and Taejoon Kim, *Senior Member, IEEE*

Abstract—In wideband sub-Terahertz (sub-THz) massive multiple-input multiple-output (MIMO) communication systems, the beam squint effect manifests as a substantial degradation in array gain. To mitigate the aforementioned beam squint effect, a hybrid precoding approach leveraging both true-time delay (TTD) and phase shifters (PS) has been proposed. However, existing methods operate under the assumption that the TTD device can generate any desired time delay value. These methods subsequently design the TTD precoder while fixing the PS precoder. This work presents a novel optimization framework for the joint TTD and PS precoder design, incorporating realistic time delay constraints for each TTD device. Unlike previous methods, our framework does not rely on the unbounded time delay assumption and optimizes the TTD and PS values jointly to cope with the practical limitations. Furthermore, within the context of our proposed framework, we mathematically determine the minimum number of TTD devices necessitated to achieve a pre-determined target array gain. Simulations confirm the proposed approach exhibits performance improvement, guarantees array gain, and achieves computational efficiency.

Index Terms—Wideband sub-THz massive mulitple-input mulitple-output (MIMO), beam squint effect, hybrid precoding, phase shifter (PS), true-time delay (TTD), and joint TTD and PS precoding.

I. INTRODUCTION

Sub-Terahertz (sub-THz) band (90-300 GHz) communication is a potential technology for the sixth generation (6G)-&-beyond wireless systems, which are expected to support various high data rate applications such as augmented reality/virtual reality (AR/VR), eHealth, and holographic telepresence [3]. The sub-THz band offers abundant spectrum resources with tens of gigahertz (GHz) bandwidths [4]–[7], compared to the current millimeter-wave (mmWave) bands in the 5G specifications [8] that utilize a few GHz bandwidths. This enables the possibility of achieving the data rates on the orders of 10 to 100 Gbps using the existing

digital modulation techniques in sub-THz frequencies [5], [9]. However, the sub-THz communication faces several challenges such as high path losses, large power consumption, and inter-symbol interference. To overcome these challenges, hybrid massive multiple-input multiple-output (MIMO) and orthogonal frequency-division multiplexing (OFDM) technologies have been widely investigated recently [10]–[17].

Massive MIMO technology has received significant attention previously [18]. Strong theoretical analyses have justified the use of a very large number of antennas at the base station [18]–[20]. This has raised significant interest in massive MIMO systems at sub-6 GHz [18]–[20] and mmWave frequencies [21]–[23]. The underlying assumption behind [18]–[23] was, however, narrowband. Wideband high-frequency massive MIMO OFDM systems may experience significant array gain loss across different OFDM subcarriers due to the spatial wideband effect [24]–[30], also known as the beam squint effect. Beam squint refers to the phenomenon in which the deviation occurs in the spatial direction of each OFDM subcarrier when the wideband OFDM is used in a very large antenna array system. This implies that beam squint can cause severe degradation of the achievable rate, which potentially demotivates the use of OFDM in the wideband sub-THz massive MIMO systems. Therefore, an efficient beam squint compensation is essential for the realization of wideband sub-THz massive MIMO communications.

A. Related Works

Previous works on wideband mmWave massive MIMO [24], [25] addressed the beam squint by designing the beamforming weights to generate adaptive-beamwidth beams that cover the squinted angles. However, these techniques are not applicable to the sub-THz bands due to the extremely narrow pencil beam requirement imposed by the much higher carrier frequencies [31]–[33]. Furthermore, several hybrid precoding approaches [16], [17] have attempted to address beam squint through digital signal processing techniques. These approaches design the digital precoder by projecting it onto the subspace spanned by the analog beamforming vectors. However, this approach can only partially compensate for the beam squint due to the low-rank characteristic of the analog beamforming.

The beam squint effect of phased array antennas has been independently studied in the radar community (e.g., see [34]–

[36], and references therein). A common method is to use true-time delay (TTD) lines instead of using phase shifters (PSs) for analog beamforming [34]–[37]. Unlike the PS-based analog beamforming that produces frequency-independent phase rotations, the TTD-based analog beamforming generates frequency-dependent phase rotations that can be used for compensating the squinted beams in the spatial domain. However, this method is not directly applicable to massive MIMO systems because it requires a large number of TTD lines. Specifically, each transmit antenna needs to be connected to a dedicated TTD, resulting in a high hardware cost and huge power consumption[1].

TTD lines have been recently proposed for sub-THz hybrid massive MIMO OFDM systems [10]–[14]. These systems [10], [12], [13] use a significantly smaller number of TTD lines than that in radar systems [34]–[36]. These methods can be seen as combining a small number of TTD lines with a layer of PSs to form an analog precoder that is able to generate frequency-dependent phase rotations while consuming less power than the conventional TTD architecture in radar systems. However, these approaches still incur a large amount of power consumption when a large number of TTD lines is needed to combat the severe beam squint. Determining the minimum number of TTDs to achieve a desired beam squint compensation capability remains challenging.

The TTD-based hybrid precoding architecture has also been utilized to address fast beam training [40], beam tracking [41], frequency multiplexing [42]–[44], and user localization [45] problems. Unlike the beam squint compensation that aligns the beam directions at every OFDM subcarrier to a same spatial direction, these works [40]–[45] exploit the beam squint effect to spread the beams across different OFDM subcarriers simultaneously. By tracking users' locations simultaneously, the communication overhead for channel sounding is substantially reduced [42]–[45].

Most of these previous works have focused on designing the TTD precoder while keeping the PS precoder fixed. This simplifies the analog precoding design by decoupling the PS and TTD precoders, and reduces the number of design variables since the number of deployed TTDs is usually less than the number of PSs. Additionally, these studies have assumed that the TTD values increase linearly with the number of antennas without any limits. This assumption is, however, unrealistic, as the range of the time delay values that a TTD can produce is strictly limited [1], [46], [47]. For instance, [39] and [38] design TTDs with the maximum time delay values of 400 ps and 508 ps, respectively. The former occupies a smaller circuit board area of 4 mm^2 while the latter occupies 5.45 mm^2. This raises a critical issue; for a given circuit board size (e.g., 128 mm^2 [48]), we need to limit both the time delay

range and the number of TTDs used to mitigate the beam squint.

We note that hardware components are typically static once deployed. Depending on the deployment scenarios, the time delay range of the deployed TTDs may be insufficient for beam squint compensation. It is sustainable to optimize the tunable parameters, such as TTD and PS values, to effectively deal with the beam squint instead of replacing them with new TTDs, which is costly.

B. Contributions and Synopsis

We propose a signal processing framework to mitigate the beam squint effect based on joint TTD and PS optimization. The practical TTD constraint is imposed such that the time delay values are restricted in a fixed interval. Our main contributions are summarized as follows.

- We first determine the ideal analog precoder that fully compensates for the beam squint. Then, we formulate the joint TTD and PS precoder optimization problem based on minimizing the distance between the ideal analog precoder and the product of TTD and PS precoders under the TTD constraints. Unlike the previous approaches [10], [12], [13] that only optimize the TTD precoders under the unbounded time delay values assumption, we jointly optimize the TTD and PS precoders subject to the limited range of time delay values that a practical TTD can produce. Although the formulated problem is non-convex and difficult to solve directly, we show that by transforming the problem into the phase domain, the original problem is converted to an approximated convex problem, which allows us to find a closed-form expression of a solution. Based on the identified solution, our analysis reveals the number of transmit antennas and the amount of time delay required for the best beam squint compensation.
- Leveraging the closed-form expressions of the proposed approach, we formulate a mixed-integer optimization problem to determine the minimum number of TTDs required to achieve a given array gain performance. Although the formulated mixed-integer problem is intractable, we show that by applying a second-order approximation, the original problem can be relaxed to a tractable form, enabling us to find a solution. Our finding indicates that the number of TTDs linearly increases with respect to the system bandwidth to guarantee a required array gain performance.
- We conduct extensive simulations to evaluate the performance of the proposed joint PS and TTD precoding. The simulation results confirm that the beam squint is compensated effectively with our design. Simulations reveal that our joint optimization approach achieves superior array gain performance and computational efficiency compared to prior TTD-based precoding methods. The simulations also verify the significant array gain performance improvement with the minimum number of TTDs informed by our optimization approaches.

Synopsis: The remainder of the paper is organized as follows. Section II presents the channel model of the wideband sub-

---

[1]While the power consumption of a TTD device depends on a specific process technology (e.g., BiCMOS [38], and CMOS [39]), a typical TTD in sub-THz consumes 100 mW [10]. It is worth noting that the power consumption of a typical PS in sub-THz is 20 mW [10], which is much lower than that of a TTD.

THz massive MIMO OFDM system and analyzes the array gain loss caused by the beam squint. Section III describes the relationship between beam squint compensation by the ideal analog precoder and the achievable rate. Section IV derives the closed-form solution of the TTD and PS precoders under the practical TTD constraints and determines the minimum number of TTDs that ensures a predefined array gain performance. Section V provides simulation results to validate the developed analysis. Section VI concludes this work.

Notation: A bold lower case letter $\mathbf{x}$ is a column vector and a bold upper case letter $\mathbf{X}$ is a matrix. $\mathbf{X}^T$, $\mathbf{X}^H$, $\mathbf{X}^{-1}$, $\|\mathbf{X}\|_F$, $\text{tr}(\mathbf{X})$, $\det(\mathbf{X})$, $\mathbf{X}(i,j)$, $\|\mathbf{x}\|_2$, and $|x|$ are, respectively, the transpose, conjugate transpose, inverse, Frobenius norm, trace, determinant, ith row and jth column entry of $\mathbf{X}$, 2-norm of $\mathbf{x}$, and modulus of $x \in \mathbb{C}$. $\text{blk}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ is an $nN \times N$ block diagonal matrix such that its main-diagonal blocks contain $\mathbf{x}_i \in \mathbb{C}^{n \times 1}$, for $i = 1, \ldots, N$ and all off-diagonal blocks are zero. $\mathbf{0}_n$, $\mathbf{1}_n$, and $\mathbf{I}_n$ denote, respectively, the $n \times 1$ all-zero vector, $n \times 1$ all-one vector, and $n \times n$ identity matrix. Given $\mathbf{x} \in \mathbb{R}^{n \times 1}$, $e^{j\mathbf{x}}$ denotes the column vector $[e^{jx_1} \ e^{jx_2} \ \ldots \ e^{jx_n}]^T \in \mathbb{C}^{n \times 1}$ obtained by applying e^j element-wise.

II. CHANNEL MODEL AND BEAM SQUINT EFFECT

In this section, we present the channel model of the wide-band sub-THz massive MIMO OFDM system. Then, the array gain loss caused by the beam squint is analyzed.

A. Channel Model

We consider the downlink of a wideband sub-THz hybrid massive MIMO OFDM system where the transmitter is equipped with an N_t-element uniform linear array (ULA) with element spacing d. The transmit antenna array is fed by N_{RF} radio frequency (RF) chains to simultaneously transmit N_s data streams to an N_r-element ULA receiver. It is assumed that N_t, N_r, N_{RF}, and N_s satisfy $N_s = N_{RF} \leq N_r \ll N_t$. Herein, we let f_c, B, and K be, respectively, the central (carrier) frequency, bandwidth of the OFDM system, and the number of OFDM subcarriers (an odd number). Then, the kth subcarrier frequency is given, for $k = 1, \ldots, K$, by

$$f_k = f_c + \frac{B}{K}\Big(k - 1 - \frac{K-1}{2}\Big). \tag{1}$$

The frequency domain MIMO OFDM channel at the kth subcarrier $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is

$$\mathbf{H}_k = \sqrt{\frac{N_r N_t}{L}} \sum_{l=1}^{L} \alpha_{k,l} e^{-j2\pi\tau_l f_k} \mathbf{u}_{k,l} \mathbf{v}_{k,l}^H, \tag{2}$$

where L denotes the number of channel (spatial) paths and $\tau_l \in \mathbb{R}$ represents the delay of the lth channel path. The $\alpha_{k,l} \in \mathbb{C}$ denotes the path gain of the kth subcarrier on the lth path, which incorporates the molecular absorption loss of sub-THz wave propagation medium. The magnitude of $\alpha_{k,l}$ is modeled by $\mathbb{E}[|\alpha_{k,l}|^2] = \left(\frac{c}{4\pi f_k \widetilde{d}}\right)^2 e^{-\mathcal{K}_{\text{abs}}(f_k)\widetilde{d}}$, where $c = 3 \times 10^8$ m/s denotes the speed of light, $\widetilde{d}$ is the transmit distance, and $\mathcal{K}_{\text{abs}}(f)$ is the frequency-dependent medium

absorption coefficient [49], [50]. We note that the sub-THz channel is unlikely to have a number of spatial paths L that is larger than the number of RF chains N_{RF} [51]. For ease of exposition, we assume $L = N_{RF}$, which is equivalent to setting $\alpha_{k,L+1} = \cdots = \alpha_{k,N_{RF}} = 0$ when $L < N_{RF}$, for $k = 1, 2, \ldots, K$. This setting has the same effect as using L RF chains for transmission while turning off $(N_{RF} - L)$ RF chains. In the remainder of this paper, we use the subscript l for denoting the index of both channel path and RF chain unless specified otherwise. The vectors $\mathbf{v}_{k,l} \in \mathbb{C}^{N_t \times 1}$ and $\mathbf{u}_{k,l} \in \mathbb{C}^{N_r \times 1}$ in (2) are the normalized transmit and receive array response vectors of the kth subcarrier on the lth path, respectively, where the $\mathbf{v}_{k,l}$ is a function of angles of departure (AoDs) $\Psi_l \in [\Psi_{\min}, \Psi_{\max}]$. The exact same definition applies to the normalized receive array response vector $\mathbf{u}_{k,l}$ in (2).

In what follows, we will limit our discussion to the transmit antenna array, keeping in mind that the same applies to the receive array [10]–[13]. The antenna geometry of the transmit array is described assuming far-field spatial angles. The nth entry of $\mathbf{v}_{k,l} \in \mathbb{C}^{N_t \times 1}$ in (2) is $\mathbf{v}_{k,l}(n,1) = \frac{1}{\sqrt{N_t}} e^{-j\pi \frac{2df_k}{c}(n-1)\sin(\Psi_l)}$. We define $\psi_{k,l} = \frac{2df_k}{c}\sin(\Psi_l)$ as the spatial direction of the kth subcarrier at the transmitter on the lth path. Assuming the half-wavelength antenna spacing, i.e., $d = \frac{c}{2f_c}$, the spatial direction at the central frequency is simplified to $\psi_l = \sin(\Psi_l)$. Hence, setting $\zeta_k = \frac{f_k}{f_c}$ leads to $\psi_{k,l} = \zeta_k \psi_l$ and

$$\zeta_k = 1 + \frac{B}{f_c}\Big(\frac{k - 1 - \frac{K-1}{2}}{K}\Big), \tag{3}$$

where (3) follows from the definition of f_k in (1). As a result, the transmit array response vector $\mathbf{v}_{k,l}$ in (2) is succinctly expressed as $\mathbf{v}_{k,l} = \frac{1}{\sqrt{N_t}}\Big[1 \ e^{j\pi\psi_{k,l}} \ \ldots \ e^{j\pi(N_t-1)\psi_{k,l}}\Big]^H \in \mathbb{C}^{N_t \times 1}$.

B. Beam Squint Effect

As aforementioned in Section I, when the wideband OFDM is employed in a massive MIMO system, a substantial array gain loss at each subcarrier could occur due to beam squint. To quantify it, we focus on the lth path of the channel in (2). Denoting the frequency-independent beamforming vector matched to the transmit array response vector with the AoD Ψ_l as $\mathbf{f}^{(l)} = \mathbf{v}_{c,l}$ (i.e., array response vector at the central frequency f_c on the lth path), the corresponding array gain at the kth subcarrier is given by $g(\mathbf{f}^{(l)}, \psi_{k,l}) = |\mathbf{v}_{k,l}^H \mathbf{f}^{(l)}|$ [10], [26], i.e.,

$$g(\mathbf{f}^{(l)}, \psi_{k,l}) = \frac{1}{N_t}\left|\sum_{n=0}^{N_t-1} e^{jn\pi(\psi_{k,l}-\psi_l)}\right| = \left|\frac{\sin(N_t \Delta_{k,l})}{N_t \sin(\Delta_{k,l})}\right|, \tag{4}$$

where $\Delta_{k,l} = \frac{\pi}{2}(\psi_{k,l} - \psi_l)$. It is not difficult to observe that at the central frequency, the array gain is $g(\mathbf{f}^{(l)}, \psi_{k,l}) = 1$ because $\lim_{x \to 0} \frac{\sin(N_t x)}{\sin(x)} = N_t$. However, when $f_k \neq f_c$, $\Delta_{k,l}$ deviates from 0; the amount of deviation increases as f_k approaches to f_1 or f_K. As a result, all subcarriers except for the central subcarrier suffer from the array gain loss. The implication in the spatial domain is that the beams at

non-central subcarriers may completely split from the one generated at the central subcarrier.

The following proposition quantifies the asymptotic array gain loss as $N_t \to \infty$.

Proposition 1. Suppose that $\psi_{k,l}$ is the spatial direction at the kth subcarrier ($f_k \neq f_c$) of the lth path. Then, the array gain in (4) converges to 0 as N_t tends to infinity, i.e., $g(\mathbf{f}^{(l)}, \psi_{k,l}) \doteq 0$, where $\doteq$ denotes the equality when $N_t \to \infty$.

Proof. The array gain in (4) can be rewritten as $g(\mathbf{f}^{(l)}, \psi_{k,l}) = \frac{1}{N_t} \left| \frac{\sin(N_t \Delta_{k,l})}{\pi \Delta_{k,l}} \frac{\pi \Delta_{k,l}}{\sin(\Delta_{k,l})} \right|$, where $\Delta_{k,l} \neq 0$ because $f_k \neq f_c$. The proposition follows from the definition of Dirac delta function $\frac{\sin(N_t \Delta_{k,l})}{\pi \Delta_{k,l}} \doteq \delta(\Delta_{k,l})$ [52], completing the proof. $\square$

Fig. 1a illustrates the convergence trend of Proposition 1. The array gain patterns are calculated for $f_c = 300$ GHz, $K = 129$, $\psi_l = 0.8$, and $B = 30$ GHz. As N_t tends to be large, the maximum array gain is only achieved at the central frequency, i.e., the 65th subcarrier in Fig. 1a, while other subcarriers suffer from substantial array gain losses. This is quite the opposite of the traditional narrowband massive MIMO system in which the array gain grows as $N_t \to \infty$.

The array gain loss is also numerically understood when the bandwidth B increases. Fig. 1b illustrates those patterns while using the same parameters as in Fig. 1a except for that $N_t = 256$. As the bandwidth B grows, the maximum array gain is obtained only at the central frequency while other subcarriers experience a large amount of array gain losses. This contrasts with the information-theoretic insight that the achievable rate grows linearly with the bandwidth while keeping the signal-to-noise-ratio (SNR) fixed.

Wideband sub-THz massive MIMO communication research is in its early stages. In order to truly unleash the potential of sub-THz communications, a hybrid precoding architecture and methods that can effectively compensate for the beam squint effect is of paramount importance.

III. PRELIMINARIES AND ACHIEVABLE RATE PERFORMANCE

We consider a TTD-based hybrid precoding architecture [10], [12], where each RF chain drives M TTDs and each TTD is connected to N PSs as shown in Fig. 1c. The N_t-element ULA is divided into M subarrays with $N = \frac{N_t}{M}$ antennas per subarray. The signal at the kth subcarrier passing through a TTD is delayed by t ($0 \leq t \leq t_{\max}$) in the time domain, which corresponds to the $-2\pi f_k t$ frequency-dependent phase rotation in the frequency domain. The $t_{\max}$ is the maximum time delay value that a TTD device can produce. The kth subcarrier signal at the receiver $\mathbf{y}_k \in \mathbb{C}^{N_r \times 1}$ is then given by [1]

$$\mathbf{y}_k = \sqrt{\rho} \mathbf{H}_k \mathbf{F}_1 \mathbf{F}_{2,k}(\{\mathbf{t}_l\}_{l=1}^{N_{RF}}) \mathbf{W}_k \mathbf{s}_k + \mathbf{n}_k, \quad (5)$$

where ρ, $\mathbf{s}_k \in \mathbb{C}^{N_s \times 1}$, $\mathbf{W}_k \in \mathbb{C}^{N_{RF} \times N_s}$, and $\mathbf{n}_k \in \mathbb{C}^{N_r \times 1}$ are, respectively, the average transmit power, transmit data stream, baseband digital precoder, and normal Gaussian noise

vector with each entry being independent and identically distributed (i.i.d.) according to zero mean and variance 1. The $\mathbf{F}_1$ in (5) is the PS precoding matrix, which is the concatenation of N_{RF} PS submatrices $\mathbf{X}_l \in \mathbb{C}^{N_t \times M}$, $1 \leq l \leq N_{RF}$, specifically,

$$\mathbf{F}_1 = \frac{1}{\sqrt{N_t}}[\mathbf{X}_1 \ \mathbf{X}_2 \ \ldots \ \mathbf{X}_{N_{RF}}] \in \mathbb{C}^{N_t \times MN_{RF}}, \quad (6)$$

where $\mathbf{X}_l = \text{blk}(e^{j\pi \mathbf{x}_1^{(l)}}, e^{j\pi \mathbf{x}_2^{(l)}}, \ldots, e^{j\pi \mathbf{x}_M^{(l)}}) \in \mathbb{C}^{N_t \times M}$, $\mathbf{x}_m^{(l)} = [x_{1,m}^{(l)} \ x_{2,m}^{(l)} \ \ldots \ x_{N,m}^{(l)}]^T \in \mathbb{R}^{N \times 1}$ is a PS vector, and $x_{n,m}^{(l)}$ is the value of the nth PS that is connected to the mth TTD on the lth RF chain, for $1 \leq n \leq N$, $1 \leq l \leq N_{RF}$, and $1 \leq m \leq M$. The $\mathbf{F}_{2,k}(\{\mathbf{t}_l\}_{l=1}^{N_{RF}}) \in \mathbb{C}^{MN_{RF} \times N_{RF}}$ in (5) is the TTD precoding matrix, which is defined as

$$\mathbf{F}_{2,k}(\{\mathbf{t}_l\}_{l=1}^{N_{RF}}) = \text{blk}(e^{-j2\pi f_k \mathbf{t}_1}, e^{-j2\pi f_k \mathbf{t}_2}, \ldots, e^{-j2\pi f_k \mathbf{t}_{N_{RF}}}), \quad (7)$$

where $\mathbf{t}_l = [t_1^{(l)} \ t_2^{(l)} \ \ldots \ t_M^{(l)}]^T \in \mathbb{R}^{M \times 1}$ is the lth time delay vector and $t_m^{(l)}$ is the time delay value of the mth TTD on the lth RF chain. For ease of exposition, in what follows, we omit the time delay vectors $\{\mathbf{t}_l\}_{l=1}^{N_{RF}}$ in the TTD precoder notation $\mathbf{F}_{2,k}(\{\mathbf{t}_l\}_{l=1}^{N_{RF}})$, $\forall k$. We note that the analog precoder corresponds to the product $\mathbf{F}_1 \mathbf{F}_{2,k} \in \mathbb{C}^{N_t \times N_{RF}}$ with the constant modulus constraint

$$\mathbf{F}_1 \mathbf{F}_{2,k} \in \mathcal{F}_{N_t, N_{RF}}, \quad (8)$$

where $\mathcal{F}_{N_t, N_{RF}}$ denotes the set of all matrices $\mathbf{X} \in \mathbb{C}^{N_t \times N_{RF}}$ such that $|\mathbf{X}(i,j)| = \frac{1}{\sqrt{N_t}}, \forall i, j$. The data stream $\mathbf{s}_k$ in (5) satisfies $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] = \frac{1}{N_s} \mathbf{I}_{N_s}$. Then, the precoders are normalized such that $\|\mathbf{F}_1 \mathbf{F}_{2,k} \mathbf{W}_k\|_F^2 = N_s$, leading to $\mathbb{E}[\|\mathbf{F}_1 \mathbf{F}_{2,k} \mathbf{W}_k \mathbf{s}_k\|_2^2] = 1$, for $1 \leq k \leq K$.

A conventional approach to maximize the system's achievable rate is to jointly design the digital precoders $\{\mathbf{W}_k\}$, TTD precoders $\{\mathbf{F}_{2,k}\}$, and PS precoder $\mathbf{F}_1$ in (5). However, this approach is intractable due to the coupling between variables and non-convex constraints of TTD and PS precoders [21], [22], [53]. Therefore, we adopt a sequential hybrid precoding method [54]–[56]. Focusing on the analog domain, we propose to jointly design the TTD and PS precoders to compensate for the beam squint while fixing the digital precoders. Subsequently, the digital precoders can be optimized by standard approaches such as [10], [12] to maximize the achievable rate.

A. Preliminaries

We first describe the sign invariance property of the array gain in (4) and then identify the ideal analog precoder that completely compensates for the beam squint. The sign invariance property and the ideal analog precoder established in this subsection will then be used in Section IV for jointly optimizing TTD and PS precoders.

1) Sign Invariance of Array Gain

We define the combination of PS precoder $\mathbf{F}_1$ and TTD precoder $\mathbf{F}_{2,k}$ as

$$\mathbf{F}_k = \mathbf{F}_1 \mathbf{F}_{2,k}, \text{ for } 1 \leq k \leq K, \quad (9)$$
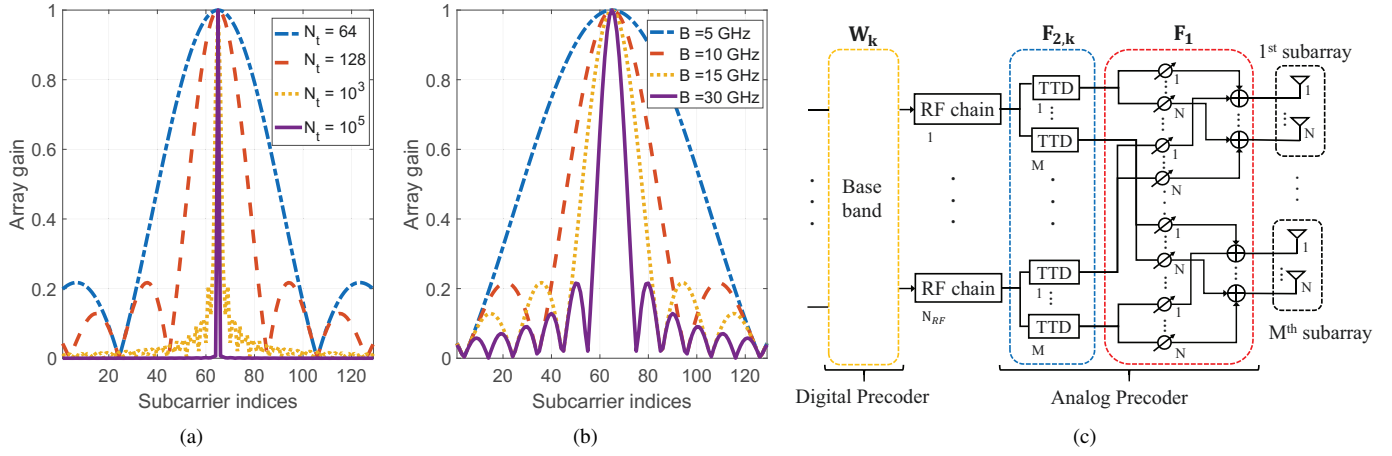
Fig. 1. (a) Array gain vs. subcarrier indices for different numbers of transmit antennas (N_t). (b) Array gain vs. subcarrier indices for different bandwidths (B). (c) TTD-based hybrid precoding architecture.

where the lth column of $\mathbf{F}_k$ is $\mathbf{f}_k^{(l)} = \frac{1}{\sqrt{N_t}} \mathbf{X}_l e^{-2\pi f_k \mathbf{t}_l}$, for $1 \leq l \leq N_{RF}$. The array gain associated with $\mathbf{f}_k^{(l)}$ is then given, based on (4), by

$$g(\mathbf{f}_k^{(l)}, \psi_{k,l}) = \frac{1}{N_t} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} e^{j\pi\zeta_k \gamma_{n,m}^{(l)}} e^{j\pi x_{n,m}^{(l)}} e^{-j\pi\zeta_k \vartheta_m^{(l)}} \right|, \quad (10)$$

where $\gamma_{n,m}^{(l)} = ((m - 1)N + n - 1)\psi_l$ and $\vartheta_m^{(l)} = 2f_c t_m^{(l)} \in [0, \vartheta_{\max}]$ with $\vartheta_{\max} = 2f_c t_{\max}$. Note that in (10), $e^{j\pi\zeta_k \gamma_{n,m}^{(l)}}$ is frequency-dependent while $e^{j\pi x_{n,m}^{(l)}}$ is frequency-independent, and $e^{-j\pi\zeta_k \vartheta_m^{(l)}}$ depends on both time and frequency. The beam squint compensation capability of the mth TTD on the lth RF chain is restricted because the phase rotation $-\pi\zeta_k \vartheta_m^{(l)}$ is within the interval $[-\pi\zeta_k \vartheta_{\max}, 0]$.

Remark 1. (Sign Invariance Property) The $g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ in (10) is invariant to the multiplication of negative signs to $\gamma_{n,m}^{(l)}$, $x_{n,m}^{(l)}$, and $\vartheta_m^{(l)}$. To be specific, given $\psi_l \geq 0$ (i.e., $\gamma_{n,m}^{(l)} \geq 0$, $\forall m, n$), we denote $\{x_{n,m}^{(l)}{}^\star\}$ and $\{\vartheta_m^{(l)}{}^\star\}$ as the optimal values of $\{x_{n,m}^{(l)}\}$ and $\{\vartheta_m^{(l)}\}$, respectively, that maximize $g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ in (10). Then, it is not difficult to observe that $\{-x_{n,m}^{(l)}{}^\star\}$ and $\{\vartheta_{\max} - \vartheta_m^{(l)}{}^\star\}$ also maximize $g(\mathbf{f}_k^{(l)}, -\psi_{k,l})$. Hence, without loss of generality, in what follows, we assume that

$$\psi_l \geq 0, \text{ for } 1 \leq l \leq N_{RF}. \quad (11)$$

The sign invariance property will be found to be useful when deriving the solution to our optimization problem for joint time delay and phase shift precoding in Section IV.

2) Ideal Analog Precoder

An ideal analog precoder can produce arbitrary frequency-dependent phase rotation values to completely mitigate the beam squint effect, which is hereafter denoted as $\widetilde{\mathbf{F}}_k^\star \in \mathcal{F}_{N_t \times N_{RF}}$, for $1 \leq k \leq K$. The purpose of invoking the ideal analog precoder is to provide a reference design for the proposed joint TTD and PS precoding method in Section IV. Denoting $\widetilde{\mathbf{f}}_k^{(l)}$ as the lth column of $\widetilde{\mathbf{F}}_k^\star$, the array

gain obtained by $\widetilde{\mathbf{f}}_k^{(l)}$ is $g(\widetilde{\mathbf{f}}_k^{(l)}, \psi_{k,l}) = |\mathbf{v}_{k,l}^H \widetilde{\mathbf{f}}_k^{(l)}| = 1$, for $1 \leq l \leq N_{RF}$, resulting in $\widetilde{\mathbf{F}}_k^\star = [\mathbf{v}_{k,1} \ \mathbf{v}_{k,2} \ \dots \ \mathbf{v}_{k,N_{RF}}]$. Thus, the $((m-1)N+n)$th row and lth column entry of $\widetilde{\mathbf{F}}_k^\star$ is

$$\widetilde{\mathbf{F}}_k^\star((m - 1)N + n, l) = \frac{1}{\sqrt{N_t}} e^{-j\pi\zeta_k \gamma_{n,m}^{(l)}}, \forall k, l, m, n. \quad (12)$$

The ideal analog precoder $\widetilde{\mathbf{F}}_k^\star$ in (12) is achievable with $\mathbf{F}_1 \mathbf{F}_{2,k}$ when each transmits antenna is equipped with a dedicated TTD, i.e., $M = N_t$ and $N = 1$. In this case, the lth column of $\mathbf{F}_1 \mathbf{F}_{2,k}$ is matched exactly to the array response vector of the kth subcarrier on the lth path, $1 \leq l \leq N_{RF}$ and $1 \leq k \leq K$. For instance, the PS and TTD values are designed as $x_{1,m}^{(l)} = 0$ and $t_m^{(l)} = \frac{(m-1)\psi_l}{2f_c}$, respectively, $1 \leq m \leq N_t$ and $1 \leq l \leq N_{RF}$. However, this design is not practically motivated because it requires N_t^2 TTDs, resulting in a high hardware complexity and huge power consumption. In Section IV, we address the latter issue by proposing a method to minimize the power consumption of the analog precoder given an array gain performance requirement.

B. Achievable Rate Performance

Given the analog precoders $\{\mathbf{F}_k\}_{k=1}^{K}$ in (9), the achievable rate (averaged over the subcarriers) of the channel in (2) is

$$R = \frac{1}{K} \sum_{k=1}^{K} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{\rho}{N_s} \mathbf{H}_k \mathbf{F}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{F}_k^H \mathbf{H}_k^H \right). \quad (13)$$

To quantify the impact of the ideal analog precoders $\{\widetilde{\mathbf{F}}_k^\star\}$ on the achievable rate R in (13), we define the singular value decomposition (SVD) of $\mathbf{H}_k$ as $\mathbf{H}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H$, where $\mathbf{U}_k \in \mathbb{C}^{N_r \times N_{RF}}$ satisfying $\mathbf{U}_k^H \mathbf{U}_k = \mathbf{I}_{N_{RF}}$, $\mathbf{V}_k \in \mathbb{C}^{N_t \times N_{RF}}$ satisfying $\mathbf{V}_k^H \mathbf{V}_k = \mathbf{I}_{N_{RF}}$, and $\mathbf{\Sigma}_k \in \mathbb{R}^{N_{RF} \times N_{RF}}$ is a diagonal matrix of non-zero singular values arranged in descending order. It is well understood that the optimal fully-digital precoders maximizing R in (13) are $\{\mathbf{V}_k\}$ [23]. Consequently, the precoders $\{\mathbf{F}_k\}$ and $\{\mathbf{W}_k\}$ that achieve this maximal rate can be designed by solving $\mathbf{F}_k \mathbf{W}_k = \mathbf{V}_k$, for $k = 1, \dots, K$. In subsequent analysis, we confirm the existence of a digital

Fig. 2. (a) Conceptual illustration of the prior methods for beam squint compensation [10], [12], [13] . (b) The proposed method of this work. (c) Conceptual illustration showing the relationship between the distance of two points on the unit modulus circle (the segment AB) and their relative phase difference (the arc corresponding to $y - x$)

precoder $\mathbf{W}_k^\star$ such that $\widetilde{\mathbf{F}}_k^\star \mathbf{W}_k^\star = \mathbf{V}_k$, for $k = 1, \ldots, K$. From the channel model in (2), we have $\mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{V}_k^H = \mathbf{G}_k \boldsymbol{\Lambda}_k \widetilde{\mathbf{F}}_k^{\star H}$, where $\mathbf{G}_k = [\mathbf{u}_{k,1} \ldots \mathbf{u}_{k,L}] \in \mathbb{C}^{N_r \times L}$ and $\boldsymbol{\Lambda}_k = \sqrt{\frac{N_t N_r}{L}} \operatorname{diag}(\alpha_{k,1} e^{-j2\pi\tau_l f_k}, \ldots, \alpha_{k,L} e^{-j2\pi\tau_l f_k})$. It can be verified that $\mathbf{V}_k = \widetilde{\mathbf{F}}_k^\star \boldsymbol{\Lambda}_k^H \mathbf{G}_k^H \mathbf{U}_k (\boldsymbol{\Sigma}_k^T)^{-1}$ due to the assumption $L = N_{RF}$. It is feasible to match the performance of the optimal fully-digital precoder $\mathbf{V}_k$ using the idea analog precoder $\widetilde{\mathbf{F}}_k^\star$ combined with the digital precoder $\mathbf{W}_k = \boldsymbol{\Lambda}_k^H \mathbf{G}_k^H \mathbf{U}_k (\boldsymbol{\Sigma}_k^T)^{-1}$. Therefore, we attempt in Section IV to design the TTD precoder $\mathbf{F}_{2,k}$ and PS precoder $\mathbf{F}_1$ in order to best approximate the ideal analog precoder $\widetilde{\mathbf{F}}_k^\star$, for $k = 1, \ldots, K$, under the per-TTD time delay constraint.

IV. JOINT DELAY AND PHASE PRECODING UNDER TTD CONSTRAINTS

We propose a novel method to design TTD and PS precoders, which is illustrated in Fig. 2b. In contrary with the prior works [10], [12], [13] that optimize TTD values while fixing the PS values (Fig. 2a), we jointly optimize both TTD and PS values (Fig. 2b). Instead of assuming the impractical TTD that produces any time delay value t (i.e., $0 \le t \le \infty$, Fig. 2a), a finite interval time delay value (i.e., $0 \le t \le t_{\max}$, Fig. 2b) is taken into account in our approach.

A. Problem Formulation

Ideally speaking, we wish to find $\{\mathbf{t}_l\}_{l=1}^{N_{RF}}$ and $\mathbf{F}_1$ satisfying $\mathbf{F}_1 \mathbf{F}_{2,k} = \widetilde{\mathbf{F}}_k^\star$, $\forall k$. However, given fixed l, m, and n, solving K-coupled matrix equations is an ill-posed problem. To overcome this, we approach to formulate a problem that optimizes $\mathbf{F}_1$ and $\{\mathbf{t}_l\}_{l=1}^{N_{RF}}$ by minimizing the difference between $\mathbf{F}_1 \mathbf{F}_{2,k}$ and $\widetilde{\mathbf{F}}_k^\star$:

$$\min_{\mathbf{F}_1, \{\mathbf{F}_{2,k}\}_{k=1}^K} \quad \frac{1}{K} \sum_{k=1}^K \left\| \widetilde{\mathbf{F}}_k^\star - \mathbf{F}_1 \mathbf{F}_{2,k} \right\|_F^2 \tag{14a}$$

$$\text{subject to} \quad 0 \le t_m^{(l)} \le t_{\max}, \forall l, m, \tag{14b}$$

$$|\mathbf{F}_1(i,j)| \in \left\{ \frac{1}{\sqrt{N_t}}, 0 \right\}, \forall i, j, \tag{14c}$$

$$|\mathbf{F}_{2,k}(p,q)| \in \{0, 1\}, \forall p, q, \tag{14d}$$

$$\mathbf{F}_1 \mathbf{F}_{2,k} \in \mathcal{F}_{N_t, N_{RF}}, \forall k, \tag{14e}$$

where the constraint in (14b) indicates the restricted range of the time delay values per-TTD device, the constraint in (14c) is due to the definition of $\mathbf{F}_1$ in (6), the constraint in (14d) is due to the definition of $\mathbf{F}_{2,k}$ in (7), and the constraint in (14e) describes the constant modulus property of the analog precoder in (8). The constraints in (14c)-(14e) in conjunction with the coupling between $\mathbf{F}_1$ and $\mathbf{F}_{2,k}$ in (14a) make the problem difficult to solve. Besides, (14) can be viewed as a matrix factorization problem with non-convex constraints, which has been studied in the context of hybrid analog-digital precoding [21], [23], [57], [58]. A common approach was applying block coordinate descent (BCD) and relaxing the constraints to deal with the non-convexity [21], [23], [57], [58]. Unlike the prior works, the original non-convex problem in (14) is approached, in this work, as an approximated convex problem.

Remark 2. We acknowledge the resemblance of the problem in (14) to the one studied in [43], which explores a generic beamforming behavior for fast frequency multiplexing. The problem in [43, Eqn. (3)] can be viewed as a regularized variant of the problem in (14). Unlike the iterative method in [43], our work specifically targets the challenges of beam squint compensation, introducing a novel, one-shot joint design for PS and TTD values. This advancement allows us to delineate the system parameters, including N_t, M, B, and $t_{\max}$ for effective beam squint compensation, which will be addressed in the later part of this section, a perspective that remains unexplored in the prior work [43].

Based on the structure of $\mathbf{F}_1$ in (6) and the block diagonal structure of $\{\mathbf{F}_{2,k}\}_{k=1}^K$ in (7), the objective function in (14a) can be rewritten as

$$\frac{1}{K} \frac{1}{N_t} \sum_{k=1}^K \sum_{l=1}^{N_{RF}} \sum_{m=1}^M \sum_{n=1}^N \left| e^{-j\pi\zeta_k \gamma_{n,m}^{(l)}} - e^{j\pi x_{n,m}^{(l)}} e^{-j\pi\zeta_k \vartheta_m^{(l)}} \right|^2. \tag{15}$$

It is still difficult to deal with the objective function in (15) due to the unit modulus constraint. To streamline the optimization process, we introduce a lemma demonstrating the equivalence of optimization on the unit circle to optimization in the phase domain.

***Lemma* 1.** For $x \in \mathbb{R}$ and $y \in \mathbb{R}$, the following equality

holds $\underset{y:\,\mathrm{mod}(y,\pi)\neq x}{\mathrm{argmin}}\;|e^{jx} - e^{jy}| = \underset{y:\,\mathrm{mod}(y,\pi)\neq x}{\mathrm{argmin}}\;|x - y|$, where $\mathrm{mod}(y, \pi)$ is y modulo π.

Proof. See Appendix A. $\qquad\square$

Fig. 2c graphically visualizes the equivalence in Lemma 1. We let points A and B represent e^{jx} and e^{jy}, respectively. It is not difficult to observe from Fig. 2c that minimizing the length of the segment AB with respect to B is equivalent to minimizing $|y - x|$ with respect to y, in which the latter is a convex optimization problem.

In what follows, Lemma 1 is exploited to convert the non-convex problem in (14) into an approximated convex problem. Incorporating Lemma 1 into (15) converts the problem in (14) to an approximation [1], [43]:

$$\min_{\{x_{n,m}^{(l)}\},\{\vartheta_m^{(l)}\}} \frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{N_{RF}}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|x_{n,m}^{(l)} - \zeta_k\vartheta_m^{(l)} + \zeta_k\gamma_{n,m}^{(l)}\right|^2, (16a)$$

$$\text{subject to}\quad 0 \leq \vartheta_m^{(l)} \leq \vartheta_{\max},\ \forall l, m. \tag{16b}$$

Next, we turn (16) to a composite matrix optimization problem. The PS and TTD variables in (16) can be collected into a matrix $\mathbf{A}_l = [\mathbf{a}_1^{(l)}\ \ldots\ \mathbf{a}_M^{(l)}] \in \mathbb{R}^{(N+1)\times M}$, where $\mathbf{a}_m^{(l)} = [(\mathbf{x}_m^{(l)})^T\ \vartheta_m^{(l)}]^T \in \mathbb{R}^{N+1}$. Containing the analog counterpart in $\widetilde{\mathbf{F}}_k^{\star}$ in a matrix $\mathbf{B}_k^{(l)} \in \mathbb{R}^{N\times M}$, where the nth row and mth column entry of $\mathbf{B}_k^{(l)}$ is $B_k^{(l)}(n, m) = -\zeta_k\gamma_{n,m}^{(l)}, \forall k, l, n, m$, the problem (16) becomes

$$\min_{\{\mathbf{A}_l\}_{l=1}^{N_{RF}}}\quad \frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{N_{RF}}\left\|\mathbf{C}_k\mathbf{A}_l - \mathbf{B}_k^{(l)}\right\|_F^2, \tag{17a}$$

$$\text{subject to}\quad \mathbf{0}_M^T \leq \mathbf{e}_{N+1}^T\mathbf{A}_l \leq \vartheta_{\max}\mathbf{1}_M^T,\ \forall l, \tag{17b}$$

where $\mathbf{C}_k = [\mathbf{I}_N\ -\zeta_k\mathbf{1}_N] \in \mathbb{R}^{N\times(N+1)}$ and $\mathbf{e}_{N+1} \in \mathbb{R}^{N+1}$ is the $(N+1)$th column of the identity matrix $\mathbf{I}_{N+1}$. The vector inequalities in (17) is the entry-wise inequalities.

By introducing $\mathbf{C} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{C}_k^T\mathbf{C}_k \in \mathbb{R}^{(N+1)\times(N+1)}$, $\mathbf{D}_l = \frac{1}{K}\sum_{k=1}^{K}\mathbf{C}_k^T\mathbf{B}_k^{(l)} \in \mathbb{R}^{(N+1)\times M}$, and $c_B^{(l)} = \frac{1}{K}\sum_{k=1}^{K}\|\mathbf{B}_k^{(l)}\|_F^2$, the objective function in (17) can be rewritten as $\frac{1}{K}\sum_{k=1}^{K}\sum_{l=1}^{N_{RF}}\left\|\mathbf{C}_k\mathbf{A}_l - \mathbf{B}_k^{(l)}\right\|_F^2 = \sum_{l=1}^{N_{RF}}\sum_{m=1}^{M}\left((\mathbf{a}_m^{(l)})^T\mathbf{C}\mathbf{a}_m^{(l)} - 2(\mathbf{d}_m^{(l)})^T\mathbf{a}_m^{(l)}\right) + \sum_{l=1}^{N_{RF}}c_B^{(l)}$, where the $\mathbf{d}_m^{(l)}$ is the mth column of $\mathbf{D}_l$. Hence, the problem (17) is equivalently

$$\min_{\mathbf{a}_m^{(l)}}\quad (\mathbf{a}_m^{(l)})^T\mathbf{C}\mathbf{a}_m^{(l)} - 2(\mathbf{d}_m^{(l)})^T\mathbf{a}_m^{(l)}, \tag{18a}$$

$$\text{subject to}\quad 0 \leq \mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} \leq \vartheta_{\max}, \forall l, m. \tag{18b}$$

The problem (18) can be viewed as a decomposition of (17) into MN_{RF} independent problems. This decomposition reveals an alignment with the TTD and PS precoding architecture in Fig. 1c, where each RF chain feeds M TTDs and each TTD feeds N PSs. Thus, the problem in (18) is equivalent to optimizing the TTD and PS values of each RF chain branch independently.

B. Optimal Closed-form Solution

In this subsection, we find the optimal closed-form solution of the convex problem (18). We start by deriving the expression of $\mathbf{C}$ in (18). To this end, we first define the constant $\Gamma = N + \eta$, where $\eta = \frac{NB^2}{f_c^2}\frac{(K^2-1)}{12K^2}$. For $\mathbf{C}_k$ in (17), we have $\mathbf{C}_k^T\mathbf{C}_k = \begin{bmatrix} \mathbf{I}_N & -\zeta_k\mathbf{1}_N \\ -\zeta_k\mathbf{1}_N^T & N\zeta_k^2 \end{bmatrix}$, where ζ_k is defined in (3). After some algebraic manipulations, it is readily verified that

$$\sum_{k=1}^{K}\zeta_k = K \text{ and } \sum_{k=1}^{K}\zeta_k^2 = K\left(1 + \frac{\eta}{N}\right). \tag{19}$$

Then, $\mathbf{C} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{C}_k^T\mathbf{C}_k$ is simplified, based on (19), to $\mathbf{C} = \begin{bmatrix} \mathbf{I}_N & -\mathbf{1}_N \\ -\mathbf{1}_N^T & \Gamma \end{bmatrix}$, which yields

$$\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{e}_{N+1} = \frac{1}{\eta}. \tag{20}$$

Defining $\mathbf{b}_{k,m}^{(l)} = [-\zeta_k\gamma_{1,m}^{(l)}\ \ldots\ -\zeta_k\gamma_{N,m}^{(l)}]^T$ as the mth column of $\mathbf{B}_k^{(l)}$ in (17), we attain

$$\mathbf{d}_m^{(l)} = \begin{bmatrix} \frac{1}{K}\sum_{k=1}^{K}\mathbf{b}_{k,m}^{(l)} \\ \frac{1}{K}\sum_{k=1}^{K}-\zeta_k\mathbf{1}_N^T\mathbf{b}_{k,m}^{(l)} \end{bmatrix} \in \mathbb{R}^{(N+1)\times 1}. \tag{21}$$

Using (21), it is readily verified that $\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} = \frac{1}{\eta}\frac{1}{K}\sum_{k=1}^{K}\left((1 - \zeta_k)\sum_{n=1}^{N}\mathbf{b}_{k,m}^{(l)}(n, 1)\right)$. Hence,

$$\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} = \frac{(2m - 1)N - 1}{2}\psi_l, \tag{22}$$

where the $\mathbf{b}_{k,m}^{(l)}(n, 1)$ is the nth entry of $\mathbf{b}_{k,m}^{(l)}$, $1 \leq n \leq N$, and the last equality in (22) follows from the facts that $\mathbf{b}_{k,m}^{(l)}(n, 1) = -\zeta_k\gamma_{n,m}^{(l)}$ and $\gamma_{n,m}^{(l)} = ((m - 1)N + n - 1)\psi_l$.

Based on (20) and (22), the optimal solution of the problem in (18) is summarized below.

***Theorem* 1.** The optimal solution $\mathbf{a}_m^{(l)\star} = [(\mathbf{x}_m^{(l)\star})^T, \vartheta_m^{(l)\star}]^T$ to (18) is given by

$$x_{n,m}^{(l)\,\star} = \begin{cases} \dfrac{N - 2n + 1}{2}\psi_l, & \text{if } 0 \leq \psi_l \leq \frac{4f_ct_{\max}}{(2m-1)N-1}, \quad (23a) \\ \vartheta_{\max} - \gamma_{n,m}^{(l)}, & \text{otherwise}, \quad (23b) \end{cases}$$

for $1 \leq l \leq N_{RF}$, $1 \leq n \leq N$, $1 \leq m \leq M$, and $\vartheta_m^{(l)\star} = 2f_ct_m^{(l)\star}$, where the $t_m^{(l)\star}$ is

$$t_m^{(l)\,\star} = \begin{cases} \dfrac{(2m-1)N-1}{4f_c}\psi_l, & \text{if } 0 \leq \psi_l \leq \frac{4f_ct_{\max}}{(2m-1)N-1}, \quad (24a) \\ t_{\max}, & \text{otherwise}. \quad (24b) \end{cases}$$

Proof. See Appendix B. $\qquad\square$

Remark 3. The solutions in (24) and (23) differentiate them from prior approaches to solving related problems. For instance, the PS values in [10] were not optimized and given by $x_{n,m}^{(l)} = -(n - 1)\psi_l$, $\forall n$. The time delay value of the mth TTD was $t_m^{(l)} = m\frac{N\psi_l}{2f_c}$ in [10]; as m increases, the $t_m^{(l)}$ could be larger than $t_{\max}$, in which case such $t_m^{(l)}$ needs to be floored to the $t_{\max}$, resulting in performance deterioration. As will be discussed in Section V, when all TTD values are smaller

than $t_{\max}$, the approaches in [10], [12] achieve the same array gain performance as the proposed approach, meaning that the designs in [10], [12] is a special case of Theorem 1.

Remark 4. The proposed approach takes advantage of the block-diagonal structure of PS and TTD precoders to decompose the MIMO precoding problem in (14) into multiple single-RF chain precoding design problems in (18). For an array architecture that does not allow introducing a tractable block-diagonal structure of the precoders, a general subspace decomposition approach for the hybrid precoders design in [21], [23] can be extended to alternatively optimize the digital precoder $\{\mathbf{W}_k^\star\}$, PS precoder $\mathbf{F}_1^\star$, and TTD precoder $\{\mathbf{F}_{2,k}^\star\}$. However, we note that satisfying the constraint in (14e) without imposing some tractable structure into $\mathbf{F}_1$ and $\{\mathbf{F}_{2,k}\}$ is still challenging. Investigation of ways of addressing general hybrid MIMO precoding architecture is subject to further research and not the focus of the present work.

In the following, the benefits of the proposed joint TTD and PS precoder optimization method are discussed, characterizing system parameters for the best beam squint compensation, which was unexplored by the prior works [10], [43]. The optimal condition in (23a) and (24a) can be rewritten as

$$\frac{(2m-1)N-1}{4f_c}\psi_l \leq t_{\max}. \tag{25}$$

When (25) is satisfied, the beam squint is compensated effectively. In this case, the squinted beams are shifted to be aligned with spatial paths. However, the condition (25) may be violated, for example, when either $t_{\max}$ becomes small or N becomes large (i.e., when N_t tends to be large while fixing M, i.e., $N = \frac{N_t}{M}$). This case results in beam misalignment at some subcarrier frequencies, which could cause array gain degradation. Based on the condition in (25), we obtain selection criteria (rule of thumb) on the required number of transmit antennas (N_t) and the value of maximum time delay ($t_{\max}$) for the best beam squint compensation as follows.

1) N_t Selection Criterion

Given the number of TTDs per RF chain M and the $t_{\max}$ values determined by the employed TTD devices, choose N_t such that

$$N_t \leq \frac{M}{2m-1} + \frac{4M}{(2m-1)}\frac{1}{\psi_l}f_c t_{\max}, \forall l, m, \tag{26}$$

where (26) is a result of substituting $N = \frac{N_t}{M}$ into (25). The inequality in (26) is rewritten as

$$N_t \leq \min_{l,m}\left(\frac{M}{2m-1} + \frac{4M}{(2m-1)}\frac{1}{\psi_l}f_c t_{\max}\right)$$
$$= \frac{M}{2M-1} + \frac{4M}{(2M-1)}\frac{1}{\max_l \psi_l}f_c t_{\max}, \tag{27}$$

where the $\min_{l,m}$ in (27) is taken so that (26) holds for all MN_{RF} TTDs. The last equality in (27) follows from substituting $m = M$ and ψ_l by its maximum value $\max_l \psi_l$.

2) $t_{\max}$ Selection Criterion

Equivalently, given the number of TTDs per RF chain M and the number of transmit antennas N_t, the TTD value should be chosen to satisfy

$$t_{\max} \geq \psi_l \frac{(2m-1)N_t - M}{4M}\frac{1}{f_c}, \forall l, m. \tag{28}$$

The inequality in (28) can be rewritten as

$$t_{\max} \geq \max_{l,m}\left(\psi_l \frac{(2m-1)N_t - M}{4M}\frac{1}{f_c}\right)$$
$$= (\max_l \psi_l)\frac{(2M-1)N_t - M}{4M}\frac{1}{f_c}. \tag{29}$$

For example, setting $f_c = 300$ GHz and $M = 16$, the right-hande side (RHS) of (27) increases linearly from 186 to 743 with respect to $t_{\max}$ with the slope being 6.1935×10^{11} as $t_{\max}$ in (27) grows from 300 ps to 1200 ps. Similarly, the RHS of (29) increases linearly from 412.5 ps to 1135.8 ps with the slope 1.6146×10^{-12} as N_t in (29) grows from 256 to 720.

Remark 5. We note that since the receiver is less likely in angle $\max_l \psi_l$, one can relax $\max_l \psi_l$ in (27) and (29) to $\varrho \max_l \psi_l$ ($0 < \varrho < 1$) to increase the upper bound of N_t to $\frac{M}{2M-1} + \frac{4M}{(2M-1)}\frac{1}{\varrho \max_l \psi_l}f_c t_{\max}$. This allows us to bring a larger N_t, which readily improves the average achievable rate performance.

C. Array Gain Loss

Next, we characterize the array gain when (29) is satisfied and array gain loss when it is not. First, when $t_{\max}$ meets the optimal condition in (29), we compute the array gain $g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ defined in (10) using the optimal PS values in (23a) and TTD values in (24a), i.e., $g(\mathbf{f}_k^{(l)}, \psi_{k,l}) = \frac{1}{N_t}\left|\sum_{m=1}^{M}\sum_{n=1}^{N}e^{j\pi\zeta_k\gamma_{n,m}^{(l)}}e^{j\pi\frac{N-2n+1}{2}\psi_l}e^{-j\pi\zeta_k\frac{(2m-1)N-1}{2}\psi_l}\right|$, yielding

$$g(\mathbf{f}_k^{(l)}, \psi_{k,l}) = \left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right|. \tag{30}$$

When the condition in (29) is not satisfied, the array gain performance is deteriorated. This can be quantified in terms of the array gain loss at the lth RF chain $\left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| - g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ as shown in the following remark.

Remark 6. Suppose that the first $M' \leq M$ TTDs are given by (24a) and the last $(M - M')$ are given by (24b), i.e., $\psi_l \frac{(2M'-1)N_t - M}{4M}\frac{1}{f_c} \leq t_{\max} \leq \psi_l \frac{(2(M'+1)-1)N_t - M}{4M}\frac{1}{f_c}$. Then, the following holds,

$$0 \leq \left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| - g(\mathbf{f}_k^{(l)}, \psi_{k,l}) \leq \frac{M-M'}{M}\left(\left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| + 1\right). \tag{31}$$

The proof of (31) is relegated to Appendix C. The bound in (31) implies that the array gain loss becomes zero when $M' = M$, which occurs when $t_{\max}$ is sufficiently large to meet the optimal condition in (29) and the TTD values $\{t_m^{(l)\star}\}_{m=1}^{M}$ are given by (24a). The array gain loss upper bound decreases by the factor of $\frac{1}{M}\left(\left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| + 1\right)$ when M' increases to

$(M' + 1)$ while satisfying $\psi_l \frac{(2(M'+1)-1)N_t - M}{4M} \frac{1}{f_c} \leq t_{\max} \leq \psi_l \frac{(2(M'+2)-1)N_t - M}{4M} \frac{1}{f_c}$.

The TTD-based hybrid precoding architecture in Fig. 1c utilizes N_{RF} RF chains, MN_{RF} TTDs, and $N_{RF}N_t$ PSs, which could increase hardware complexity substantially. Reducing the hardware complexity of the architecture in Fig. 1c while maintaining the beam squint compensation performance is an important research problem and is subject to further study. In this connection, the system parameter selection criteria in (27) and (29) give an idea of the values of N_t and $t_{\max}$ for the best beam squint compensation given a fixed number of TTDs per RF chain M. In practice, it is also critical to install an appropriate value of M to reduce the power consumption of the analog precoder, which is addressed in the next subsection.

D. M Selectrion Criterion

In this subsection, we further exploit the closed-form solution of the TTD and PS precoders in Theorem 1 to characterize the minimum number of TTDs per RF chain M given a predefined array gain performance while assuming that the constraint in (29) is always satisfied and the number of transmit antennas N_t is fixed. To this end, we formulate a problem that minimizes M to guarantee a given array gain performance:

$$\min_M M \text{ subject to } N_t = MN, \ g(\mathbf{f}_k^{(l)}, \psi_{k,l}) \geq g_0, \forall k, l, \quad (32)$$

where the second constraint in (32) describes the array gain guarantee with the threshold $0 < g_0 < 1$. We note that the total power consumption of the analog precoder can be modeled as $P_{total} = N_{RF}MP_{TTD} + N_{RF}N_tP_{PS}$, where P_{TTD} and P_{PS} are the power consumption of a TTD and a PS, respectively. Hence, given the fixed values of N_t, P_{TTD}, and P_{PS}, the objective of minimizing the power consumption (i.e., $\min_M P_{total}$) is equivalent to the objective in (32).

Incorporating $g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ in (30) into the second constraint in (32) and substituting $\frac{N_t}{M}$ for N, the problem is converted to:

$$\min_M M \text{ subject to } \left| \frac{\sin\left(\frac{N_t}{M}\Delta_{k,l}\right)}{\frac{N_t}{M}\sin(\Delta_{k,l})} \right| \geq g_0, \forall k, l. \quad (33)$$

The problem in (33) can be viewed as finding the minimum element in the feasible set $\mathcal{M} = \left(\bigcap_{k=1,l=1}^{K,L} \mathcal{M}_{k,l} \right) \cap \mathcal{N}_t$, where $\mathcal{M}_{k,l} = \left\{ M \in \mathbb{R}: \left| \frac{\sin(\frac{N_t}{M}\Delta_{k,l})}{\frac{N_t}{M}\sin(\Delta_{k,l})} \right| \geq g_0 \right\}$, for $1 \leq k \leq K$ and $1 \leq l \leq L$, and $\mathcal{N}_t = \{m \in \mathbb{N}: m|N_t\}$, where $m|N_t$ denotes that m is a divisor of N_t. Thus, solving the problem (33) using the greedy search method requires the construction of $KL+1$ sets $\{\mathcal{M}_{k,l}\}$ and $\mathcal{N}_t$, which is demanding as the number of OFDM subcarriers grows.

To cope with these difficulties, we propose to approximate the constraint in $\mathcal{M}_{k,l}$, i.e., the constraint in (33) as $|q(\Delta_{k,l})| \geq g_0, \forall k, l$, where $q(x) = 1 + \frac{1}{6}\left(1 - \frac{N_t^2}{M^2}\right)x^2$, which is obtained by the Taylor expansion $\frac{\sin\left(\frac{N_t}{M}x\right)}{\frac{N_t}{M}\sin(x)} = q(x) + \mathcal{O}(x^4)$ derived at $x = 0$. Noting that $|q(\Delta_{k,l})| \geq q(\Delta_{k,l})$, the problem

in (33) is relaxed to

$$M^\star = \underset{M}{\operatorname{argmin}} M \text{ subject to } M \in \mathcal{N}_t, \ q(\Delta_{k,l}) \geq g_0, \forall k, l. \quad (34)$$

About the optimal solution to (34), denoted as $M^\star$, we have the following theorem.

Theorem 2. Given the fixed N_t, B, f_c, and K, the $M^\star$ of (34) is $M^\star = \left\lceil \sqrt{\frac{N_t^2}{1+\Omega(g_0,B)}} \right\rceil_{N_t}$, where $\Omega(g_0, B) = \frac{6(1-g_0)}{\left(\frac{\pi}{2}\frac{B}{f_c}\frac{K-1}{2K}\right)^2 \max_l \psi_l^2}$ and $\lceil x \rceil_{N_t}$ denotes the smallest integer greater than or equal to x that is a divisor of N_t.

Proof. See Appendix D. $\qquad\square$

Remark 7. To understand the relationship between the $M^\star$ in Theorem 2 and the system parameters, we first relax its integer constraint, leading to $M^\star = \sqrt{\frac{N_t^2}{1+\Omega(g_0,B)}}$. In the regime of a large number of OFDM subcarriers ($K \gg 1$), we have $\frac{K-1}{2K} \approx \frac{1}{2}$, leading to $M^\star \approx \sqrt{\frac{N_t^2(\frac{\pi}{4}\frac{B}{f_c})^2 \max_l \psi_l^2}{6(1-g_0)}} = \left(\frac{\pi N_t}{4f_c}\sqrt{\frac{\max_l \psi_l^2}{6(1-g_0)}}\right)B$. Given a fixed N_t and $t_{\max}$, it reveals that the minimum required number of TTDs that ensures a predefined array gain performance grows linearly with the bandwidth B. For instance, setting $N_t = 720$ and $t_{\max} = 1200$ ps meets the optimal condition in (31), the required number of TTDs $M^\star$ increases approximately linearly from 20 to 180 as B grows from 10 GHz to 80 GHz with the slope 2.4335×10^{-9} for $g_0 = 0.9$.

The proposed approach embodies a fundamental engineering design principle: maximizing system performance within existing hardware constraints. Striking this balance is crucial as it not only improves resource utilization but also enhances the overall performance and feasibility of TTD-based hybrid precoding systems. The proposed approach aligns with the resource-conscious design principle.

V. SIMULATION RESULTS

In this section, we demonstrate the benefits of the proposed joint TTD and PS precoding by comparing it with the state-of-the-art closed-form solution in [10] and iterative design with a least square solution in [43, Algorithm 1] in terms of array gain and computation time required to design TTD and PS values. We set the number of iterations for the benchmark [43, Algorithm 1] as $N_{iter} = 10$ to alternatingly design PS and TTD values. The numerical results are conducted on an Intel(R) Core(TM) i7-9850H CPU using MATLAB 2023a.

Throughout the simulation, the system parameters are set as follows unless otherwise stated: the central carrier frequency $f_c = 300$ GHz, bandwidth $B = 30$ GHz, the number of OFDM subcarriers $K = 129$, number of transmit antennas $N_t = 256$, number of receive antennas $N_r = 4$, number of RF chain $N_{RF} = 4$, and number of data stream $N_s = 4$. The number of TTDs per RF chain $M = 16$, the number of PSs per TTD $N = 16$, and the maximum time delay $t_{\max} = 320$ ps. The AoDs and AoAs are uniformly drawn in $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$.
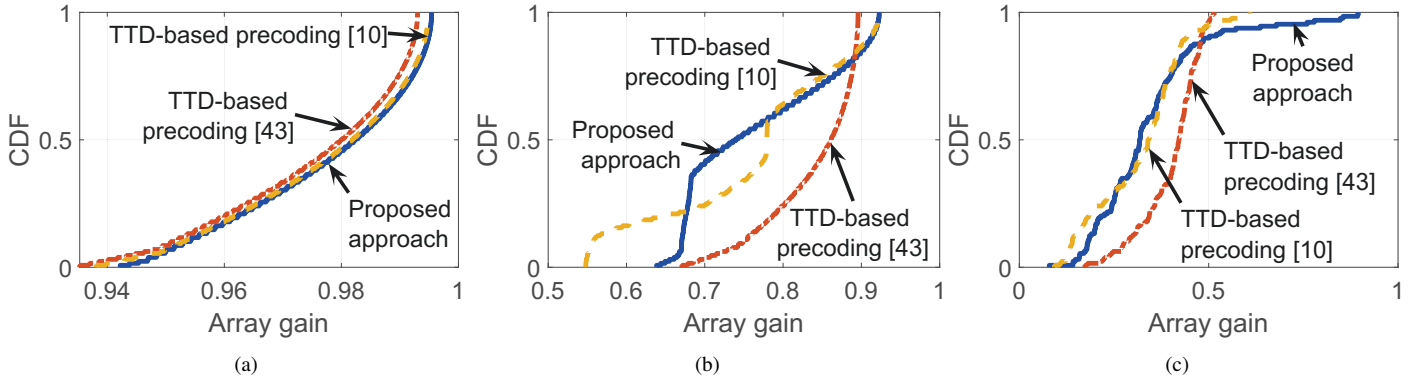
Fig. 3. CDF of the array gain with different number of transmit antennas (N_t) values: (a) $N_t = 128$, (b) $N_t = 256$, and c) $N_t = 512$.
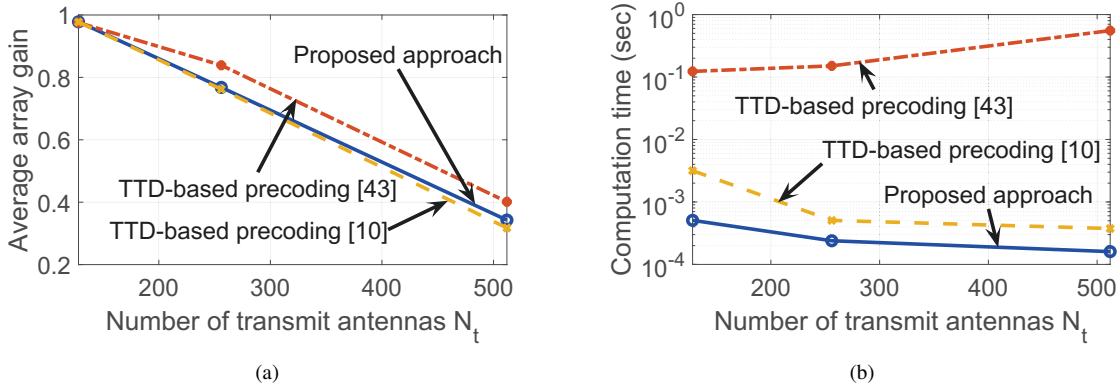

Fig. 4. (a) Average array gain vs. N_t and (b) computation time to design PS and TTD values when $N_t \in \{128, 256, 512\}$.

A. Array Gain Performance and Computation Time

Although the derivations in the paper assume continuous PS and TTD values, the following numerical evaluations adopt finite-resolution PS and TTD values to be realistic. In particular, we specify the set of feasible PS values as $\mathcal{P}$ with 4-bit quantization, represented by $\mathcal{P} = \{0, \frac{2\pi}{16}, \dots, \frac{15}{16}2\pi\}$. Similarly, the set of feasible TTD values $\mathcal{T}$ is defined with a 2-ps-step size, such that $\mathcal{T} = \{0, 2, \dots, 318, 320\}$ ps, accommodating the practical limitations of TTD devices. Quantization is applied directly to the PS and TTD values derived in Theorem 1, aligning them with the nearest values in $\mathcal{P}$ and $\mathcal{T}$, respectively, to obtain a solution of discrete PS and TTD values. To ensure a fair evaluation, this quantization method is also applied to the closed-form solution in benchmark [10] and the iterative least squares solution detailed in [43, Algorithm 1].

We measure the empirical cumulative distribution function (CDF) of the array gain $G_\psi(x)$ and the average array gain $\frac{1}{K} \sum_{k=1}^{K} g(\mathbf{f}_k^{(l)}, \psi_k)$ at the central spatial direction ψ. We also measure the corresponding computation time to design PS and TTD values. Herein, the CDF of the array gain is computed as $G_\psi(x) = \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_g(k, x)$, where x represents the array gain $0 \leq x \leq 1$ in (4), the indicator function $\mathbb{1}_g(k, x)$ takes values 1 if $g(\mathbf{f}_k^{(l)}, \psi_k) \leq x$ and 0 otherwise, and $g(\mathbf{f}_k^{(l)}, \psi_k)$ is computed as in (4) with the spatial direction $\psi = 0.9$.

Fig. 3 displays the empirical CDF curves of the array gain when the number of transmit antennas N_t takes the values from $\{128, 256, 512\}$. When $N_t = 128$, the proposed approach exhibits a superior array gain performance compared to the benchmarks [10], [43]. When $N_t = 256$ (Fig. 3b) and $N_t = 512$ (Fig. 3c), the proposed approach and benchmarks [10], [43] suffer from the array gain loss because the degree of beam squint increases as N_t grows (i.e., Proposition 1). Nevertheless, the proposed approach provides an enhanced beam squint compensation capability compared to the benchmark [10]. This is due to the fact that as N_t grows, some of the time delay values of the benchmark [10] become larger than $t_{\max} = 320$ ps, in which case the time delay value needs to be quantized to the $t_{\max}$. Yet, when $N_t = 256$ and $N_t = 512$, the benchmark [43] outperforms the proposed approach and benchmark [10]. Notably, in Fig. 3c, approximately 65% of OFDM subcarriers of the benchmark [43] achieve array gains ≥ 0.4, while only 25% of OFDM subcarriers of the proposed approach achieve similar performance.

Fig. 4a displays the corresponding average array gain performances of the array gain CDF curves displayed in Fig. 3. Observations from Fig. 4a reveal the proposed approach achieves slightly higher average array gain than the benchmarks [10], [43] when $N_t = 128$. When beam squint becomes more severe
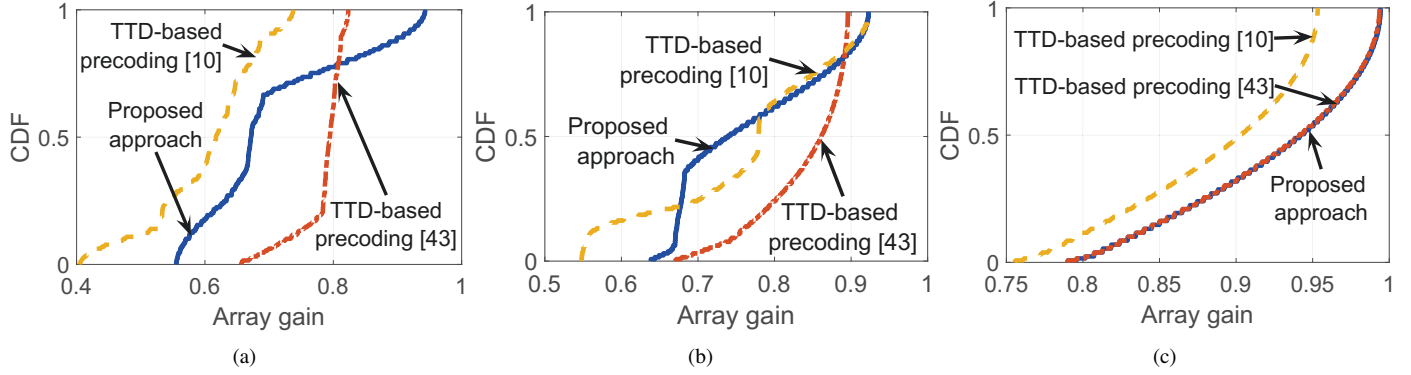
Fig. 5. CDF of array gain with different maximum time delay (t_max) values: (a) $t_\mathrm{max}=280$ ps, (b) $t_\mathrm{max}=320$ ps, and (c) $t_\mathrm{max}=380$ ps.
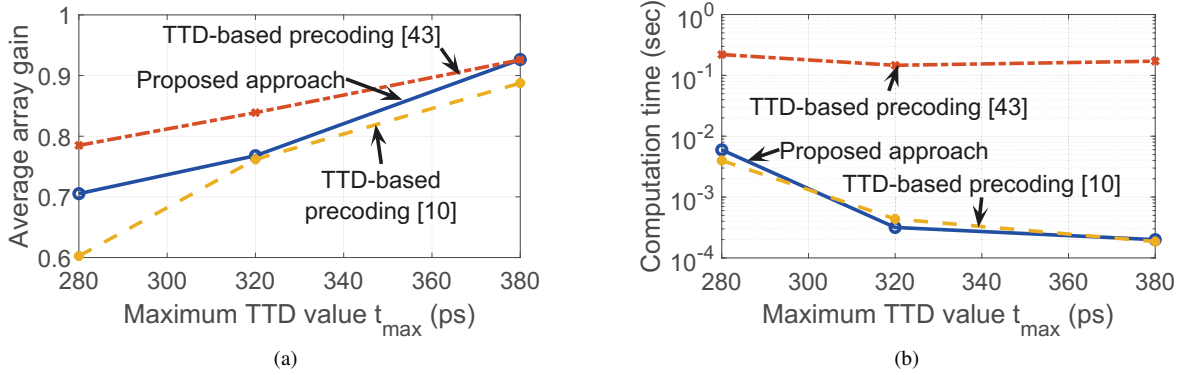

Fig. 6. (a) Average array gain vs. t_max and (b) computation time to design PS and TTD values when $t_\mathrm{max} \in \{280, 320, 380\}$ ps and TTD step size 2 ps.

($N_t = 256$ and $N_t = 512$), the benchmark [43] achieves a moderately higher average array gain than our proposed approach and benchmark [10]. This is also consistent with the CDF curves in Fig. 3c.

Fig. 4b illustrates the computation time required by our proposed approach and the benchmarks [10], [43] to achieve the average array gain performances shown in Fig. 4a. The comparison reveals that the proposed approach and benchmark [10] maintain a rather consistent time complexity trend as N_t grows, underscoring the efficiency of one-shot low-complexity designs. Conversely, the computation time of the benchmark [43] increases substantially with N_t, attributed to its iterative design process which necessitates matrix inversion at each step—a computationally intensive operation. For instance, at $N_t = 256$, the proposed approach records a swift computation time of 2.4×10^{-4} sec, markedly faster compared to the 0.1508 sec required by the benchmark [43]. This highlights the computational advantage of our one-shot design over iterative methods that involve complex calculations for each round of iteration.

Fig. 5 demonstrates the empirical CDF curves of the array gain when the maximum time delay t_max takes values in $\{280, 320, 380\}$ ps while fixing $N_t = 256$ and setting the TTD step size to 2 ps. It is observed in Fig. 5 that the proposed approach outperforms the benchmark [10]. For example, it can

be observed from Fig. 5a that 33% of OFDM subcarriers of the proposed approach achieve ≥ 0.7 array gain compared to that 10% of OFDM subcarriers of the benchmark [10] attain the similar performance. Similar to the trend in Figs. 3b-3c, it can be observed from Figs. 5a-5b that the benchmark [43] outperforms the proposed approach when the maximum TTD value is not sufficient to compensate for the beam squint. The proposed approach and the benchmark [43] converge to the same array gain performance when $t_\mathrm{max} = 380$ (ps).

Fig. 6a demonstrates the average array gain performance of the corresponding CDF curves shown in Fig. 5. Seen from Fig. 6a, the proposed approach exhibits a superior average array gain performance compared to the benchmark [10]. When $t_\mathrm{max} = 380$ ps, the proposed approach and the benchmark [43] converge to a similar average array gain performance. These observations are consistent with the trends in Fig. 5.

Fig. 6b illustrates the computation time needed by the proposed approach and benchmarks [10], [43] to achieve the average array gain performance depicted in Fig. 6a. In alignment with the trends observed in Fig. 4b, Fig. 6b corroborates that the proposed approach and the benchmark [10] demand significantly less computation time compared to the benchmark [43]. Notably, at $t_\mathrm{max} = 380$ ps, the benchmark [43] necessitates a computation time of 0.146 sec while the proposed approach achieves comparable performance with an

impressively reduced computation time of only 3.16×10^{-4} sec.

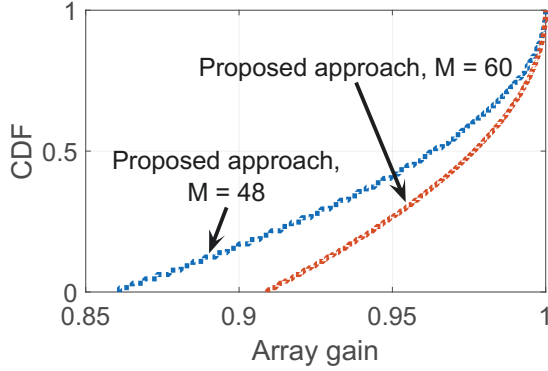### B. Array Gain Performance Guarantee of $\widetilde{M}^\star$ in Theorem 2

Fig. 7. CDF of array gain of the proposed approach with different number of TTDs per RF chain (M) values.

In this subsection, we demonstrate the array gain performance guarantee provided by the minimum required number of TTDs per RF chain $\widetilde{M}^\star$ in Theorem 2. Similar to Section V-A, we evaluate the empirical CDF of the array gain $G_\psi(x)$ at the spatial direction $\psi = 0.8$ except that we assume $N_t = 720$ and $t_{\max} = 1000$ ps in this simulation with 8-bit quantization for the PS values and 2-ps-step size for the TTD values. The array gain threshold is set to $g_0 = 0.9$ in Theorem 2. Incorporating g_0 and the system parameters into Theorem 2 yields $\widetilde{M}^\star = \left\lceil \sqrt{\frac{720^2}{1+\Omega(0.9,30)}} \right\rceil_{720} = 60$.

Fig. 7 demonstrates the CDF of the array gain when $M = \widetilde{M}^\star = 60$ and $M = 48$, where $M = 48$ is chosen to be the largest divisor of N_t that is smaller than $\widetilde{M}^\star = 60$. We note that the condition in (31) is satisfied for both values of M when $t_{\max} = 1000$ ps. As shown in Fig. 7, for the optimized $\widetilde{M}^\star = 60$, every OFDM subcarrier satisfies the array gain ≥ 0.9. However, when $M = 48$, there are 18% of the OFDM subcarriers that have the array gain < 0.9. Hence, the curves in Fig. 7 verify that the $\widetilde{M}^\star = 60$ is the minimum number of TTDs per RF chain to provide the minimum array gain $g_0 = 0.9$.

VI. CONCLUSION

We presented a new framework to the problem of compensating the beam squint effect arising in wideband sub-THz hybrid massive MIMO systems. We determined the ideal analog precoder that fully compensates for the beam squint. A novel TTD-based hybrid precoding approach was proposed by jointly optimizing the TTD and PS precoders under the per-TTD time delay constraints. The joint optimization problem was formulated in the context of minimizing the distance between the ideal analog precoder and the product of the PS and TTD precoders. By transforming the original problem into the phase domain, the original problem was converted to an approximated problem, which allowed us to find a closed-form solution. Based on the closed-form expression of our solution, we presented the selection criteria for the required number of transmit antennas and the value of maximum time delay. Exploiting the proposed joint TTD and PS precoder optimization approach, we determined the minimum number of TTDs required to achieve an array gain performance guarantee while minimizing the analog precoder power consumption. Simulations corroborated the advantages of our joint TTD and PS optimization approach by demonstrating guaranteed array gain performance and reduced computational time.

APPENDIX A
PROOF OF LEMMA 1

Without loss of generality, we assume $0 < |x - y| < \pi$. Then, $\mathrm{argmin}_{0<|x-y|<\pi}|e^{jx} - e^{jy}| = \mathrm{argmin}_{0<|x-y|<\pi}\left| \sin\left(\frac{x-y}{2}\right) \right| = \mathrm{argmin}_{0<|x-y|<\pi} \sin\left(\left| \frac{x-y}{2} \right| \right) = \mathrm{argmin}_{0<|x-y|<\pi} |x - y|$, where the last step follows from the fact that $\sin(z)$ is an increasing function of $0 < z < \frac{\pi}{2}$. This completes the proof.

APPENDIX B
PROOF OF THEOREM 1

To prove Theorem 1, we start by formulating the Lagrangian of (18), which is given by

$$\mathcal{L}(\mathbf{a}_m^{(l)}, \lambda_1, \lambda_2) = \mathbf{a}_m^{(l)T}\mathbf{C}\mathbf{a}_m^{(l)} - 2\mathbf{d}_m^{(l)T}\mathbf{a}_m^{(l)} + \lambda_1(\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} - \vartheta_{\max}) + \lambda_2(-\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)}), \quad (35)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the Largrangian multipliers. After incorporating the first order necessary condition for $\mathbf{a}_m^{(l)}$ in (35), the Karush–Kuhn–Tucker (KKT) conditions of (18) are given by:

$$\begin{cases} 2\mathbf{C}\mathbf{a}_m^{(l)} - 2\mathbf{d}_m^{(l)} + \lambda_1\mathbf{e}_{N+1} - \lambda_2\mathbf{e}_{N+1} &= 0, \\ \lambda_1(\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} - \vartheta_{\max}) &= 0, \\ \lambda_2(-\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)}) &= 0, \\ \lambda_1 \geq 0, \lambda_2 \geq 0. \end{cases} \quad (36)$$

When $\lambda_1 = \lambda_2 = 0$, (36) yields $\mathbf{a}_m^{(l)} = \mathbf{C}^{-1}\mathbf{d}_m^{(l)}$ for $0 \leq \mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} \leq \vartheta_{\max}$ because $0 \leq \mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} \leq \vartheta_{\max}$. When $\lambda_1 > 0$ and $\lambda_2 = 0$, (36) gives $\mathbf{a}_m^{(l)} = \mathbf{C}^{-1}(\mathbf{d}_m^{(l)} - \frac{1}{2}\lambda_1\mathbf{e}_{N+1})$ for $\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} > \vartheta_{\max}$ because $\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} = \vartheta_{\max}$ and $\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} = \mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} - \frac{1}{2}\lambda_1\frac{1}{\eta}$, where the last equality follows from (20). When $\lambda_1 = 0$ and $\lambda_2 > 0$, we obtain $\mathbf{a}_m^{(l)} = \mathbf{C}^{-1}(\mathbf{d}_m^{(l)} + \frac{1}{2}\lambda_2\mathbf{e}_{N+1})$ for $\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} < 0$ because $\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} = 0$ and $\mathbf{e}_{N+1}^T\mathbf{a}_m^{(l)} = \mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} + \frac{1}{2}\lambda_2\frac{1}{\eta}$, where the last equality follows from (20). However, by (22) we have $\mathbf{e}_{N+1}^T\mathbf{C}^{-1}\mathbf{d}_m^{(l)} = \frac{(2m-1)N-1}{2}\psi_l \geq 0$, where the last inequality is due to the sign invariance property in (11). Therefore, the

last case leads to a contradiction. In summary, solving (36) gives

$$
\mathbf{a}_m^{(l)\star} = \begin{cases} \mathbf{C}^{-1}\mathbf{d}_m^{(l)}, & \text{if } 0 \le \frac{(2m-1)N-1}{2}\psi_l \le \vartheta_{\max}, \quad (37a) \\ \mathbf{C}^{-1}\left(\mathbf{d}_m^{(l)} - \frac{1}{2}\lambda_1\mathbf{e}_{N+1}\right), & \text{otherwise}, \quad (37b) \end{cases}
$$

where $\lambda_1 = 2\eta\left(\frac{(2m-1)N-1}{2}\psi_l - \vartheta_{\max}\right)$.

To further simplify the closed-form solution in (37), we deduce $\mathbf{a}_m^{(l)\star}$ in (37a) by simplifying $\mathbf{C}^{-1}\mathbf{d}_m^{(l)}$. Using (21) gives

$$
\mathbf{C}^{-1}\mathbf{d}_m^{(l)} = \begin{bmatrix} \frac{1}{K}\sum_{k=1}^{K}\mathbf{b}_{k,m}^{(l)} + \frac{1}{\eta}\frac{1}{K}\sum_{k=1}^{K}(1-\zeta_k)\mathbf{1}_N\mathbf{1}_N^T\mathbf{b}_{k,m}^{(l)} \\ \frac{1}{\eta}\frac{1}{K}\sum_{k=1}^{K}(1-\zeta_k)\mathbf{1}_N^T\mathbf{b}_{k,m}^{(l)} \end{bmatrix}. \quad (38)
$$

Using the definition of $\mathbf{b}_{k,m}^{(l)} = [-\zeta_k\gamma_{1,m}^{(l)} \ \cdots \ -\zeta_k\gamma_{N,m}^{(l)}]^T$, the nth entry of $\widetilde{\mathbf{d}}_m^{(l)} \triangleq \mathbf{C}^{-1}\mathbf{d}_m^{(l)}$ in (38), where $1 \le n \le N$, is given by

$$
\begin{aligned}
\widetilde{\mathbf{d}}_m^{(l)}(n,1) &= \frac{1}{K}\gamma_{n,m}^{(l)}\sum_{k=1}^{K}-\zeta_k + \frac{1}{\eta K}\sum_{n=1}^{N}\gamma_{n,m}^{(l)}\sum_{k=1}^{K}(\zeta_k^2-\zeta_k) \\
&= \frac{N-2n+1}{2}\psi_l,
\end{aligned} \quad (39)
$$

where the last equality in (39) is due to (19) and the fact that $\gamma_{n,m}^{(l)} = ((m-1)N+n-1)\psi_l$. The $(N+1)$th entry of $\widetilde{\mathbf{d}}_m^{(l)}$ in (38) is

$$
\widetilde{\mathbf{d}}_m^{(l)}(N+1,1) = \frac{1}{N}\sum_{n=1}^{N}\gamma_{n,m}^{(l)} = \frac{(2m-1)N-1}{2}\psi_l. \quad (40)
$$

Therefore, substituting (39) and (40) into (37a) leads to (23a) and (24a), respectively. Next, (37b) can be rewritten as

$$
\mathbf{C}^{-1}(\mathbf{d}_m^{(l)} - \frac{1}{2}\lambda_1\mathbf{e}_{N+1}) = \widetilde{\mathbf{d}}_m^{(l)} - \left(\frac{(2m-1)N-1}{2}\psi_l - \vartheta_{\max}\right)\mathbf{1}_{N+1}. \quad (41)
$$

Incorporating (39) and (40) into (41) leads to (23b) and (24b), respectively, completing the proof.

Appendix C
Proof of Remark 6

From the assumption in Remark 6, we have $t_m^{(l)\star} = \frac{(2m-1)N-1}{4f_c}\psi_l$, for $m = 1,\ldots,M'$ and $t_m^{(l)\star} = t_{\max}$ for $m = M'+1,\ldots,M$ at the lth RF chain. The PS values are, respectively, given by (23a) and (23b), $x_{n,m}^{(l)\star} = \frac{N-2n+1}{2}\psi_l$, for $m = 1,\ldots,M'$ and $x_{n,m'}^{(l)\star} = \vartheta_{\max} - \gamma_{n,m}^{(l)}$, for $m' = M'+1,\ldots,M$. Then, the array gain $g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ defined in

(4) is given by

$$
\begin{aligned}
&g(\mathbf{f}_k^{(l)}, \psi_{k,l}) \\
&= \frac{1}{N_t}\Bigg| \sum_{m=1}^{M'}\sum_{n=1}^{N} e^{j\pi\xi_k\gamma_{n,m}^{(l)}} e^{j\pi\frac{N-2n+1}{2}f_c} e^{-j\pi\xi_k\frac{(2m-1)N-1}{2}\psi_l} \\
&\quad + \sum_{m'=M'+1}^{M}\sum_{n=1}^{N} e^{j\pi\xi_k\gamma_{n,m'}^{(l)}} e^{j\pi(\vartheta_{\max}-\gamma_{n,m'}^{(l)})} e^{-j\pi\xi_k\vartheta_{\max}} \Bigg| \\
&= \frac{1}{N_t}\Bigg| M'\frac{\sin(N\Delta_{k,l})}{\sin(\Delta_{k,l})} + e^{-j\pi(\xi_k-1)\vartheta_{\max}}\sum_{m'=M'+1}^{M}\sum_{n=1}^{N} e^{j\pi(\xi_k-1)\gamma_{n,m'}^{(l)}} \Bigg|.
\end{aligned} \quad (42)
$$

Applying the triangle inequality to (42) yields

$$
\frac{M'}{N_t}\left|\frac{\sin(N\Delta_{k,l})}{\sin(\Delta_{k,l})}\right| - \frac{(M-M')N}{N_t} \le g(\mathbf{f}_k^{(l)}, \psi_{k,l}), \quad (43)
$$

where $\frac{M'}{N_t}\left|\frac{\sin(N\Delta_{k,l})}{\sin(\Delta_{k,l})}\right|$ is the array gain achieved by the first M' TTD precoding $\{t_m^{(l)\star}\}_{m=1}^{M'}$ and $\frac{(M-M')N}{N_t}$ is an upper bound of the array gain obtained by the last $(M-M')$ TTDs precoding $\{t^{(l)\star}\}_{m=M'+1}^{M}$. From (43) and noting that $N_t = MN$, the array gain loss $\left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| - g(\mathbf{f}_k^{(l)}, \psi_{k,l})$ is bounded by

$$
\begin{aligned}
0 &\le \left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| - g(\mathbf{f}_k^{(l)}, \psi_{k,l}) \\
&\le \frac{M-M'}{M}\left|\frac{\sin(N\Delta_{k,l})}{N\sin(\Delta_{k,l})}\right| + \frac{(M-M')}{M},
\end{aligned} \quad (44)
$$

which completes the proof.

Appendix D
Proof of Theorem 2

The second constraint in (34) can be rewritten as

$$
\begin{aligned}
g_0 &\le 1 + \frac{1}{6}\left(1 - \frac{N_t^2}{M^2}\right)\max_{k,l}\left(\frac{\pi}{2}\frac{B}{f_c}\left(\frac{k-1-\frac{K-1}{2}}{K}\right)\psi_l\right)^2 \\
&= 1 + \frac{1}{6}\left(1 - \frac{N_t^2}{M^2}\right)\left(\frac{\pi}{2}\frac{B}{f_c}\frac{K-1}{2K}\right)^2\max_l\psi_l^2,
\end{aligned} \quad (45)
$$

where the upper bound in (45) follows from the substitution of $\Delta_{k,l} = \frac{\pi}{2}(\zeta_k-1)\psi_l$ with ζ_k defined in (3) and the last equality in (45) is due to the substitutions $k = K$ and $\psi_l^2 = \max_l\psi_l^2$. Then, the inequality in (45) can be rewritten as $M \ge \sqrt{\frac{N_t^2}{1+\Omega(g_0,B)}}$, where $\Omega(g_0,B) = \frac{6(1-g_0)}{\left(\frac{\pi}{2}\frac{B}{f_c}\frac{K-1}{2K}\right)^2\max_l\psi_l^2}$, leading to $M^\star = \left\lceil\sqrt{\frac{N_t^2}{1+\Omega(g_0,B)}}\right\rceil_{N_t}$ because $M \in \mathcal{N}_t$. This completes the proof.

References

[1] D. Q. Nguyen and T. Kim, "Joint delay and phase precoding under true-time delay constraints for THz massive MIMO," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 3496–3501.

[2] ——, "Joint hybrid delay-phase precoding under true-time delay constraints in wideband THz massive MIMO systems," 2023. [Online]. Available: https://arxiv.org/abs/2212.07484

[3] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "To-

ward 6G networks: Use cases and technologies," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, 2020.

[4] I. F. Akyildiz, J. M. Jornet, and C. Han, "Terahertz band: Next frontier for wireless communications," *Physical Communication*, vol. 12, pp. 16–32, 2014.

[5] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, 2019.

[6] Z. Chen, C. Han, Y. Wu, L. Li, C. Huang, Z. Zhang, G. Wang, and W. Tong, "Terahertz wireless communications for 2030 and beyond: A cutting-edge frontier," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 66–72, 2021.

[7] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4250–4285, 2022.

[8] 3rd Generation Partnership Project (3GPP), *3GPP TR 21.917 - Release 17 Description*, 09 2022, version 1.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3937

[9] S. Koenig, D. Lopez Diaz, J. Antes, F. Boes, R. Henneberger, A. Leuther, A. Tessmann, R. Schmogrow, D. Hillerkuss, R. Palmer, T. Zwick, C. Koos, W. Freude, O. Ambacher, J. Leuthold, and I. Kallfass, "Wireless sub-THz communication system with high data rate," *Nature Photonics*, vol. 7, pp. 977–981, 10 2013.

[10] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband THz massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7271–7286, 2022.

[11] L. Yan, C. Han, and J. Yuan, "Energy-efficient dynamic-subarray with fixed true-time-delay design for Terahertz wideband hybrid beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 2840–2854, 2022.

[12] F. Gao, B. Wang, C. Xing, J. An, and G. Y. Li, "Wideband beamforming for hybrid massive MIMO Terahertz communications," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1725–1740, 2021.

[13] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband Terahertz massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1604–1620, 2021.

[14] Y. Wu, G. Song, H. Liu, L. Xiao, and T. Jiang, "3-D hybrid beamforming for Terahertz broadband communication system with beam squint," *IEEE Transactions on Broadcasting*, vol. 69, no. 1, pp. 264–275, 2023.

[15] Q. Yuan, H. Liu, M. Xu, Y. Wu, L. Xiao, and T. Jiang, "Deep learning-based hybrid precoding for Terahertz massive MIMO communication with beam squint," *IEEE Communications Letters*, vol. 27, no. 1, pp. 175–179, 2023.

[16] A. M. Elbir, K. V. Mishra, and S. Chatzinotas, "Terahertz-band joint ultra-massive MIMO radar-communications: Model-based and model-free hybrid beamforming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1468–1483, 2021.

[17] A. M. Elbir, "A unified approach for beam-split mitigation in Terahertz wideband hybrid beamforming," *IEEE Transactions on Vehicular Technology*, pp. 1–6, 2023.

[18] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[19] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[20] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[21] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[22] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[23] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems,"

[24] M. Cai, K. Gao, D. Nie, B. Hochwald, J. N. Laneman, H. Huang, and K. Liu, "Effect of wideband beam squint on codebook design in phased-array wireless systems," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.

[25] X. Liu and D. Qiao, "Space-time block coding-based beamforming for beam squint compensation," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 241–244, 2019.

[26] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.

[27] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, S. Sun, and T. S. Rappaport, "Spatial-wideband effect in massive MIMO with application in mmwave systems," *IEEE Communications Magazine*, vol. 56, no. 12, pp. 134–141, 2018.

[28] B. Wang, M. Jian, F. Gao, G. Y. Li, and H. Lin, "Beam squint and channel estimation for wideband mmwave massive MIMO-OFDM systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 5893–5908, 2019.

[29] C. Han, L. Yan, and J. Yuan, "Hybrid beamforming for Terahertz wireless communications: Challenges, architectures, and open problems," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 198–204, 2021.

[30] J. Tan and L. Dai, "THz precoding for 6G: Challenges, solutions, and opportunities," *IEEE Wireless Communications*, pp. 1–8, 2022.

[31] N. M. Monroe, G. C. Dogiamis, R. Stingel, P. Myers, X. Chen, and R. Han, "Electronic THz pencil beam forming and 2D steering for high angular-resolution operation: A 98 × 98-unit 265GHz CMOS reflectarray with in-unit digital beam shaping and squint correction," in *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, vol. 65, 2022, pp. 1–3.

[32] Y. Wu, L. Xiao, G. Liu, and T. Jiang, "Hybrid beamforming for Terahertz wireless communications with beam squint: A survey," *Chinese Journal of Electronics*, vol. 31, no. E220309, p. 1043, 2022.

[33] C. Han, L. Yan, and J. Yuan, "Hybrid beamforming for Terahertz wireless communications: Challenges, architectures, and open problems," *IEEE Wireless Communications*, vol. 28, pp. 198–204, 2021.

[34] R. Mailloux, *Phased Array Antenna Handbook, Third Edition*, 2017.

[35] M. Longbrake, "True time-delay beamsteering for radar," in *2012 IEEE National Aerospace and Electronics Conference (NAECON)*, 2012, pp. 246–249.

[36] R. Rotman, M. Tur, and L. Yaron, "True time delay in phased arrays," *Proceedings of the IEEE*, vol. 104, no. 3, pp. 504–518, 2016.

[37] E. Ghaderi, A. Sivadhasan Ramani, A. A. Rahimi, D. Heo, S. Shekhar, and S. Gupta, "An integrated discrete-time delay-compensating technique for large-array beamformers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 9, pp. 3296–3306, 2019.

[38] M.-K. Cho, I. Song, and J. D. Cressler, "A true time delay-based sige bi-directional T/R chipset for large-scale wideband timed array antennas," in *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2018, pp. 272–275.

[39] F. Hu and K. Mouthaan, "A 1–20 ghz 400 ps true-time delay with small delay error in 0.13 μm cmos for broadband phased array antennas," in *2015 IEEE MTT-S International Microwave Symposium*, 2015, pp. 1–3.

[40] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, "Near-field rainbow: Wideband beam training for XL-MIMO," 2022. [Online]. Available: https://arxiv.org/abs/2205.03543

[41] J. Tan and L. Dai, "Wideband beam tracking in THz massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1693–1710, 2021.

[42] R. Li, H. Yan, and D. Cabric, "Rainbow-link: Beam-alignment-free and grant-free mmW multiple access using true-time-delay array," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1692–1705, 2022.

[43] V. V. Ratnam, J. Mo, A. Alammouri, B. L. Ng, J. Zhang, and A. F. Molisch, "Joint phase-time arrays: A paradigm for frequency-dependent analog beamforming in 6G," *IEEE Access*, vol. 10, pp. 73 364–73 377, 2022.

[44] I. K. Jain, R. Reddy Vennam, R. Subbaraman, and D. Bharadia, "mmflexible: Flexible directional frequency multiplexing for multi-user mmWave networks," in *IEEE INFOCOM 2023 - IEEE Conference on Computer Communications*, 2023, pp. 1–10.

[45] H. Luo and F. Gao, "Beam squint assisted user localization in near-field communications systems," 2022. [Online]. Available: https://arxiv.org/abs/2205.11392

[46] C.-C. Lin, C. Puglisi, V. Boljanovic, H. Yan, E. Ghaderi, J. J. Gaddis, Q. Xu, S. Poolakkal, D. Cabric, and S. Gupta, "Multi-mode spatial signal processor with rainbow-like fast beam training and wideband communications using true-time-delay arrays," *IEEE Journal of Solid-State Circuits*, vol. 57, pp. 3348–3360, 2022.

[47] C.-C. Lin, V. Boljanovic, H. Yan, E. Ghaderi, M. A. Mokri, J. J. Gaddis, A. Wadaskar, C. Puglisi, S. Mohapatra, Q. Xu, S. Poolakkal, D. Heo, S. Gupta, and D. Cabric, "Wideband beamforming with rainbow beam training using reconfigurable true-time-delay arrays for millimeter-wave wireless," *IEEE Circuits and Systems Magazine*, vol. 22, no. 4, pp. 6–25, 2022.

[48] O. Castañeda, S. Jacobsson, G. Durisi, T. Goldstein, and C. Studer, "Finite-alphabet MMSE equalization for all-digital massive MU-MIMO mmWave communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2128–2141, 2020.

[49] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-ray channel modeling and wideband characterization for wireless communications in the Terahertz band," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2402–2412, 2015.

[50] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the Terahertz band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, 2011.

[51] H. Sarieddeen, M.-S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for Terahertz communications," *Proceedings of the IEEE*, vol. 109, no. 10, pp. 1628–1665, 2021.

[52] R. Bracewell, *The Fourier Transform and its Applications*, 3rd ed. McGraw-Hill, New York, 2000.

[53] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, 2016.

[54] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, 2015.

[55] W. Zhang, T. Kim, and S.-H. Leung, "A sequential subspace method for millimeter wave MIMO channel estimation," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5355–5368, 2020.

[56] W. Zhang, M. Dong, and T. Kim, "MMV-based sequential AoA and AoD estimation for millimeter wave MIMO channels," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4063–4077, 2022.

[57] W. Zhang, T. Kim, D. J. Love, and E. Perrins, "Leveraging the restricted isometry property: Improved low-rank subspace decomposition for hybrid millimeter-wave systems," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5814–5827, 2018.

[58] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1432–1443, 2017.

Dang Qua Nguyen (Student Member, IEEE) received a B.Eng. degree in Electronics and Telecommunications from Hanoi University of Science and Technology (HUST), Hanoi, Vietnam in 2019. He is currently pursuing a Ph.D. degree in Electrical Engineering and Computer Science at The University of Kansas (KU), KS, USA, starting from September 2020. His research interests are statistical signal processing, wireless communication, and federated learning.

Taejoon Kim (Senior Member, IEEE) received a Ph.D. in electrical and computer engineering from Purdue University, West Lafayette, IN, USA. He is an Associate Professor and the Chair's Council Faculty of Electrical Engineering and Computer Science at the University of Kansas (KU). Before joining KU, he held positions at Nokia Bell Laboratories, Berkeley, KTH Royal Institute of Technology, and the City University of Hong Kong. He was an Associate Editor of the IEEE Transactions on Communications. He holds 29 issued U.S. patents. His research spans communications, machine learning, and security. He was a recipient of the Harry Talley Excellence in Teaching Award and the Miller Professional Development Award in Research from the KU School of Engineering. Along with the coauthors, he won the IEEE Communications Society Stephen O. Rice Prize in 2016 and the IEEE PIMRC 2012 Best Paper Award.