Vandermonde Constrained Tensor Decomposition for Hybrid Beamforming in Multi-Carrier MIMO Systems

Mohamed Salah Ibrahim[†], Akshay Malhotra[‡], Mihaela Beluri^{*}, Arnab Roy[†], and Shahab Hamidi-Rad[‡]

[†]InterDigital Communications Inc., Conshohocken, PA, USA, [‡]InterDigital Communications Inc., Los Altos, CA, USA

*InterDigital Communications Inc., New York, NY, USA

Abstract-Hybrid beamforming has evolved as a promising technology that offers the balance between system performance and design complexity in mmWave MIMO systems. Existing hybrid beamforming methods either impose unitmodulus constraints or a codebook constraint on the analog precoders/combiners, which in turn results in a performanceoverhead tradeoff. This paper puts forth a tensor framework to handle the wideband hybrid beamforming problem, with Vandermonde constraints on the analog precoders/combiners. The proposed method strikes the balance between performance, overhead and complexity. Numerical results on a 3GPP link-level test bench reveal the efficacy of the proposed approach relative to the codebook-based method while attaining the same feedback overhead. Moreover, the proposed method is shown to achieve comparable performance to the unit-modulus approaches, with substantial reductions in overhead.

I. INTRODUCTION

Millimeter wave (mmWave) has emerged as a powerful technology, that can handle the unprecedented demands on wireless connectivity, through offering large available bandwidth [1]. However, the high propagation loss inherent to mmWave bands, if not mitigated, can severely impact the system performance. Large antenna arrays which achieve high beamforming gains are used to compensate the propagation loss [2].

Large scale antenna systems implementation, on the other hand, incurs several practical challenges including the high energy consumption and cost of radio frequency (RF) chains, as each antenna element requires a dedicated RF chain. Such hurdles limit the possibility of employing a fully digital beamforming design. As an efficient surrogate, hybrid (analog/digital) beamforming has been introduced in [3], [4] as means of attaining favorable complexity-performance tradeoff in mmWave multicarrier massive MIMO systems. Hybrid beamforming relies on using a small number of RF chains to design high-dimensional analog precoders (implemented with only phase shifters) together with a low-dimensional (digital) baseband precoder. The combination of analog and digital precoders has the potential to approach the performance of a purely digital solution while providing substantial savings in energy consumption and design complexity.

Although maximizing the system spectral efficiency in the case of digital beamforming design admits a simple algebraic solution via singular value decomposition (SVD) [5], hybrid

beamforming yields a highly non-convex problem that requires joint optimization of the hybrid precoders and combiners [3]. A more tractable formulation is to transform the hybrid beamforming design to a matrix factorization problem. In particular, the optimal SVD-based digital solution is first derived to maximize the spectral efficiency. Then, the hybrid beamforming is posed towards factorizing the fully digital precoder (combiner) as the hybrid precoding (combining) components. The factorization is usually solved either under unit modulus constraints [3], [6] or with codebook constraints [7] on the analog precoder (combiner), to ensure that the analog precoder can be modeled using phase shifters. While considering the unit modulus constraints, in general, result in a much better solution compared to the codebook constraint [6], the resulting communication overhead of the latter is considerably lower [7], rendering it more appropriate for limited feedback systems [8]. Further, compared to the codebook constraints approach, the feedback overhead for unit magnitude constraints scales linearly with the number of Tx/Rx antennas, thereby precluding its use in massive MIMO systems.

This begs the question whether it is possible to achieve a comparable performance to the unit-modulus based methods while yielding the feedback associated with the codebookbased approaches? This is the central question that this paper seeks to address. We answer the stated question in the affirmative by modeling the wideband hybrid beamforming as a low rank tensor decomposition problem with Vandermonde constraints on the analog precoders/combiners. Invoking the socalled parallel factor (PARAFAC) analysis, to decompose the resulting tensor, we show that PARAFAC yields high-quality hybrid precoders/combiners, with identifiability guarantees on the resulting factors. This paper adds to the broad variety of tensors applications in wireless communications [9]–[11]. Different from all prior hybrid beamforming works that adopt the spectral efficiency formula for performance evaluation, this paper evaluates the practical impact of the proposed method by integrating hybrid beamforming to an end-to-end communication scenario with time-varying channels. Numerical results demonstrate that the end-to-end performance of the proposed approach considerably outperforms the codebook based method while achieving comparable performance to the unit modulus based approaches. Further, the proposed method yields significantly lower communication overhead compared

to unit modulus approaches.

II. SYSTEM MODEL

Consider a downlink transmission in a multi-carrier MIMO system comprising a base station (BS) and a single user equipment (UE). The BS is equipped with N_t transmit antennas and N_t^{RF} transmit radio frequency (RF) chains while the UE is equipped with N_r receive antennas and $N_r^{\rm RF}$ receive RF chains. The BS aims at communicating N_s data streams to the UE over K subcarriers, where $N_s \leq N_t^{\text{RF}} \leq N_t$ and $N_s \leq N_r^{\text{RF}} \leq N_r$ [3]. The BS first employs a digital baseband precoding matrix $\mathbf{F}_{\text{BB}}[k] \in \mathbb{C}^{N_t^{\text{RF}} \times N_s}$ on the transmitted symbols $\mathbf{s}[k] \in \mathbb{C}^{N_s}, \forall k \in [K] := \{0, \cdots, K-1\}, \text{ as}$ shown in Fig. 1. Then, the data symbols are transformed to the time domain using N-point inverse fast Fourier transform (IFFT). After a cyclic prefix (CP) is added to the time-domain signal, the BS applies an analog precoder $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N_t imes N_t^{\mathrm{RF}}}$ (implemented using analog phase shifters), i.e., $|\mathbf{F}_{RF}(i, j)| =$ 1, $\forall i = 0, \dots, N_t - 1$ and $j = 0, \dots, N_t^{\text{RF}} - 1$. Notice that same \mathbf{F}_{RF} is applied across all subcarriers, i.e., \mathbf{F}_{RF} is frequency independent. Towards this end, the transmitted complex signal from the BS can be expressed as,

$$\mathbf{x}[k] = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[k] \mathbf{s}[k], \ \forall k \in [K].$$
(1)

It is assumed that i) $\mathbb{E}[\mathbf{s}[k]\mathbf{s}^{H}[k]] = \frac{\alpha}{KN_{s}}\mathbf{I}_{N_{s}}$, and ii) the total power budget constraint α is satisfied by enforcing the constraint $\|\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_{F}^{2} = N_{s} \ \forall k \in [K].$

At the receiver, the UE first employs an analog combiner $\mathbf{W}_{\mathrm{RF}} \in \mathbb{C}^{N_r \times N_r^{\mathrm{RF}}}$ followed by a digital baseband combiner $\mathbf{W}_{\mathrm{BB}} \in \mathbb{C}^{N_r^{\mathrm{RF}} \times N_s}$ after CP removal and frequency transformation using N-point FFT. Similar to the unit modulus constraint on the entries of \mathbf{F}_{RF} , it assumed that the (i, j)-th entry of \mathbf{W}_{RF} has a unit modulus, i.e., $|\mathbf{W}_{\mathrm{RF}}(i, j)| = 1$, $\forall i = 0, \cdots, N_r^{\mathrm{RF}} - 1$ and $j = 0, \cdots, N_r - 1$. Thus, the N_s -dimensional complex baseband signal at the UE at the k-th subcarrier is given by,

$$\mathbf{y}[k] = \mathbf{W}_{BB}^{H}[k]\mathbf{W}_{RF}^{H}\mathbf{H}[k]\mathbf{x}[k] + \mathbf{W}_{BB}^{H}[k]\mathbf{W}_{RF}^{H}\mathbf{v}[k].$$
(2)

where $\mathbf{H}[k] \in \mathbb{C}^{N_r \times N_t}$ represents the downlink channel at the k-th subcarrier, and $\mathbf{v}[k] \in \mathbb{C}^{N_r}$ is the additive white Guassian noise vector associated with the k-th subcarrier, $\forall k \in [K]$. It is assumed that the entries of $\mathbf{v}[k]$ are independent and identically distributed (i.i.d) random variables with zero mean and variance σ^2 , i.e., $\mathbf{v}[k] \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_{N_r}), \ \forall i = 0, \cdots, N_r - 1$. Throughout this work, we assume that the channel matrices across the subcarriers $\{\mathbf{H}[k]\}_{k=1}^K$ are perfectly known at the UE.

Remark 1. It is worth pointing out that, in practical wireless systems, there is one representative channel matrix for each group of subcarriers or resource blocks, referred to as subband size, and hence, there is one baseband precoder/combiner for each subband as opposed to each subcarrier. The reason behind that is primarily to reduce the overhead associated with the channel and/or precoding-related feedback. In the hybrid beamforming context, this will reduce the overhead associated with the baseband precoders/combiners, and will



Fig. 1: Block diagram of OFDM-MIMO system with a BS and a UE employing hybrid precoding and combining.

also reduce the complexity as obviously smaller number of baseband precoders and combiners need to be computed. This fact will be utilized later in the simulations (Section VI) and also in the overhead computations associated with the proposed approach and the existing hybrid beamforming methods.

III. PROBLEM DEFINITION

The wideband hybrid beamforming problem seeks to find the set of hybrid precoders and combiners $(\mathbf{F}_{RF}, \{\mathbf{F}_{BB}[k]\}_{k=1}^{K}, \mathbf{W}_{RF}, \{\mathbf{W}_{BB}[k]\}_{k=1}^{K})$ that can maximize the spectral efficiency. Assuming the transmitted symbols follow a Gaussian distribution, the achievable spectral efficiency associated with the *k*-th subcarrier can be expressed as [5]

$$\mathbf{R}[k] = \log_2(\det(\mathbf{I}_{N_s} + \frac{\alpha}{N_s} \mathbf{\Gamma}^{-1}[k] \mathbf{W}_{BB}^H[k] \mathbf{W}_{RF}^H \mathbf{H}[k] \mathbf{F}_{RF} \mathbf{F}_{BB}[k] \times \mathbf{F}_{BB}^H[k] \mathbf{F}_{RF}^H \mathbf{H}^H[k] \mathbf{W}_{RF} \mathbf{W}_{BB}[k]))$$
(3)

where $\Gamma[k] := \sigma^2 \mathbf{W}_{BB}^H[k] \mathbf{W}_{RF}^H \mathbf{W}_{RF} \mathbf{W}_{BB}[k]$ represents the covariance matrix of the post-processing noise term in (2). The goal is then to design the hybrid precoders $(\mathbf{F}_{RF}, \{\mathbf{F}_{BB}[k]\}_{k=1}^K, \mathbf{W}_{RF}, \{\mathbf{W}_{BB}[k]\}_{k=1}^K)$ that aim at maximizing the overall spectral efficiency while satisfying the imposed constraints on the analog and digital precoders/combiners. Maximizing the spectral efficiency, though, yields a highly non tractable optimization problem that requires the hybrid precoders and combiners to be jointly optimized.

Instead of maximizing the spectral efficiency, one can decouple the precoders and combiners design, and formulate the hybrid beamforming problem as two separate low-rank matrix factorization problems [3], [6], [12]. The precoder problem aims at factorizing the optimal digital precoder $\mathbf{F}_{opt}[k] \in \mathbb{C}^{N_t \times N_s}$ to $\mathbf{F}_{RF}\mathbf{F}_{BB}[k]$, where the columns of $\mathbf{F}_{opt}[k]$ are the N_s dominant right singular vectors of $\mathbf{H}[k], \forall k \in [K]$. On the other hand, the combiner problem seeks to factorize $\mathbf{W}_{opt}[k] \in \mathbb{C}^{N_r \times N_s}$ to $\mathbf{W}_{RF}\mathbf{W}_{BB}[k]$, where $\mathbf{W}_{opt}[k]$ is the WMMSE solution, i.e., $\mathbf{W}_{opt}[k] = (\overline{\mathbf{H}}^H[k]\overline{\mathbf{H}}[k] + \sigma^2 \mathbf{I}_{N_s})^{-1}\overline{\mathbf{H}}^H$ and $\overline{\mathbf{H}}[k] = \mathbf{H}[k]\mathbf{F}_{opt}[k]$. Interestingly, it has been shown in [12] that solving the factorization problems implicitly leads to maximizing the system spectral efficiency. Since both problems exhibit similar mathematical formulation, except that the precoder problem has an additional sum power constraint, we will focus on the precoder factorization problem. However, the proposed method may be easily applied to solve the combiner problem. From an optimization perspective, given the fully digital SVD-based precoder $\mathbf{F}_{opt}[k]$, the hybrid beamforming problem can be posed as [3], [6], [12]

s.t.

$$\min_{\mathbf{F}_{\mathsf{RF}},\{\mathbf{F}_{\mathsf{BB}}[k]\}_{k=1}^{K}} \sum_{k=1}^{K} \|\mathbf{F}_{\mathsf{opt}}[k] - \mathbf{F}_{\mathsf{RF}}\mathbf{F}_{\mathsf{BB}}[k]\|_{F}^{2}$$
(4a)

$$\mathbf{F}_{\mathrm{RF}} \in \mathcal{F} \tag{4b}$$

$$\|\mathbf{F}_{\mathsf{RF}}\mathbf{F}_{\mathsf{BB}}[k]\|_F^2 = N_s, \ \forall k \in [K]$$
(4c)

where \mathcal{F} is the feasible set of the analog precoders. In the wideband hybrid beamforming literature, the feasible set either includes unit modulus constraints on the entries of \mathbf{F}_{RF} [6] (denoted as \mathcal{F}_U), or code-book based selection of the columns of \mathbf{F}_{RF} [7] (denoted as \mathcal{F}_C). The two feasible sets yield an interesting overhead-performance trade-off. While considering the feasible set \mathcal{F}_C results in much lower overhead relative to the set \mathcal{F}_U , the solution associated with \mathcal{F}_U performs much better than that of \mathcal{F}_C . The intuition is that \mathcal{F}_U provides a much wider search space compared to \mathcal{F}_C , i.e., $\mathcal{F}_C \subset \mathcal{F}_U$, and hence, better performance is expected.

In this paper, we will introduce a new feasible set (denoted as \mathcal{F}_V) to the wideband hybrid beamforming problem in (4) by enforcing a Vandermonde structure on the columns of \mathbf{F}_{RF} , i.e., $\mathcal{F}_V := \{\mathbf{x} \in \mathbb{C}^{N_t} \mid \mathbf{x} = [1, e^{j\phi}, \cdots, e^{j(N_t-1)\phi}]^T\}$, and $\phi \in [-\pi, \pi]$. Towards this end, the problem that this paper seeks to solve is the following low-rank matrix optimization problem,

Find
$$\mathbf{F}_{\mathrm{RF}}(\phi_0, \cdots, \phi_{N_{\mathrm{RF}}-1}), \{\mathbf{F}_{\mathrm{BB}}[k]\}_{k=1}^K$$
 (5a)

s.t.
$$\mathbf{F}_{opt}[k] \approx \mathbf{F}_{RF}\mathbf{F}_{BB}[k], \ \mathbf{F}_{RF} \in \mathcal{F}_V, \ \forall \ k \in [K].$$
 (5b)

Notice that the sum power constraint in (4c) is temporarily omitted as it has been shown that such a constraint can be satisfied via a simple normalization step to the resulting baseband precoders [3]. To our best knowledge, the formulation in (5) has not been considered before in the hybrid beamforming literature. Such a formulation strikes the balance between the obtained solution quality and the resulting overhead. In particular, the resulting solution achieves the same overhead associated with the \mathcal{F}_C set while achieving comparable performance to the solutions associated with the \mathcal{F}_U set. In the subsequent section, we will show that (5) can be reformulated as a tensor factorization problem where efficient tensor decomposition methods can be applied.

IV. PARAFAC DECOMPOSITION

Before reformulating (5) as a tensor factorization problem and to facilitate our discussion, we briefly review some key concepts that will be used in the proposed tensor approach. *A. Tensor Preliminaries*

A third order tensor $\mathcal{X} \in \mathbb{C}^{I \times J \times P}$ is a three way array whose elements are indexed by three indices (i, j, p). The so-called Parallel Factor decomposition (PARAFAC), a.k.a Canonical Polyadic Decomposition (CPD), is one powerful tensor decomposition method. A tensor \mathcal{X} admits a PARAFAC decomposition if it can be written as the sum of vector outer products [13],

$$\mathcal{X} = \sum_{f=1}^{F} \mathbf{a}_f \circ \mathbf{b}_f \circ \mathbf{c}_f.$$
(6)

where \circ denotes the vector outer product, and F is a positive integer that we refer to as the tensor rank or CPD rank (the smallest value such that (6) holds). The terms $\mathbf{a}_f \in \mathbb{C}^I$, $\mathbf{b}_f \in \mathbb{C}^J$ and $\mathbf{c}_f \in \mathbb{C}^P$ are the *f*-th columns of the so-called low-rank factors $\mathbf{A} \in \mathbb{C}^{I \times F}$, $\mathbf{B} \in \mathbb{C}^{J \times F}$, and $\mathbf{C} \in \mathbb{C}^{P \times F}$, respectively, of the tensor \mathcal{X} .

Different from the tensor format in (6), PARAFAC can also be written in slab format. Let $\mathbf{X}_p := \mathcal{X}(:, :, p)$ represent the *p*-th frontal slab of $\mathcal{X}, \forall p \in [P] := \{0, \dots, P-1\}$.¹ The PARAFAC decomposition of \mathcal{X} in the slab-format is given by

$$\mathcal{X}(:,:,p) = \mathbf{A}\mathbf{D}_p(\mathbf{C})\mathbf{B}^T, \quad \forall p \in [P].$$
(7)

where $\mathbf{D}_p(\mathbf{C}) := \text{Diag}(\mathbf{C}(p,:)) \in \mathbb{C}^{F \times F}$ with the elements on the diagonal be the *p*-th row of **C**. Throughout this paper, we will use the notation $\mathcal{X} := [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ to denote (7).

B. Identifiability

One distinctive property of tensors is that the PARAFAC model is essentially unique under mild conditions even if F is greater than $\max(I, J, P)$. The definition of essential uniqueness is presented as follows.

Definition 1. The PARAFAC decomposition of a tensor \mathcal{X} is said to be essentially unique, $\mathcal{X} := \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$, if \mathbf{A}, \mathbf{B} and \mathbf{C} are identifiable up to scaling and permutation. This means that if $\mathcal{X} := \llbracket \overline{\mathbf{A}}, \overline{\mathbf{B}}, \overline{\mathbf{C}} \rrbracket$, for some $\overline{\mathbf{A}} \in \mathbb{C}^{I \times F}$, $\overline{\mathbf{B}} \in \mathbb{C}^{J \times F}$, and $\overline{\mathbf{C}} \in \mathbb{C}^{P \times F}$, then there exists a permutation matrix $\mathbf{\Pi} \in \mathbb{R}^{F \times F}$ and diagonal scaling matrices $\{\Lambda_i\}_{i=1}^3$ such that,

$$\mathbf{A} = \overline{\mathbf{A}}\Pi\mathbf{\Lambda}_1, \mathbf{B} = \overline{\mathbf{B}}\Pi\mathbf{\Lambda}_2, \mathbf{C} = \overline{\mathbf{C}}\Pi\mathbf{\Lambda}_3, \mathbf{\Lambda}_1\mathbf{\Lambda}_2\mathbf{\Lambda}_3 = \mathbf{I}.$$
 (8)

If there is no structure imposed on the low rank factors, then a generic identifiability condition on PARAFAC uniqueness is given in [14]. If, however, one or more of the low rank factor matrices have a Vandermonde structure, then more relaxed uniqueness conditions based on the Kruskal rank can be found in [10], [13], [15]. The latest and the most relevant identifiability results to the problem considered herein is given as follows.

Theorem 1. [16] Consider the data model in (7) and assume that the factors $\mathbf{A} \in \mathbb{C}^{I \times F}$ and $\mathbf{C} \in \mathbb{C}^{P \times F}$ are Vandermonde and that $\mathbf{B} \in \mathbb{C}^{J \times F}$ is tall and full rank. If,

$$k_{\mathbf{A}} + \min(P - 1, F) \ge F + 1.$$
 (9)

then the PARAFAC decomposition of \mathcal{X} in terms of \mathbf{A} , \mathbf{B} , and \mathbf{C} is essentially unique, where $k_{\mathbf{A}}$ denotes the Kruskal rank (k-rank) of the matrix \mathbf{A} .

¹Note that we used the MATLAB notation $\mathcal{X}(:,:,p)$ to read the frontal slab of a three-way tensor.

It has been shown in [16] that a matrix with Vandermonde structure has full k-rank, i.e., $k_{\mathbf{A}} = \min(I, F)$. The condition in (9) will be interpreted later in the context of hybrid beamforming.

V. HYBRID BEAMFORMING VIA PARAFAC

In this section, it will be shown how the wideband hybrid beamforming problem in (5) can be reformulated as a tensor decomposition problem. Let us define the matrices $\mathbf{X} = [\mathbf{F}_{opt}[1], \cdots, \mathbf{F}_{opt}[K]] \in \mathbb{C}^{N_t \times KN_s}$ and $\mathbf{B} = [\mathbf{F}_{BB}^T[1], \cdots, \mathbf{F}_{BB}^T[K]]^T \in \mathbb{C}^{KN_s \times N_t^{RF}}$, then it can be easily seen that (5) can be expressed in more compact form as

Find
$$\mathbf{F}_{\mathrm{RF}}(\phi_0, \cdots, \phi_{N_{\star}^{\mathrm{RF}}-1}), \mathbf{B}$$
 (10a)

s.t.
$$\mathbf{X} \approx \mathbf{F}_{\text{RF}} \mathbf{B}^T, \ \mathbf{F}_{\text{RF}} \in \mathcal{F}_V.$$
 (10b)

Remark 2. Notice that while (10) assumes a uniform linear array (ULA) structure on the columns of the analog beamformer \mathbf{F}_{RF} , the proposed tensor method can be further extended to handle other array structures, for e.g., uniform planar array (UPA) [16]. In that sense, the proposed method can be used to recover azimuth and elevation estimates for each column of \mathbf{F}_{RF} . This is in fact a big advantage of the proposed approach relative to the state-of-the-art. Owing to space limitations, we will present only the ULA structure here.

Let us construct the following two subarrays,

$$\mathbf{A} = \mathbf{F}_{\mathsf{RF}}(1:end-1,:), \text{ (all rows except last)}$$
(11a)

$$\mathbf{A} = \mathbf{F}_{\mathsf{RF}}(2:end,:), \text{ (all rows except first)}$$
(11b)

Then, it follows that by exploiting the Vandermonde structure of the columns of the matrix \mathbf{F}_{RF} , the $(N_t - 1) \times N_t^{\text{RF}}$ matrices **A** and $\overline{\mathbf{A}}$ are displaced but otherwise identical subarrays, i.e.,

$$\overline{\mathbf{A}} = \mathbf{A} \boldsymbol{\Phi}_1. \tag{12}$$

where $\Phi_1 := \text{Diag}([e^{-\phi_0}, \cdots, e^{-\phi_{N_t^{\text{RF}}-1}}])$. Further, for consistency, let $\mathbf{A} = \mathbf{A} \Phi_0$, where $\Phi_0 = \mathbf{I}_{N_t^{\text{RF}}}$. Let $\mathbf{C} \in \mathbb{C}^{P \times N_t^{\text{RF}}}$ be a matrix holding the diagonal of Φ_p on its *p*-th row, for $p = 0, \cdots, P - 1$ and P = 2. Then, upon defining $\mathbf{X}_0 = \mathbf{X}(1 : end - 1, :) \in \mathbb{C}^{(N_t - 1) \times KN_s}$, $\mathbf{X}_1 = \mathbf{X}(2 : end, :) \in \mathbb{C}^{(N_t - 1) \times KN_s}$ and $\mathbf{D}_p(\Phi) = \Phi_p$, for p = 0, 1, we can write the following,

$$\mathbf{X}_0 = \mathbf{A} \mathbf{D}_0(\mathbf{C}) \mathbf{B}^T, \tag{13}$$

$$\mathbf{X}_1 = \mathbf{A} \mathbf{D}_1(\mathbf{C}) \mathbf{B}^T. \tag{14}$$

From the PARAFAC decomposition slab format defined in (7), it is easy to see that (13) and (14) form a two-slab, i.e., P = 2, PARAFAC model with Vandermonde structure in one mode. Thus, solving (10) is tantamount to decomposing the tensor $\mathcal{X} \in \mathbb{C}^{(N_t-1)\times KN_s \times P}$ with its *p*-th slab defined as $\mathcal{X}(:,:,p) := \mathbf{X}_p$, for $p = 0, \dots, P - 1$ and P = 2. From an optimization perspective, this can be expressed as

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} \|\mathcal{X} - [\![\mathbf{A},\mathbf{B},\mathbf{C}]\!]\|_F^2$$
(15)

Several algorithms have been developed to tackle the optimization problem (15) [17]. In this work, we adopt the trilinear alternating least square (TALS) algorithm implemented in the widely known Tensorlab MATLAB toolbox [18].

Considering the condition in (9) in the context of hybrid beamforming, one can easily see that with P = 2 and given that \mathbf{A} is tall (i.e., $N_t \ge N_t^{\mathrm{RF}} + 1$) and Vandermonde, the condition in (9) is always satisfied. The only requirement though to ensure essential uniqueness of \mathcal{X} is that $\mathbf{B} \in \mathbb{C}^{KN_s \times N_t^{\mathrm{RF}}}$ needs to be tall and full rank. This requires the number of subcarriers multiplied by the number of streams be greater than or equal to the number of transmit RF chains. This renders our proposed method not applicable for single carrier systems, i.e., K = 1, with $N_t^{\mathrm{RF}} > N_s$, otherwise, such a condition can be easily satisfied with a modest number of subcarriers. Let $\overline{\mathbf{A}} \in \mathbb{C}^{(N_t-1) \times N_t^{\mathrm{RF}}}$, $\overline{\mathbf{B}} \in \mathbb{C}^{KN_s \times N_t^{\mathrm{RF}}}$ and $\overline{\mathbf{C}} \in \mathbb{C}^{2 \times N_t^{\mathrm{RF}}}$

Let $\overline{\mathbf{A}} \in \mathbb{C}^{(N_t-1)\times N_t^{\mathrm{RF}}}$, $\overline{\mathbf{B}} \in \mathbb{C}^{KN_s \times N_t^{\mathrm{RF}}}$ and $\overline{\mathbf{C}} \in \mathbb{C}^{2\times N_t^{\mathrm{RF}}}$ be the resulting solution of (15). The goal now is to find $(\phi_0, \dots, \phi_{N_t^{\mathrm{RF}}-1})$ and $\{\mathbf{F}_{\mathrm{BB}}[k]\}_{k=1}^K$ given $(\overline{\mathbf{A}}, \overline{\mathbf{B}}, \overline{\mathbf{C}})$. To do so, we first recover the N_t^{RF} phases $\{\phi_i\}_{i=0}^{N_t^{\mathrm{RF}-1}}$ from the columns of $\overline{\mathbf{A}}$ by simply reading the angles of first elements of the columns of $\overline{\mathbf{A}}$.

Algorithm 1 V-TPAR: Vandermonde Two-slab PARAFAC
Input : $\{\mathbf{F}_{opt}[k] \in \mathbb{C}^{N_t \times N_s}\}_{k=1}^K, N_t^{RF},$
Construct $\mathbf{X} = [\mathbf{F}_{opt}[1], \cdots, \mathbf{F}_{opt}[K]] \in \mathbb{C}^{N_t \times KN_s}$
Construct $\mathcal{X} \in \mathbb{C}^{(N_t-1) \times KN_s \times 2}$ as $\mathcal{X}(:,:,1) = \mathbf{X}(1:end - 1)$
1,:), and $\mathcal{X}(:,:,2) = \mathbf{X}(2:end,:)$
Decompose $\mathcal{X} := [\![\overline{\mathbf{A}}, \overline{\mathbf{B}}, \overline{\mathbf{C}}]\!]$ using TALS
for $i = 0: N_t^{RF} - \ddot{1}$ do
Recover ϕ_i by computing the angle of the first element of
$\overline{\mathbf{A}}(:,i)$
Form $\mathbf{F}_{\text{RF}}(:,i) = [1, e^{j\phi_i}, \cdots, e^{j(N_t-1)\phi_i}]^T$
Obtain $\lambda_1^i = \overline{\mathbf{A}}(1,i), \ \lambda_3^i = \overline{\mathbf{C}}(1,i), \ \lambda_2^i = \frac{1}{\lambda_2^i \lambda_3^i}$
Obtain $\mathbf{B}(:,i) = \overline{\mathbf{B}}(:,i)/\lambda_2^i$
end
Reshape B to retrieve $\{\mathbf{F}_{BB}[k] \in \mathbb{C}^{N_t^{RF} \times N_s}\}_{k=1}^K$.

To obtain $\{\mathbf{F}_{BB}[k]\}_{k=1}^{K}$, we need to resolve the complex scaling ambiguity that is inherent to PARAFAC (see Definition 1 for the essential uniqueness of PARAFAC). Note that we ignore the permutation ambiguity, as in the hybrid beamforming context, finding the analog and baseband precoders up to a common permutation ambiguity is irrelevant since it merely amounts to shuffling the RF chains. The complex scale ambiguity though is important as it amounts to entirely changing the directions of the precoders. Fortunately, since the columns of both matrices A and C exhibit a Vandermonde structure, the column-wise scale ambiguity in both matrices can be resolved by simply dividing the elements of each column by the first element. Once the complex scale ambiguities associated with the columns of $\overline{\mathbf{A}}$ and $\overline{\mathbf{C}}$, denoted as Λ_1 and Λ_3 , respectively, are resolved, it can be seen from (8) that the column-wise scale ambiguity of **B**, denoted as Λ_2 , can be easily obtained as $\Lambda_2 = (\Lambda_3 \Lambda_1)^{-1}$. The above procedures for solving the wideband hybrid beamforming problem using Vandermonde-



Fig. 2: The BER performance with $N^{\text{RF}} = 2$ and $N_s = 1$.

Parameter	Value	
Carrier frequency	28 GHz	
Subcarrier spacing	60 kHz	
Modulation	16-QAM	
Code rate	0.49	
Number of transmit antennas	32	
Number of receive antennas	8	
UE speed	0.5 km/hr	
Delay spread	300 ns	
Channel model	CDL-C	

TABLE I: Parameter settings for the simulations.

constrained Two-slab PARAFAC (V-TPAR) are outlined in Algorithm 1.

The complexity of Algorithm 1 is incurred in decomposing the tensor \mathcal{X} using the iterative TALS algorithm. The per iteration complexity of TALS is equal to the cost of inverting an $(N_t^{\text{RF}})^2 \times (N_t^{\text{RF}})^2$ matrix. The overall complexity then depends on the total number of iterations which in turn depends on the problem and the size of the tensor (see [17] and references therein for convergence properties of TALS). As we will see later, for the considered problem, a few iterations of TALS seem to be sufficient to obtain high-quality solution.

VI. SIMULATIONS

In this section, we will provide numerical results on 3GPP link-level channel model to assess the performance of the proposed method. The adopted simulation parameters are listed in Table I. We use the CDL-C channel model with the delay spread set to 300 ns. Both BS and UE are equipped with uniform linear array where the antenna elements are separated by a half wavelength. All results are averaged out over 200 realizations. The number of subbands is set to 30, i.e., K = 30, where each subband consists of one resource block (RB), i.e., 12 subcarriers. The channel matrix for each subband is obtained by averaging out the channels across the 12 subcarriers. For the proposed method implementation, we used the TALS algorithm implemented in the Tensorlab MATLAB toolbox. Finally, all simulations were performed on an Intel(R) Xeon(R) Gold 6234 CPU.

To benchmark the performance of the proposed method, we use the manifold optimization (MO) alternating minimization



Fig. 3: The BER performance $N^{\text{RF}} = 3$ and $N_s = 2$.

\mathbf{F}_{RF} feasible set	\mathcal{F}_{u}	\mathcal{F}_{c}	\mathcal{F}_v
Num. of parameters	$N_t N_t^{\rm RF}$	$N_t^{\rm RF}$	$N_t^{\rm RF}$
Method	MO and PE	OMP	T-VPAR

TABLE II: Feedback overhead associated with the different feasible sets of the analog beamformers.

algorithm [6], the phase extraction (PE) alternating minimization algorithm [6] and the OMP algorithm [3] as baselines. Both MO and PE solve the wideband hybrid beamforming problem (4) with unit modulus constraints on the entries of the analog beamformers, while the OMP algorithm solve (4) with codebook constraint on the columns of the analog beamformers. For OMP, we use the DFT codebook for both \mathbf{F}_{RF} and \mathbf{W}_{RF} .

From the feedback overhead perspective, one can see from Table II that the Vandermonde feasible set (our proposed method) attains the same overhead of the codebook one (OMP). In particular, the number of parameters to feed back is independent of the number of transmit (receive) antennas and is equal to the number of transmit (receive) antennas if the the analog precoders and combiners are computed at the UE (BS). On the other side, the unit-modulus feasible set (MO and PE) suffers from the large overhead that scales up with the number of transmit/receive antennas, thereby limiting their use in limited feedback systems.

To evaluate the practical impact of the different hybrid beamforming algorithms, we report the coded BER in an endto-end system. First, we consider a scenario with $N^{\text{RF}} = 2$ and $N_s = 1$, i.e., $N^{\text{RF}} = 2N_s$, while the rest of the parameters are as listed in Table I. It is known from [6] that when $N^{\text{RF}} = N_s$, PE achieves the same performance of MO at much lower complexity while the performance of the former degrades when $N^{\text{RF}} > N_s$. Fig. 2 shows the end-to-end coded BER performance of the different methods. One can see that, for this case, the proposed method achieves more than 1 dB SNR gain relative to OMP. More interestingly, the proposed approach outperforms the PE method with more than an order of magnitude reduction in BER at -8 dB. Further, when $N_s < N^{\text{RF}} < 2N_s$ as shown in Fig. 3, the performance of the proposed method significantly outperforms OMP with roughly



Fig. 4: The BER performance $N^{\text{RF}} = N_s = 2$.

4 dB SNR gain. Finally, one can see that both the tensor method and PE attain approximately the same performance, with 1 dB loss relative to MO.

Next, we simulated another scenario with $N^{\text{RF}} = N_s = 2$. It can be seen that now PE achieves the same performance as MO while the proposed method incurs roughly 2 dB SNR loss, as Fig. 4 depicts. In addition, one can see that the proposed algorithm considerably outperforms the OMP algorithm with more than an order of magnitude reduction in BER when the SNR exceeds -10 dB.

Finally, to assess the complexity of the proposed tensor approach, Fig. 5 depicts the average run time of the proposed method relative to the considered baselines, when $N^{\text{RF}} = N_s = 2$ and $N^{\text{RF}} = 3$ and $N_s = 2$. We observe that the run time of the proposed method is comparable to PE while achieving more than an order of magnitude reduction in run time compared to MO in both setups. Finally, OMP features the lowest run time but this obviously comes at the expense of performance.

VII. CONCLUSIONS

This paper has considered single user hybrid precoding and combining in wideband mmWave MIMO systems under Vandermonde constraints on the hybrid precoders and combiners. The problem is formulated as a tensor factorization problem where PARAFAC is invoked to find the Vandermondeconstrained analog beamformers and the set of baseband precoders - with identifiability guarantees. Numerical results on a 3GPP link-level test bench have revealed the superiority of the proposed method relative to the state-of-the-art. In particular, the proposed method has shown to be striking the balance between performance, overhead and complexity. As a future work, we aim at expanding the applicability of the proposed framework to other array structures such as uniform plannar array (UPA). Further, we plan to explore the impact of increasing the number of subarrays (multi-slab PARAFAC as opposed to two slab) on the estimation accuracy, and its trade-off with computational complexity.



REFERENCES

- F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. of the IEEE*, vol. 102, no. 3, pp. 366–385, Feb. 2014.
- [3] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. on Wir. Commun.*, vol. 13, no. 3, pp. 1499–1513, Nov. 2014.
- [4] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. of Sel. Topics in Sig. Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [5] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. on Sel. Areas in Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.
- [6] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. of Sel. Topics in Sig. Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.
- [7] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. on Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
- [8] D. J. Love, R. W. Heath, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. on Sel. Areas in Commun.*, vol. 26, no. 8, pp. 1341– 1365, Oct. 2008.
- [9] A. L. de Almeida, G. Favier, and J. C. M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Processing*, vol. 87, no. 2, pp. 337–351, Feb. 2007.
- [10] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind parafac receivers for DS-CDMA systems," *IEEE Trans. on Sig. Process.*, vol. 48, no. 3, pp. 810–823, June 2000.
- [11] M. S. Ibrahim, C. I. Kanatsoulis, and N. D. Sidiropoulos, "Downlink channel feedback for FDD massive MIMO systems via tensor compression and sampling," in 54th Asilomar Conference on Sig., Sys., and Comp., Pacific Grove, CA, USA, Nov. 2020, pp. 27–31.
- [12] O. El Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *IEEE Int. Conf. on Commun. (ICC)*, Ottawa, Canada, June 2012, pp. 3724–3729.
- [13] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *Journal of Chemometrics*, vol. 14, no. 3, pp. 229–239, Nov. 2000.
- [14] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 1018–1037, Sept. 2012.
- [15] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 3, pp. 642– 666, Sept. 2006.
- [16] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. on Sig. Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

- [17] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalex-akis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. on Sig. Process.*, vol. 65, no. 13, pp. 3551–3582, Apr. 2017.
 [18] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab v3.0," Mar. 2016.



