

Cooperative Hybrid Transmit Beamforming in Cell-free mmWave MIMO Networks

Meesam Jafri, *Student Member, IEEE*, Suraj Srivastava, *Student Member, IEEE*, Naveen K. D. Venkategowda, *Student Member, IEEE*, Aditya K. Jagannatham, *Member, IEEE*, Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Hybrid precoders and combiners are designed for cooperative cell-free multi-user millimeter wave (mmWave) multiple-input multiple-output (MIMO) cellular networks for low complexity interference mitigation. Initially, we derive an optimal hybrid transmit beamformer (HTBF) for a broadcast scenario considering both total and per access point (AP) power constraints. Next, an optimal successive hybrid beamformer technique is proposed for unicast and multicast scenarios which relies on the optimal minimum variance distortionless response (MVDR). We demonstrate that it mitigates both the interuser and intergroup interference, while successively ensuring orthogonality to the previously scheduled users/user groups. Furthermore, it is shown theoretically that the proposed schemes are capable of supporting a large number of users. Subsequently, a Bayesian learning (BL) based method is conceived for jointly designing the RF and baseband precoders/combiners for the various scenarios considered. Furthermore, we also conceive the uplink counterpart of our HTBF scheme, which is based on maximizing the signal-to-interference-plus noise ratio (SINR) of each individual user. Finally, the efficacy of the proposed schemes is characterized by our extensive simulation results in terms of cancelling the interuser/intergroup interference, which improves the spectral efficiency.

Index Terms—mmWave, hybrid beamforming, multi-cell multi-user, broadcast, and multicast, Bayesian learning.

I. INTRODUCTION

The next-generation wireless communication systems aim for increased data rates, reduced lower latency and large-scale access. The mmWave technology has shown significant promise in meeting these requirements [1]–[3]. Communication at mmWave frequencies has recently attracted significant attention due to the fact that the sub-6 GHz band faces acute spectrum shortage [4], [5]. In this context, the

under-utilized mmWave band has the potential of satisfying the requirements of next generation wireless networks. Furthermore, the abundance of free spectrum in the mmWave band is capable of supporting the massive connectivity of devices required by the Internet of Things (IoT) [6]. However, as shown via practical measurements in [7], [8], mmWave networks suffer from signal blockage and penetration losses, which degrades the received signal strength. The dense deployment of access points (APs) for creating small cells has shown significant promise for overcoming these challenges. However, inter-cell interference (ICI) becomes a significant bottleneck with the ever-increasing network densification. Several cooperative signal processing concepts, such as cooperative MIMO networks, network MIMO, and coordinated multipoint (CoMP) transmission [9]–[11], are presented in the literature to address the issue of ICI. Note that the network architectures mentioned above are conventionally implemented in a network-centric fashion by dividing the APs into disjoint clusters. In [12]–[14], a user-centric cell-free MIMO structure relying on a large number of distributed APs is considered, which are connected to a central processing unit (CPU). The network jointly serves multiple users in conjunction with no cell boundaries, similar to the cooperative MIMO structure. However, cooperative MIMO networks mitigate the ICI by grouping multiple cells into fixed cooperative clusters, which shifts the interference management from the cell level to the cluster level. The resultant inter-cluster interference imposes a fundamental limit on cooperation. Hence, users close to the cluster edge might not benefit from the cooperative MIMO clustering. By contrast, cell-free MIMO facilitates improved service for cell-edge users due to the elimination of the inter-cluster interference. Furthermore, the other significant aspect of cell-free MIMO is the deployment of a higher number of APs than that of the users in comparison to the cooperative MIMO. Thus, cell-free networks promise a further improvement in spectral efficiency by mitigating the interference through intelligent, cooperative beamforming.

Apart from the benefit of wider bandwidth availability, the short wavelength of mmWave frequencies allows us to pack large antenna array into a compact space for supporting highly directional beamforming, while simultaneously serving numerous users. Thus, the multiple-input-multiple-output (MIMO) technology can be exploited to obtain high beamforming gains for mitigating the effects of path loss, atmospheric absorption and penetration loss in the mmWave regime [15]. However, the high power consumption and

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The work of Aditya K. Jagannatham was supported in part by the Qualcomm Innovation Fellowship, and in part by the Arun Kumar Chair Professorship.

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

M. Jafri, S. Srivastava and A. K. Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology, Kanpur, Kanpur, 208016, India (e-mail: meesam@iitk.ac.in, ssvrast@iitk.ac.in, adityaj@iitk.ac.in.)

N. K. D. Venkategowda is with the Department of Science and Technology, Linköping University, 60174 Norrköping, Sweden (e-mail: naveen.venkategowda@liu.se)

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk)

associated cost/complexity of assigning a dedicated per-antenna RF chain prohibits the employment of the conventional fully-digital (FD) transceiver architecture in mmWave MIMO systems [7]. This has motivated the conception of the novel hybrid MIMO transceiver architecture as a promising solution, which employs much fewer RF chains in comparison to the antenna elements, thus easing its practical implementation. The overall MIMO signal processing in such a hybrid architecture is divided into the RF and baseband domains, where RF transmit precoding (TPC) yields a directional beamforming gain, whereas spatial multiplexing and interference cancellation are achieved via suitable baseband TPC. It can be readily observed that the design of the baseband and RF precoders and combiners is the key to address the various challenges associated with the proposed cooperative cell-free multi-user (CFMU) mmWave MIMO system, and also to harness the significant gains promised by the mmWave band for 5G/B5G.

A. Related Works

In the recent on mmWave literature, sophisticated techniques have been proposed for designing hybrid TPCs and receive combiners (RCs) for mmWave MIMO systems [16]–[19]. In particular, the early contributions focused more on single-user (SU) scenarios. Omar *et al.*, in [16] considered SU mmWave hybrid MIMO systems and derived the hybrid transmit beamformers (HTBF) in which the sparse scattering nature of mmWave wireless channel has been exploited to obtain a two-stage solution via the popular orthogonal matching pursuit (OMP) algorithm. Wang *et al.*, [17], proposed another hybrid TPC design that decomposes the original HTBF design problem into several sub-problems to render the optimization tractable. In order to further improve the spectral efficiency and to mitigate the inter-user interference (IUI) in practical mmWave cellular systems a multi-user (MU) mmWave MIMO system has also been considered in the pioneering works [18], [19], wherein a BS simultaneously serves multiple users/devices.

In a mmWave MU MIMO system, the digital component of the hybrid architecture provides higher degrees of freedom than analog-only beamforming, and can hence be employed for mitigating the interference among the different users. Note that the design of HTBF is a key challenge in MU mmWave MIMO systems, since the resultant IUI degrades the overall system performance [20], [21]. Therefore, there is an urgent need for minimizing the IUI, while the HTBF gain has to be simultaneously maximized for each user. Thus, it is of pivotal importance to design efficient HTBF for mmWave MU MIMO systems that have a low complexity. Hence, sophisticated HTBF techniques have been developed for optimizing their performance. In [19], a low-complexity iterative algorithm was proposed, which exploited the sparse scattering nature of the mmWave wireless channel for designing the HTBF of the BS and analog RC of each user. The proposed solution was shown to be asymptotically optimal for the large-scale antenna regime, which also demonstrates

the well-known fact that MU MIMO systems attain a high capacity than SU MIMO systems.

In [22]–[26], the coordination of different base stations (BSs) has been proposed for conventional sub-6 GHz systems for improving their performance. However, the proposed schemes targeted exclusively fully-digital sub-6 GHz MU MIMO systems. Hence, they are unsuitable for mmWave MU MIMO cellular systems. Michaloliakos *et al.* [27], proposed a novel scheme relying on jointly selecting the analog TPCs for maximizing the data rate of the users by harnessing a set of predefined beam patterns. Zhu *et al.* [28], proposed a novel cooperative HTBF design for multi-cell multi-user (MCMU) scenarios based on the Kronecker decomposition. First, analog transmit beamformer (TBF) is obtained by decomposing the beamformer as well as the data and interference path-vectors into the corresponding Kronecker products of unit-modulus phase-shift vectors. Next, the baseband TBF is designed by using the minimum mean square error (MMSE) TBF technique. Shu *et al.* [29], presented both signal-to-leakage-plus-noise-ratio (SLNR) and regularized zero-forcing based HTBF techniques for multi-stream MCMU systems to study the benefits of interference coordination in terms of the spectral efficiency. Bai *et al.* [30], presented a two-step hybrid cooperative TPC design for a distributed antenna scenario combined with user selection for energy-efficient transmissions. While their work has immense utility, it is restricted to a single-cell scenario, which does not exploit the additional advantages offered by multi-cell cooperation. Fang *et al.* [31], maximize the sum-power consumption of the APs supporting single-antenna users relying on cooperative HTBF.

While several authors have addressed the problem of hybrid beamformer design for classic cellular systems, the research of their cell-free counterparts is limited. In [32], the authors proposed a hybrid beamformer scheme, where the analog and digital beamformers are alternately optimized based on the weighted minimum mean squared error (WMMSE) criterion. In [33], the authors proposed a two-step procedure for their hybrid beamformer design. In the first step, the analog TPC is obtained from a codebook, followed by semidefinite relaxation (SDR) and convex approximation-aided baseband TPC design. In [34], the authors designed a hybrid beamforming scheme for CFMU massive MIMO systems toward maximizing the user rate. Chenghao *et al.* [35] proposed a block coordinated descent (BCD) algorithm based hybrid transmit beamformer (HTBF) maximizing the weighted sum-rate of cell-free mmWave MIMO systems. However, the aforementioned treatises assume the availability of either perfect knowledge of CSI or that of its second-order statistics. The HTBF design of a cell-free network requires the full knowledge of AoAs/AoDs of the multipath components and their complex-valued path gains. However, in practice, acquiring accurate estimates of these quantities is challenging. In the recent studies of cell-free MIMO systems, sophisticated techniques have been proposed for channel estimation in CFMU mmWave MIMO

systems [36]–[40]. Furthermore, the channel matrix in the mmWave regime is sparse, which can be jointly attributed to the reduced scattering and diffraction effects as well as to the focused beam emanating from the large antenna arrays. Exploiting this sparsity can lead to a significant reduction in the number of pilot transmissions required and also a substantial improvement in the estimation accuracy. In this context, the compressed sensing (CS) techniques proposed in [41], [42] utilize a sparse signal recovery model for channel estimation. The authors of [43], [44] have demonstrated the applicability of some classical CS based schemes for sparse channel estimation at a reduced pilot overhead. One can readily employ these existing schemes for efficient sparse channel estimation in CFMU mmWave MIMO networks.

In wireless communication systems, there are several scenarios, where common data has to be sent to a group of users, for instance, during the transmission of content pertaining to a live event. In such scenarios, multicast is an efficient framework for improving both the spectral and power efficiency by simultaneously transmitting the common data to a group of users [45]–[47]. This has huge potential in numerous applications, such as mobile TV, live video streaming, video conferencing etc. As a further advance in [45], a low-complexity uplink-downlink duality based suboptimal algorithm is presented to jointly design the beamformers for multiple users in a single-cell wireless multicast network. Chengjun *et al.* [46] conceived a low-complexity multi-quality multicast beamforming schemes. The authors of [47] proposed a weighted MMSE based multicast beamformer employing the popular successive convex approximation and Lagrangian duality based techniques. Note that all the above contributions only consider conventional sub-6 GHz fully digital transceiver architectures, which cannot be used in mmWave CFMU MIMO systems. Multicast transmission, specifically designed for mmWave MIMO networks, was studied in [48]–[50]. In [48], a single group scenario is considered for HTBF-aided multicast transmission, wherein an iterative algorithm was designed based on the principle of alternating minimization. In the seminal work [49], the author presented a single-cell single-group HTBF design relying on max-min fairness for multicasting using a limited number of RF chains. In [50], a suboptimal low-complexity HTBF design was proposed for the downlink of unicast and multicast mmWave MIMO systems. However, all these papers only considered a single AP scenario for multicast transmission. To the best of our knowledge, none of the works in the existing literature have sophisticatedly addressed the problem of cooperative HTBF design in CFMU mmWave MIMO systems. Thus, motivated by these limitations of the existing literature reviewed above, this paper presents a novel cooperative beamformer design conceived for efficiently mitigating the interference in CFMU mmWave MIMO networks. Table I boldly and explicitly contrasts to the literature discussed above. Next, we further detail the various contributions of our work.

B. Contributions

Again, we design HTBFs and RCs for broadcast, unicast and multicast scenarios in cooperative CFMU mmWave MIMO cellular systems.

- Initially, we design an optimal hybrid broadcast beamformer for CFMU mmWave MIMO systems. Furthermore, the optimal TBFs having per-AP power constraints are also derived for this scenario. At the users, a low-complexity maximal ratio combiner (MRC) is used in the baseband for maximizing the effective SNR.
- Next, the design and analysis is extended to both unicast and multicast transmission scenarios. Toward this, the minimum variance distortionless response (MVDR) [51], [52] based optimal successive minimum variance hybrid beamforming (OSHB) schemes are designed for mitigating both the inter-user and inter-group interference. It has been demonstrated theoretically that the proposed schemes can support a higher number of users than popular block diagonalization (BD) method.
- Subsequently, a Bayesian learning (BL)-based method is developed for jointly designing the RF and baseband TPCs/RCs, which is of significant utility, when perfect knowledge of the angles of arrival/departure (AoAs/AoDs) and path gains of the multipath components is not available.
- Furthermore, in order to address the cooperative beamformer design problem of the uplink, a successive CFMU uplink hybrid beamforming (SCUHBF) scheme is developed that maximizes the SINR. It is shown that the proposed scheme achieves an improved data rate as a benefit of interference cancellation.
- Simulation results are provided for characterizing the OSHB and SCUHBF schemes in cancelling the interuser/intergroup interference together with the improved spectral efficiency compared to a no-coordination scenario, the performance attained is close to that of the ideal fully-digital design.

C. Organization of the paper

The remainder of the paper is arranged as follows. Section II introduces the system model and channel model. In Section III, we formulate the problem of HBBF for CFMU mmWave MIMO systems. Section IV and V describe the problem of cooperative TPC design of unicast and multicast scenarios, respectively, and develop the OSHB schemes for it. The SCUHBF scheme of an uplink scenario is derived in Section VI, followed by the BL-based joint precoder design in Section-VII. Finally, Section VIII presents our simulation results for illustrating the merits of the proposed work, followed by our conclusions in Section IX.

D. Notation

Vectors and matrices are denoted by boldfaced lower and upper case letters respectively, while scalars are represented by lower case letters. The transpose, Hermitian, conjugate,

Table I: Contrasting our contribution to the existing literature

	[16]	[17]	[19]	[22]	[28]	[29]	[34]	[35]	[45]	[48]	[49]	[50]	Proposed Work
mmWave communication	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓	✓
Hybrid architecture	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓	✓
Multi-user			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Downlink	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
Uplink					✓				✓				✓
Cooperative beamforming				✓	✓	✓	✓	✓	✓				✓
Multicast single-group									✓	✓	✓	✓	✓
Multicast multi-group									✓				✓
Broadcast										✓		✓	✓
Unicast				✓	✓	✓	✓	✓	✓	✓		✓	✓

and inverse of a matrix \mathbf{F} are denoted by \mathbf{F}^T , \mathbf{F}^H , \mathbf{F}^* , and \mathbf{F}^{-1} , respectively. Furthermore, $\text{Tr}(\mathbf{A})$ denotes the trace of a matrix \mathbf{A} , $\|\mathbf{x}\|_2$ represents the l_2 -norm of a vector \mathbf{x} , and $\mathbb{E}[\cdot]$ is the expectation operator. Furthermore, $\mathbf{x} \sim \mathcal{CN}(\mathbf{b}, \mathbf{A})$ denotes a complex Gaussian random vector having mean \mathbf{b} and covariance matrix \mathbf{A} . Finally, \mathbf{I}_N represents an $N \times N$ identity matrix and $\mathbf{A} \succcurlyeq 0$ indicates that \mathbf{A} is a positive semi-definite (PSD) matrix.

II. MMWAVE CFMU SYSTEM AND CHANNEL MODEL

Consider a downlink cell-free mmWave hybrid MIMO system with M cooperative APs, where the m th AP has N_T transmit antennas (TAs) and $N_{\text{RF},m}$ RF chains so that $1 \leq N_{\text{RF},m} \ll N_T$, $\forall m, 1 \leq m \leq M$. These M APs are connected to a central processing unit (CPU) and serve G user-groups, where the g th group is comprised of U_g users, as shown in Fig. 1. Each user is equipped with N_R receive antennas (RAs) and $N_{\text{RF},u}$ RF chains, which satisfy $1 \leq u \leq U_g$, $\forall 1 \leq g \leq G$. We consider a next-generation multicast scenario, where several users request identical services relying on the same data. Such a group of users is referred to as a multicast group in the literature, and many such groups may coexist as shown in Fig. 1. For convenience, user u in the g th multicast group is denoted as $U_u^{(g)}$. The signal, $\mathbf{r}_u^{(g)} \in \mathbb{C}^{N_R \times 1}$, received at the user u in the g th multicast group can be formulated as

$$\mathbf{r}_u^{(g)} = \sum_{m=1}^M \mathbf{H}_{u,m}^{(g)} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m}^{(g)} x^{(g)} + \sum_{\substack{l=1 \\ l \neq g}}^G \sum_{m=1}^M \mathbf{H}_{u,m}^{(g)} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m}^{(l)} x^{(l)} + \boldsymbol{\delta}_u^{(g)}, \quad (1)$$

where $\mathbf{H}_{u,m}^{(g)} \in \mathbb{C}^{N_R \times N_T}$ represent the mmWave MIMO channel matrix from the m th AP to the user $U_u^{(g)}$, $\mathbf{F}_{\text{RF},m} \in \mathbb{C}^{N_T \times N_{\text{RF},m}}$ is the RF TPC of the m th AP, and $\mathbf{f}_{\text{BB},m}^{(g)} \in \mathbb{C}^{N_{\text{RF},m} \times 1}$ is the baseband TPC of the m th AP for the user-group g . Furthermore, $\boldsymbol{\delta}_u^{(g)} \in \mathbb{C}^{N_R \times 1}$ represents the complex additive Gaussian noise vector with distribution $\boldsymbol{\delta}_u^{(g)} \sim \mathcal{CN}(0, \sigma_\delta^2 \mathbf{I}_{N_R})$ and $x^{(g)}$ signifies the transmitted symbol that has unit power i.e., $\mathbb{E}[|x^{(g)}|^2] = 1$. The

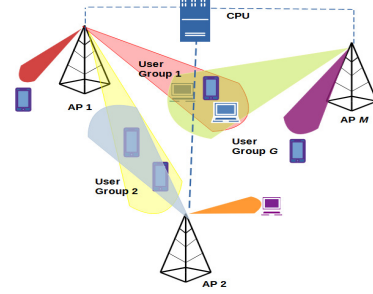


Figure 1: Cooperative beamforming in cell-free mmWave MIMO for multicast scenario.

received signal $\tilde{\mathbf{r}}_u^{(g)} \in \mathbb{C}^{N_{\text{RF},u} \times 1}$ after RF combining at the user u in the g th group is given by

$$\tilde{\mathbf{r}}_u^{(g)} = \left(\mathbf{W}_{\text{RF},u}^{(g)} \right)^H \sum_{m=1}^M \mathbf{H}_{u,m}^{(g)} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m}^{(g)} x^{(g)} + \left(\mathbf{W}_{\text{RF},u}^{(g)} \right)^H \sum_{\substack{l=1 \\ l \neq g}}^G \sum_{m=1}^M \mathbf{H}_{u,m}^{(g)} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m}^{(l)} x^{(l)} + \tilde{\boldsymbol{\delta}}_u^{(g)},$$

where $\mathbf{W}_{\text{RF},u}^{(g)} \in \mathbb{C}^{N_R \times N_{\text{RF},u}}$ denotes the RF RC of the user $U_u^{(g)}$ and $\tilde{\boldsymbol{\delta}}_u^{(g)} = \left(\mathbf{W}_{\text{RF},u}^{(g)} \right)^H \boldsymbol{\delta}_u^{(g)}$. Let us define the effective channel matrix between the user $U_u^{(g)}$ and m th AP as $\mathbf{H}_{\text{eff},u,m}^{(g)} = \left(\mathbf{W}_{\text{RF},u}^{(g)} \right)^H \mathbf{H}_{u,m}^{(g)} \mathbf{F}_{\text{RF},m} \in \mathbb{C}^{N_{\text{RF},u} \times N_{\text{RF},m}}$, and also define the concatenated baseband TPC $\mathbf{f}_{\text{BB}}^{(g)}$ of the g th group as $\mathbf{f}_{\text{BB}}^{(g)} = \left[\left(\mathbf{f}_{\text{BB},1}^{(g)} \right)^H, \left(\mathbf{f}_{\text{BB},2}^{(g)} \right)^H, \dots, \left(\mathbf{f}_{\text{BB},M}^{(g)} \right)^H \right]^H$. The concatenated effective channel matrix between the user $U_u^{(g)}$ and all the cooperating APs is denoted as $\mathbf{H}_{\text{eff},u}^{(g)} = \left[\mathbf{H}_{\text{eff},u,1}^{(g)}, \mathbf{H}_{\text{eff},u,2}^{(g)}, \dots, \mathbf{H}_{\text{eff},u,M}^{(g)} \right] \in \mathbb{C}^{N_{\text{RF},u} \times \left(\sum_{m=1}^M N_{\text{RF},m} \right)}$. Therefore, the effective received signal after baseband combining, $\hat{\mathbf{r}}_u^{(g)}$, at user $U_u^{(g)}$ can be written as

$$\hat{\mathbf{r}}_u^{(g)} = \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \mathbf{H}_{\text{eff},u}^{(g)} \mathbf{f}_{\text{BB}}^{(g)} x^{(g)} + \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \sum_{\substack{l=1 \\ l \neq m}}^G \mathbf{H}_{\text{eff},k}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} x^{(l)} + \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \tilde{\boldsymbol{\delta}}_u^{(g)},$$

where $\mathbf{w}_{\text{BB},u}^{(g)} \in \mathbb{C}^{N_{\text{RF},u} \times 1}$ denotes the baseband RC of the user $U_u^{(g)}$. The SINR of the user $U_u^{(g)}$ can be expressed as

$$\text{SINR}_u^{(g)} = \frac{\left| \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \mathbf{H}_{\text{eff},u}^{(g)} \mathbf{f}_{\text{BB}}^{(g)} \right|^2}{\sum_{l=1, l \neq m}^G \left| \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \mathbf{H}_{\text{eff},u}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} \right|^2 + \sigma_\delta^2}.$$

Furthermore, the downlink capacity represented as C_{DL} is given by

$$C_{\text{DL}} = \sum_{g=1}^G \sum_{u=1}^{U_g} \log_2 \left(1 + \text{SINR}_u^{(g)} \right).$$

A. mmWave MIMO channel model

Using the geometric channel model described in [16], [53] for the mmWave MIMO system, the channel matrix between the m th AP and u th user, denoted by $\mathbf{H}_{u,m} \in \mathbb{C}^{N_R \times N_T}$, can be expressed as

$$\mathbf{H}_{u,m} = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^L \alpha_{u,m}^{(l)} \mathbf{a}_R(\phi_{l,u,m}^R) \mathbf{a}_T^H(\phi_{l,u,m}^T), \quad (2)$$

where L represents the number of dominant multipath components between the m th AP and user u , $\alpha_{u,m}^{(l)}$ denotes the complex-valued path gain of the l th multipath component. Furthermore, the terms $\mathbf{a}_T(\phi_l^T)$ and $\mathbf{a}_R(\phi_l^R)$ denote the transmit and receive array response vectors, respectively, with ϕ_l^R being the angle of arrival (AoA) and ϕ_l^T the angle of departure (AoD). For uniformly spaced linear arrays (ULA), where λ and d denote the signal wavelength and distance between the antenna elements, respectively, the array response vectors $\mathbf{a}_R(\phi_l^R)$ and $\mathbf{a}_T(\phi_l^T)$ are given by

$$\begin{aligned} \mathbf{a}_R(\phi_l^R) &= \frac{1}{\sqrt{N_R}} [1, e^{j \frac{2\pi}{\lambda} d \sin(\phi_l^R)}, \dots, e^{j \frac{2\pi}{\lambda} (N_R-1) d \sin(\phi_l^R)}]^T, \\ \mathbf{a}_T(\phi_l^T) &= \frac{1}{\sqrt{N_T}} [1, e^{j \frac{2\pi}{\lambda} d \sin(\phi_l^T)}, \dots, e^{j \frac{2\pi}{\lambda} (N_T-1) d \sin(\phi_l^T)}]^T. \end{aligned}$$

One can also represent this mmWave MIMO channel in a compact form as

$$\mathbf{H}_{u,m} = \mathbf{A}_{R,u,m} \mathbf{H}_{u,m}^\alpha \mathbf{A}_{T,u,m}^H,$$

where $\mathbf{A}_R = [\mathbf{a}_R(\phi_{1,u,m}^R), \dots, \mathbf{a}_R(\phi_{L,u,m}^R)] \in \mathbb{C}^{N_R \times L}$ and $\mathbf{A}_T = [\mathbf{a}_T(\phi_{1,u,m}^T), \dots, \mathbf{a}_T(\phi_{L,u,m}^T)] \in \mathbb{C}^{N_T \times L}$ contain the array response vectors of the AP and user, respectively, and $\mathbf{H}_{u,m}^\alpha = \text{diag}(\alpha_{u,m}^{(1)}, \alpha_{u,m}^{(2)}, \dots, \alpha_{u,m}^{(L)})$ denotes the diagonal matrix of path gains. Although this channel model is described for a uniform linear array (ULA), one can also readily extend the framework of this paper to uniform planar arrays (UPA). The algorithms and results derived in this work are also valid for a UPA, the corresponding array response vector can be written as

$$\begin{aligned} \mathbf{a}_x(\phi, \theta) &= \frac{1}{\sqrt{N_x}} \left[1, e^{j \frac{2\pi}{\lambda} d (\sin \phi \sin \theta + \cos \theta)}, \dots, \right. \\ &\quad \left. \dots, e^{j \frac{2\pi}{\lambda} d ((N_x^h - 1) \sin \phi \sin \theta + (N_x^v - 1) \cos \theta)} \right], \quad (3) \end{aligned}$$

where $x \in \{R, T\}$, while N_x^h and N_x^v denote the number of horizontal and vertical elements of the UPA in the two-dimension plane of the array. The UPA structure at each AP has a total of N_T elements associated with $N_T = N_T^h \times N_T^v$.

Note that the system model developed in Section-II describes a general multicast scenario, where a group of users may request identical services. A special case of the multicast scenario is popularly termed as a unicast scenario in which each user requests different information. The next section designs HTBFs for a broadcast system, where all the users request the same information, i.e., there is only a single group. Thus, it can also be viewed as a special instance of the multicast scenario, where there is no inter-group or inter-user interference. This interference-free scenario simplifies the HTBF design problem, as described in the next section.

III. HYBRID BROADCAST BEAMFORMING (HBBF)

In a broadcast scenario, the sum signal-to-noise ratio (SSNR) of all the users in the system is maximized subject to the total power constraint (P_T) in order to obtain the optimal HBBF. The motivation for SSNR maximization can be summarised as follows. Note that at low-SNRs, we have $\log_2(1 + \text{SNR}_u) \approx \frac{1}{\ln(2)} \text{SNR}_u$. Therefore, one can deduce that the maximization of capacity is identical to SSNR maximization. Furthermore, it can be observed from the following relationship, which exploits Jensen's inequality, that SSNR maximization provides an upper bound of the true capacity:

$$\frac{1}{U} \sum_{u=1}^U \log_2(1 + \text{SNR}_u) \leq \log_2 \left(1 + \frac{1}{U} \sum_{u=1}^U \text{SNR}_u \right).$$

Considering a broadcast scenario to U users, the signal $\mathbf{r}_u \in \mathbb{C}^{N_R \times 1}$ received at user u can be expressed as

$$\mathbf{r}_u = \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m} x + \delta_u,$$

where $\mathbf{H}_{u,m} \in \mathbb{C}^{N_R \times N_T}$ represents the channel matrix between the user u and the m th AP, $\mathbf{F}_{\text{RF},m} \in \mathbb{C}^{N_T \times N_{\text{RF},m}}$ and $\mathbf{f}_{\text{BB},m} \in \mathbb{C}^{N_{\text{RF},m} \times 1}$ denote the RF and baseband TPCs respectively, at the m th AP. Furthermore, $\delta_u \in \mathbb{C}^{N_R \times 1}$ represents the AWGN at user u and x denotes the transmitted symbol corresponding to the broadcast scenario, which is normalized to have unit power i.e., $\mathbb{E}[|x|^2] = 1$. After RF combining using the combiner matrix $\mathbf{W}_{\text{RF},u} \in \mathbb{C}^{N_R \times N_{\text{RF},u}}$, the received signal $\tilde{\mathbf{r}}_u$ can be written as

$$\tilde{\mathbf{r}}_u = \mathbf{W}_{\text{RF},u}^H \left(\sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},m} x \right) + \tilde{\delta}_u, \quad (4)$$

where $\tilde{\delta}_u = \mathbf{W}_{\text{RF},u}^H \delta_u$. Note that the columns of the RF TPC $\mathbf{F}_{\text{RF},m}$ can be chosen as the dominant $N_{\text{RF},m}$ transmit array steering vectors corresponding to each user. Similarly, the RF RC $\mathbf{W}_{\text{RF},u}$ can be configured with the $N_{\text{RF},u}$ dominant receive array steering vectors corresponding to each AP. Thus, to design the RF precoder for the m th AP, one can

first find the index $\mathcal{K}_{u,m}$ of the multipath component corresponding to the u th user, which has the maximum magnitude path gain, i.e., $\mathcal{K}_{u,m} = \max \left\{ \left| \alpha_{u,m}^{(1)} \right|, \left| \alpha_{u,m}^{(2)} \right|, \dots, \left| \alpha_{u,m}^{(L)} \right| \right\}$. Then, the RF TPC $\mathbf{F}_{\text{RF},m}$ for the m th AP can be obtained by extracting the $N_{\text{RF},m}$ columns indexed by the set $\mathcal{K}_m = \{\mathcal{K}_{1,m}, \mathcal{K}_{2,m}, \dots, \mathcal{K}_{U,m}\}$ from the matrices $\mathbf{A}_{T,u,m}$. This can be mathematically represented as

$$\mathbf{F}_{\text{RF},m} = [\mathbf{A}_{T,1,m}(:, \mathcal{K}_{1,m}), \dots, \mathbf{A}_{T,U,m}(:, \mathcal{K}_{U,m})]. \quad (5)$$

In a similar fashion, one can also obtain the RF RC $\mathbf{W}_{\text{RF},u}$ for each user. Let the effective channel matrix $\mathbf{H}_{\text{eff},u} \in \mathbb{C}^{N_{\text{RF},u} \times N_{\text{RF},m}}$ between the user u and the m th AP be defined as $\mathbf{H}_{\text{eff},u,m} = \mathbf{W}_{\text{RF},u}^H \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m}$. Furthermore, let the concatenated effective channel matrix $\mathbf{H}_{\text{eff},u} \in \mathbb{C}^{N_{\text{RF},u} \times \sum_{m=1}^M N_{\text{RF},m}}$ for user u be defined as

$$\mathbf{H}_{\text{eff},u} = [\mathbf{H}_{\text{eff},u,1} \mathbf{H}_{\text{eff},u,2} \cdots \mathbf{H}_{\text{eff},u,M}].$$

The received signal $\tilde{\mathbf{r}}_u$ of (4) can be completely rewritten as

$$\tilde{\mathbf{r}}_u = \sum_{m=1}^M \mathbf{H}_{\text{eff},u,m} \mathbf{f}_{\text{BB},m} x + \tilde{\boldsymbol{\delta}}_u = \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB}} x + \tilde{\boldsymbol{\delta}}_u, \quad (6)$$

where $\mathbf{f}_{\text{BB}} = [\mathbf{f}_{\text{BB},1}^T, \mathbf{f}_{\text{BB},2}^T, \dots, \mathbf{f}_{\text{BB},M}^T]^T \in \mathbb{C}^{\sum_{m=1}^M N_{\text{RF},m} \times 1}$ denotes the stacked baseband TPs of all the APs. Note that the signal power received at user u is given as $\mathbf{f}_{\text{BB}}^H \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB}}$, which implies that the total received power for all the users in the system can be written as $\sum_{u=1}^U \mathbf{f}_{\text{BB}}^H \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB}} = \mathbf{f}_{\text{BB}}^H \mathbf{H} \mathbf{f}_{\text{BB}}$, where $\mathbf{H} = \sum_{u=1}^U \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u}$. Our objective is to find the optimal stacked baseband TPC \mathbf{f}_{BB} for the HBBF in order to maximize the SSNR. Toward this, the optimization problem can be formulated as

$$\begin{aligned} & \underset{\mathbf{f}_{\text{BB}}}{\text{maximize}} && \mathbf{f}_{\text{BB}}^H \mathbf{H} \mathbf{f}_{\text{BB}} \\ & \text{subject to} && \|\mathbf{f}_{\text{BB}}\|^2 \leq P_T. \end{aligned}$$

The optimal baseband TPC \mathbf{f}_{BB} for the optimization problem above is given by $\mathbf{f}_{\text{BB}} = \sqrt{P_T} \boldsymbol{\lambda}_{\max}(\mathbf{H})$, where $\boldsymbol{\lambda}_{\max}(\mathbf{H})$ represents the eigenvector corresponding to the largest eigenvalue of the matrix \mathbf{H} . Interestingly, it can be observed that the SSNR of the mmWave broadcast system can be maximized by sending the information in the direction of the leading eigenmode of the effective channel matrix \mathbf{H} . Substituting this optimal \mathbf{f}_{BB} into (6), we have

$$\tilde{\mathbf{r}}_u = \sqrt{P_T} \mathbf{H}_{\text{eff},u} \boldsymbol{\lambda}_{\max}(\mathbf{H}) x + \tilde{\boldsymbol{\delta}}_u. \quad (7)$$

The covariance matrix $\mathbf{R}_{\tilde{\boldsymbol{\delta}}_u} \in \mathbb{C}^{N_{\text{RF},u} \times N_{\text{RF},u}}$ of the combined noise vector $\tilde{\boldsymbol{\delta}}_u$ can be expressed as

$$\mathbf{R}_{\tilde{\boldsymbol{\delta}}_u} = \mathbb{E} [\tilde{\boldsymbol{\delta}}_u \tilde{\boldsymbol{\delta}}_u^H] = \sigma_{\tilde{\boldsymbol{\delta}}}^2 \mathbf{W}_{\text{RF},u}^H \mathbf{W}_{\text{RF},u}. \quad (8)$$

Therefore, the whitened system model can be written as

$$\mathbf{R}_{\tilde{\boldsymbol{\delta}}_u}^{-\frac{1}{2}} \tilde{\mathbf{r}}_u = \sqrt{P_T} \mathbf{R}_{\tilde{\boldsymbol{\delta}}_u}^{-\frac{1}{2}} \mathbf{H}_{\text{eff},u} \boldsymbol{\lambda}_{\max}(\mathbf{H}) x + \mathbf{R}_{\tilde{\boldsymbol{\delta}}_u}^{-\frac{1}{2}} \tilde{\boldsymbol{\delta}}_u. \quad (9)$$

The optimal baseband RC $\mathbf{w}_{\text{BB},u}$ for user u can now be readily found using the MRC criterion as

$$\mathbf{w}_{\text{BB},u} = \frac{\mathbf{R}_{\tilde{\boldsymbol{\delta}}_u}^{-\frac{1}{2}} \mathbf{H}_{\text{eff},u} \boldsymbol{\lambda}_{\max}(\mathbf{H})}{\left\| \mathbf{R}_{\tilde{\boldsymbol{\delta}}_u}^{-\frac{1}{2}} \mathbf{H}_{\text{eff},u} \boldsymbol{\lambda}_{\max}(\mathbf{H}) \right\|}. \quad (10)$$

Next, we derive the cooperative HBBF for a broadcast system having per AP-power constraints. Furthermore, we also maximize the minimum SNR of the users in the system to incorporate fairness in the power allocation.

A. HBBF-P

This sub-section considers per-AP power constraints for performing cooperative HBBF for CFMU systems that maximizes the sum-SNR. Let us assume the power among the APs to be equally distributed as $\frac{P_T}{M}$ in the broadcast group. The associated optimization problem of the CFMU HBBF, which maximizes the sum-SNR, is given by

$$\begin{aligned} & \underset{\mathbf{F}_{\text{BB},u}}{\text{maximize}} && \sum_{u=1}^U \text{Tr}(\mathbf{f}_{\text{BB}}^H \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB}}) \\ & \equiv && \sum_{u=1}^U \sum_{m=1}^M \text{Tr}(\mathbf{H}_{\text{eff},u,m}^H \mathbf{H}_{\text{eff},u,m} \mathbf{F}_{\text{BB},m}) \end{aligned}$$

$$\text{subject to } \text{Tr}(\mathbf{F}_{\text{BB},m}) \leq \frac{P_T}{M}, \quad m = 1, 2, \dots, M$$

$$\mathbf{F}_{\text{BB},m} \succeq 0 \quad \text{and} \quad \text{rank}(\mathbf{F}_{\text{BB},m}) = 1 \quad \forall m,$$

where $\mathbf{F}_{\text{BB},m} = \mathbf{f}_{\text{BB},m} \mathbf{f}_{\text{BB},m}^H \in \mathbb{C}^{N_{\text{RF},m} \times N_{\text{RF},m}}$ denotes a rank-1 positive semi-definite (PSD) matrix. The above optimization problem is non-convex due to the rank-1 constraint, hence it is difficult to solve. Upon relaxing this rank-1 constraint, the equivalent semi-definite programme (SDP) can be expressed as

$$\begin{aligned} & \underset{\mathbf{F}_{\text{BB},m}}{\text{maximize}} && \sum_{u=1}^U \sum_{m=1}^M \text{Tr}(\mathbf{H}_{\text{eff},u,m}^H \mathbf{H}_{\text{eff},u,m} \mathbf{F}_{\text{BB},m}) \\ & \text{subject to} && \text{Tr}(\mathbf{F}_{\text{BB},m}) \leq \frac{P_T}{M}, \quad m = 1, 2, \dots, M, \\ & && \mathbf{F}_{\text{BB},m} \succeq 0 \quad \forall m. \end{aligned} \quad (11)$$

B. HBBF-P with a max-min per-AP power constraint

In order to ensure fairness in the power allocation, the minimum SNR of the users can be maximized. Towards this, let the target SNR of each user be denoted by γ . To perform HBBF, the max-min fairness based optimization problem can be framed as

$$\begin{aligned} & \underset{\gamma, \mathbf{F}_{\text{BB},m}}{\text{maximize}} && \gamma \\ & \text{s.t.} && \sum_{m=1}^M \text{Tr}(\mathbf{H}_{\text{eff},u,m}^H \mathbf{H}_{\text{eff},u,m} \mathbf{F}_{\text{BB},m}) \geq \gamma, \quad \forall u, \\ & && \text{Tr}(\mathbf{F}_{\text{BB},m}) \leq \frac{P_T}{M}, \quad m = 1, 2, \dots, M, \\ & && \mathbf{F}_{\text{BB},m} \succeq 0, \quad \forall m. \end{aligned} \quad (12)$$

It can be observed that the optimization problems (11) and (12) can be efficiently solved using widely available tools, such as CVX [54]. The baseband beamformer $\mathbf{f}_{\text{BB},m}$ can be finally determined as the eigenvector corresponding to the largest eigenvalue of $\mathbf{F}_{\text{BB},m}^*$, i.e., $\mathbf{f}_{\text{BB},m}^{\text{opt}} = \sqrt{\lambda_{\max}} \hat{\mathbf{f}}_{\text{BB},m}$, where $\hat{\mathbf{f}}_{\text{BB},m}$ represents the unit-norm eigenvector corresponding to the eigenvalue λ_{\max} . Note that the relaxed SDP based beamformer obtained is an approximation of the optimal beamformer. It has been shown that the SDP relaxation results in a close approximation of the optimization problems involving non-convex quadratically constrained quadratic programs (QCQPs) (11) and max-min-based quadratic programs (12). Furthermore, it has been demonstrated that for SDP relaxation, the rank of an optimal solution is upper bounded by \sqrt{M} [55]. For practical values of M , the experimental studies in [55] demonstrate that the average approximation ratio is close to unity. The computational complexity order of HBBF-P scheme is $\mathcal{O} \left[(MN_{\text{RF},m})^{3.5} \right]$ [56], [57].

IV. UNICAST SCENARIO: OPTIMAL SUCCESSIVE HYBRID TRANSMIT BEAMFORMING

This section jointly designs the optimal HTBFs and RCs for the mmWave unicast scenarios by maximizing the SINR in successive order. The algorithm developed herein will be extended to multicast systems in the next section. In this setting, there are $G = U$ user groups with one user in each group, i.e., $U_g = 1$, where U denotes the total number of users present in the system. The signal $\mathbf{r}_u \in \mathbb{C}^{N_R \times 1}$ received at the u th user can be written as

$$\mathbf{r}_u = \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},u,m} x_u + \sum_{\substack{i=1 \\ i \neq u}}^U \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},i,m} x_i + \boldsymbol{\delta}_u, \quad (13)$$

where $\boldsymbol{\delta}_u \in \mathbb{C}^{N_R \times 1}$ represents the AWGN at user u and x_u denotes the symbol transmitted to user u , which is normalized to have unit power. Furthermore, $\mathbf{f}_{\text{BB},u,m} \in \mathbb{C}^{N_{\text{RF},m} \times 1}$ represents the baseband TPC at the m th AP for the user u . After employing the RF RC $\mathbf{W}_{\text{RF},u} \in \mathbb{C}^{N_u \times N_{\text{RF},u}}$, the received signal, $\tilde{\mathbf{r}}_u = \mathbf{W}_{\text{RF},u}^H \mathbf{r}_u$, is given as:

$$\tilde{\mathbf{r}}_u = \mathbf{W}_{\text{RF},u}^H \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},u,m} x_u + \mathbf{W}_{\text{RF},u}^H \sum_{\substack{i=1 \\ i \neq u}}^U \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m} \mathbf{f}_{\text{BB},i,m} x_i + \tilde{\boldsymbol{\delta}}_u, \quad (14)$$

where $\tilde{\boldsymbol{\delta}}_u = \mathbf{W}_{\text{RF},u}^H \boldsymbol{\delta}_u$. Let $\mathbf{H}_{\text{eff},u,m} = \mathbf{W}_{\text{RF},u}^H \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},m}$ denote the effective channel matrix between the m th AP and user u . Let the concatenated baseband TPC and the effective channel matrix of the u th user be defined as

$$\mathbf{f}_{\text{BB},u} = [\mathbf{f}_{\text{BB},u,1}^T, \mathbf{f}_{\text{BB},u,2}^T, \dots, \mathbf{f}_{\text{BB},u,M}^T]^T,$$

$$\mathbf{H}_{\text{eff},u} = [\mathbf{H}_{\text{eff},u,1}, \mathbf{H}_{\text{eff},u,2}, \dots, \mathbf{H}_{\text{eff},u,M}],$$

respectively. After baseband combining at user u employing the RC $\mathbf{w}_{\text{BB},u} \in \mathbb{C}^{N_{\text{RF},u} \times 1}$, the estimate \hat{x}_u of the symbol x_u is given by

$$\hat{x}_u = \mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},u} x_u + \sum_{i=1, i \neq u}^U \mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},i} x_i + \mathbf{w}_{\text{BB},u}^H \boldsymbol{\delta}_u. \quad (15)$$

Note that the first and second terms in (15) correspond to the desired signal and inter-user interference, respectively. Next, we develop the OSHB-U scheme similar to the successive optimization method in multi-user MIMO systems [58] that employs cooperative beamforming for iteratively multiplexing multiple users in the downlink, together with exploiting the MVDR beamformer of [59] for interference minimization in the mmWave cellular system. Furthermore, it maximizes the SINR of each user while cancelling the interferences imposed on the users, which have already been scheduled. Toward this, we now derive the proposed OSHB-U technique for a unicast scenario.

To begin with, it can be readily observed that while scheduling the u th user, the zero-forcing (ZF) interference constraint of $\mathbf{w}_{\text{BB},v}^H \mathbf{H}_{\text{eff},v} \mathbf{f}_{\text{BB},u} = 0$, $1 \leq v \leq u-1$, successfully eliminates the interference imposed by the previously scheduled users. Hence, the baseband TPC is given by $\mathbf{f}_{\text{BB},u} = \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u$, where $\boldsymbol{\nu}_u \in \mathbb{C}^{\sum_{m=1}^M N_{\text{RF},m} - u + 1}$, and $\mathbf{M}_{u-1}^\perp \in \mathbb{C}^{\sum_{m=1}^M N_{\text{RF},m} \times (\sum_{m=1}^M N_{\text{RF},m} - u + 1)}$ represents the orthonormal basis to the null space of the matrix $\mathbf{M}_{u-1} \in \mathbb{C}^{(u-1) \times \sum_{m=1}^M N_{\text{RF},m}}$, which is given by

$$\mathbf{M}_{u-1} = \left[(\mathbf{w}_{\text{BB},1}^H \mathbf{H}_{\text{eff},1})^T, \dots, (\mathbf{w}_{\text{BB},u-1}^H \mathbf{H}_{\text{eff},u-1})^T \right]^T. \quad (16)$$

The matrix \mathbf{M}_{u-1}^\perp can be readily derived from the singular value decomposition (SVD) of the matrix \mathbf{M}_{u-1} , and it corresponds to the right singular vectors matrix composed of $\left(\sum_{m=1}^M N_{\text{RF},m} - u + 1 \right)$. Furthermore, the transmissions of the successive users $u+1, u+2, \dots, U$ are also orthogonal to those of the u th user, i.e., $\mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},v} = 0$, $u+1 \leq v \leq U$. Therefore, the effective received signal after RF precoding and combining at user u can be expressed as

$$\begin{aligned} \tilde{\mathbf{r}}_u &= \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},u} x_u + \sum_{i=1}^{u-1} \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},i} x_i + \tilde{\boldsymbol{\delta}}_u \\ &= \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u x_u + \sum_{i=1}^{u-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{i-1}^\perp \boldsymbol{\nu}_i x_i + \tilde{\boldsymbol{\delta}}_u. \end{aligned} \quad (17)$$

Let the interference and noise terms in (17) be collectively denoted by $\boldsymbol{\tau}_u = \sum_{i=1}^{u-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{i-1}^\perp \boldsymbol{\nu}_i x_i + \tilde{\boldsymbol{\delta}}_u$. Its covariance matrix $\mathbf{C}_{\boldsymbol{\tau}_u} \in \mathbb{C}^{N_{\text{RF},u} \times N_{\text{RF},u}}$ is given by

$$\begin{aligned}
\mathbf{C}_{\tau_u} &= \sum_{i=1}^{u-1} \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},i} \left(\mathbb{E} [|x_i|^2] \right) \mathbf{f}_{\text{BB},i}^H \mathbf{H}_{\text{eff},u}^H + \mathbb{E} \left[\delta_u \delta_u^H \right] \\
&= \sum_{i=1}^{u-1} \mathbf{H}_{\text{eff},u} \left(\mathbf{M}_{i-1}^\perp \boldsymbol{\nu}_i \boldsymbol{\nu}_i^H \left(\mathbf{M}_{i-1}^\perp \right)^H \right) \mathbf{H}_{\text{eff},u}^H \\
&\quad + \sigma_\delta^2 \mathbf{W}_{\text{RF},u}^H \mathbf{W}_{\text{RF},u}. \tag{18}
\end{aligned}$$

To obtain the optimal baseband TPC and RC, one can write the effective interference-whitened system model corresponding to (17) as

$$\mathbf{C}_{\tau_u}^{-\frac{1}{2}} \tilde{\mathbf{r}}_u = \mathbf{C}_{\tau_u}^{-\frac{1}{2}} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u x_u + \mathbf{C}_{\tau_u}^{-\frac{1}{2}} \boldsymbol{\tau}_u. \tag{19}$$

Therefore, the optimal baseband TPC and RC for user u , which maximize the SINR are given by $\mathbf{f}_{\text{BB},u} = \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u$ and $\mathbf{w}_{\text{BB},u} = \left(\mathbf{C}_{\tau_u}^{-1/2} \right)^H \tilde{\mathbf{w}}_{\text{BB},u}$, where $\tilde{\mathbf{w}}_{\text{BB},u}$ and $\boldsymbol{\nu}_u$ are the principal left and right singular vectors, respectively, of the matrix $\mathbf{C}_{\tau_u}^{-1/2} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \in \mathbb{C}^{N_{\text{RF},u} \times (MN_{\text{RF},m} - u + 1)}$. The detailed proofs of the results obtained above are given in Appendix A. The optimal vector $\mathbf{w}_{\text{BB},u} = \left(\mathbf{C}_{\tau_u}^{-1/2} \right)^H \tilde{\mathbf{w}}_{\text{BB},u}$ is known as the MVDR beamformer that maximizes the SINR. This is termed as MVDR, because the output signal at user u has the minimum variance and the desired signal is not distorted. Therefore, the minimum variance distortionless response (MVDR) together with the optimal successive minimum variance hybrid beamforming (OSHB) results in the optimal TPC that maximizes the SINR, while also nulling the interference emanating from the previously scheduled $u-1$ users. Note that the proposed optimal hybrid beamformer is obtained in two steps. In the first step, we employ the CSI toward design of the optimal RF TPCs. In the second step, the proposed baseband TPC is derived based on the MVDR beamformer. Therefore, this results in an optimal hybrid beamformer that maximizes the SINR along with interference cancellation. Algorithm 1 presents the step-by-step procedure of the OSHB algorithm for obtaining the baseband TPC and RC. The key features of the proposed OSHB technique for the mmWave unicast CFMU MIMO system are described next.

To begin with, note that the interference cancellation procedure of the proposed OSHB method is similar to the classic successive interference cancellation (SIC) methods, wherein once the signal of user u is successfully recovered, one can remodulate it and then subtract it from the received composite signal for decontaminating it and for reducing the decoding complexity of the remaining users at the APs. Hence, the baseband TPC configured for the u th user is specifically chosen for ensuring that the interference emanating from the formerly selected $(u-1)$ users is zero. This can be realized by selecting the TPC for ensuring that it lies in the subspace orthogonal to the column space of the matrix \mathbf{M}_{u-1} . Furthermore, one can observe that the maximum number of users supported by the OSHB-U scheme is $\sum_{m=1}^M N_{\text{RF},m}$. In order to cancel the interference imposed on user u , the number of columns in the null space of the matrix $\mathbf{M}_{u-1}^\perp \in \mathbb{C}^{\sum_{m=1}^M N_{\text{RF},m} \times (\sum_{m=1}^M N_{\text{RF},m} - u + 1)}$ must be

Algorithm 1: OSHB algorithm

- 1 **Input:** Effective channel matrix $\mathbf{H}_{\text{eff},u}$, $\mathbf{W}_{\text{RF},u}$
 - 2 **Initialization:** $\mathbf{M}_0 = \mathbf{I}_{1 \times MN_{\text{RF},m}}$,
 $\mathbf{C}_{\tau_1} = \mathbf{W}_{\text{RF},1}^H \mathbf{W}_{\text{RF},1}$
 - 3 **for** $u = 1 : U$ **do**
 - 1) $[\mathbf{U} \ \mathbf{S} \ \mathbf{V}] = \text{svd}(\mathbf{M}_{u-1})$
 - 2) $\mathbf{M}_{u-1}^\perp = \mathbf{V}(:, u : MN_{\text{RF},m})$
 - 3) $[\mathbf{U}' \ \mathbf{S}' \ \mathbf{V}'] = \text{svd}\left(\mathbf{C}_{\tau_u}^{-\frac{1}{2}} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp\right)$
 - 4) $\boldsymbol{\nu}_u = \mathbf{V}'(:, 1)$
 - 5) $\mathbf{f}_{\text{BB},u} = \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u$
 - 6) $\mathbf{w}_{\text{BB},u} = \left(\mathbf{C}_{\tau_u}^{-\frac{1}{2}}\right)^H \mathbf{U}'(:, 1)$
 - 7) $\mathbf{C}_{\tau_u} \rightarrow$ update according to (18)
 - 8) $\mathbf{M}_{u-1} \rightarrow$ update according to (16)
-
- end for**
-

larger than or equal to one, i.e. $\sum_{m=1}^M N_{\text{RF},m} - u + 1 \geq 1$, which yields $U \leq \sum_{m=1}^M N_{\text{RF},m}$. This implies that the number of RF chains at each AP has to be increased for multiplexing a higher number of users. On the other hand, if one considers the overall effective channel matrix $\mathbf{H}_{\text{eff}} = [\mathbf{H}_{\text{eff},1}^T, \mathbf{H}_{\text{eff},u}^T, \dots, \mathbf{H}_{\text{eff},U}^T]^T \in \mathbb{C}^{UN_{\text{RF},u} \times MN_{\text{RF},m}}$ to perform block diagonalization, then one has to satisfy the constraint $UN_{\text{RF},u} \leq \sum_{m=1}^M N_{\text{RF},m}$. Hence, the block diagonalization method specifically designed for perfect interference removal can only support $\lfloor \sum_{m=1}^M N_{\text{RF},m} / N_{\text{RF},u} \rfloor$ users. Therefore, it can be readily observed that a substantially higher number of users can be served using the proposed OSHB scheme than by the conventional block diagonalization approach, which demonstrates its practical suitability.

V. OPTIMAL SUCCESSIVE HYBRID BEAMFORMING FOR MULTICAST SCENARIOS

As described previously, in a multicast scenario, the users belonging to the same multicast group tend to request the same data. For such a system, one can develop the OSHB-M scheme, similar to the OSHB-U scheme derived for the unicast scenario, which successfully cancels the interference arising from the previously chosen groups $1, 2, \dots, g-1$ contaminating the g th group. Furthermore, to minimize the interference generated by the previously chosen groups, the MVDR-based beamformer can be employed for the g th group.

As described in Section II, the signal $\tilde{\mathbf{r}}_u^{(g)} \in \mathbb{C}^{N_{\text{RF},u} \times 1}$ received for the user $U_u^{(g)}$ at the output of the RF RC is given as

$$\tilde{\mathbf{r}}_u^{(g)} = \mathbf{H}_{\text{eff},u}^{(g)} \mathbf{f}_{\text{BB}}^{(g)} x^{(g)} + \sum_{\substack{l=1 \\ l \neq g}}^G \mathbf{H}_{\text{eff},u}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} x^{(l)} + \check{\delta}_u^{(g)}. \tag{20}$$

The received signal concatenated for all the users in group g can be formulated as

$$\underbrace{\begin{bmatrix} \mathbf{r}_1^{(g)} \\ \mathbf{r}_2^{(g)} \\ \vdots \\ \mathbf{r}_{U_g}^{(g)} \end{bmatrix}}_{\mathbf{r}^{(g)}} = \underbrace{\begin{bmatrix} \mathbf{H}_{\text{eff},1}^{(g)} \\ \mathbf{H}_{\text{eff},2}^{(g)} \\ \vdots \\ \mathbf{H}_{\text{eff},U_g}^{(g)} \end{bmatrix}}_{\mathbf{H}_{\text{eff}}^{(g)}} \mathbf{f}_{\text{BB}}^{(g)} x^{(g)} + \sum_{\substack{l=1 \\ l \neq g}}^G \mathbf{H}_{\text{eff}}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} x^{(l)} + \underbrace{\begin{bmatrix} \delta_1^{(g)} \\ \delta_2^{(g)} \\ \vdots \\ \delta_{U_g}^{(g)} \end{bmatrix}}_{\boldsymbol{\delta}^{(g)}}. \quad (21)$$

Let $\mathbf{w}_{\text{BB}}^{(g)} = \left[\left(\mathbf{w}_{\text{BB},1}^{(g)} \right)^T, \left(\mathbf{w}_{\text{BB},2}^{(g)} \right)^T, \dots, \left(\mathbf{w}_{\text{BB},U_g}^{(g)} \right)^T \right]^T$ denote the concatenated baseband combiner of all the users for the g th group. The signal $\tilde{r}_u^{(g)} = \left(\mathbf{w}_{\text{BB},u}^{(g)} \right)^H \hat{\mathbf{r}}_u^{(g)}$ received by the g th group after baseband RC is given by

$$\tilde{r}_u^{(g)} = \left(\mathbf{w}_{\text{BB}}^{(g)} \right)^H \mathbf{H}_{\text{eff}}^{(g)} \mathbf{f}_{\text{BB}}^{(g)} x^{(g)} + \left(\mathbf{w}_{\text{BB}}^{(g)} \right)^H \sum_{\substack{l=1 \\ l \neq g}}^G \mathbf{H}_{\text{eff}}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} x^{(l)} + \tilde{\delta}^{(g)}. \quad (22)$$

The proposed OSHB-M algorithm employs the ZF constraint $\mathbf{G}^{(g)} \mathbf{f}_{\text{BB}}^{(g)} = 0$ for cancelling the interference denoted by the second term in (22), where the matrix $\mathbf{G}^{(g)} \in \mathbb{C}^{(\sum_{j=1}^{g-1} U_j) \times \sum_{m=1}^M N_{\text{RF},m}}$ is defined as $\mathbf{G}^{(g)} = \left[\left(\mathbf{F}^{(1)} \right)^T, \left(\mathbf{F}^{(2)} \right)^T, \dots, \left(\mathbf{F}^{(g-1)} \right)^T \right]^T$, with

$$\mathbf{F}^j = \begin{bmatrix} \left(\mathbf{w}_{\text{BB},1}^{(g)} \right)^H \mathbf{H}_{\text{eff},1}^{(j)} \\ \left(\mathbf{w}_{\text{BB},2}^{(g)} \right)^H \mathbf{H}_{\text{eff},2}^{(j)} \\ \vdots \\ \left(\mathbf{w}_{\text{BB},U_j}^{(g)} \right)^H \mathbf{H}_{\text{eff},U_j}^{(j)} \end{bmatrix} \in \mathbb{C}^{U_j \times M N_{\text{RF},m}}. \quad (23)$$

Hence, the baseband TPC constructed for the g th group is given by $\mathbf{f}_{\text{BB}}^{(g)} = \left(\mathbf{G}^{(g)} \right)^\perp \boldsymbol{\lambda}^{(g)}$, where $\boldsymbol{\lambda}^{(g)} \in \mathbb{C}^{(\sum_{m=1}^M N_{\text{RF},m} - \sum_{j=1}^{g-1} U_j) \times 1}$, and the columns of the matrix $\left(\mathbf{G}^{(g)} \right)^\perp \in \mathbb{C}^{\sum_{m=1}^M N_{\text{RF},m} \times (M N_{\text{RF},m} - \sum_{j=1}^{g-1} U_j)}$ constitute an orthonormal basis for the null-space of the matrix $\mathbf{G}^{(g)}$, and can be obtained as described in Section IV for the matrix \mathbf{G}_{u-1}^\perp . Next, we denote the interference-plus-noise part $\boldsymbol{\tau}^{(g)}$ in (21) as $\boldsymbol{\tau}^{(g)} = \sum_{l=1}^{g-1} \mathbf{H}_{\text{eff}}^{(g)} \mathbf{f}_{\text{BB}}^{(l)} x^{(l)} + \boldsymbol{\delta}^{(g)}$. The covariance matrix $\mathbf{C}_\tau^{(g)} = \mathbb{E}[\boldsymbol{\tau}^{(g)} (\boldsymbol{\tau}^{(g)})^H] \in \mathbb{C}^{U_g N_{\text{RF},u} \times U_g N_{\text{RF},u}}$ of $\boldsymbol{\tau}^{(g)}$ for the g th multicast group can be written as

$$\mathbf{C}_\tau^{(g)} = \sum_{l=1}^{g-1} \mathbf{H}_{\text{eff}}^{(g)} \left(\mathbf{f}_{\text{BB}}^{(l)} \left(\mathbf{f}_{\text{BB}}^{(l)} \right)^H \right) \left(\mathbf{H}_{\text{eff}}^{(g)} \right)^H + \sigma_\delta^2 \mathbf{I}_{U_g N_{\text{RF},u}}.$$

Therefore, the optimal baseband TPC $\mathbf{f}_{\text{BB}}^{(g)}$ and RC $\mathbf{w}_{\text{BB}}^{(g)}$ conditioned for the g th multicast group are given as

$$\mathbf{f}_{\text{BB}}^{(g)} = \left(\mathbf{G}^{(g)} \right)^\perp \boldsymbol{\lambda}^{(g)} \quad \text{and} \quad \mathbf{w}_{\text{BB}}^{(g)} = \left(\left(\mathbf{C}_\tau^{(g)} \right)^{-1/2} \right)^H \tilde{\mathbf{w}}_{\text{BB}}^{(g)},$$

where $\tilde{\mathbf{w}}_{\text{BB}}^{(g)}$ and $\boldsymbol{\lambda}^{(g)}$ represent the principal left and right singular vectors, respectively, of the matrix $\tilde{\mathbf{H}}_{\text{eff}}^{(g)} = \left(\mathbf{C}_\tau^{(g)} \right)^{-1/2} \mathbf{H}_{\text{eff}}^{(g)} \left(\mathbf{G}^{(g)} \right)^\perp \in$

$\mathbb{C}^{U_g N_{\text{RF},u} \times (M N_{\text{RF},m} - \sum_{j=1}^{g-1} U_j)}$. This result follows from our analysis similar to that provided for the unicast scenario described in Appendix A. Finally, it can be readily observed that the proposed OSHB-U and OSHB-M schemes can be implemented at the complexity order of $\mathcal{O} \left(\left(\sum_{m=1}^M N_{\text{RF},m} \right)^3 \right)$. This is due to the fact that these schemes are based on the singular value decomposition of the matrices \mathbf{M}_{u-1} and $\mathbf{G}^{(g)}$. Note that while deriving the multicast beamforming solutions we assumed that the multicast users are in the same geographical location. However, if the users that request the same data are in different geographical regions, it is more efficient to transmit separately, i.e., unicast, to each user [60].

VI. HYBRID PRECODER DESIGN USING THE BAYESIAN LEARNING (BL)-BASED APPROACH

One can note that in the previous sections, the designs of the RF TPCs and RCs require the knowledge of the dominant array response vectors of the APs and users, respectively, which eventually necessitates the full knowledge of the AoAs/AoDs of the multipath components and their complex-valued path gains. However, in practice, it is challenging to acquire accurate estimates of these quantities, which motivates us to develop a scheme that can efficiently design the hybrid TPCs and RCs without requiring the knowledge of the full channel state information (CSI).

To this end, in this section, we jointly design the baseband and RF TPCs and RCs for a unicast scenario. However, a similar approach can be readily developed for a general multicast scenario. The proposed joint design first obtains the optimal digital baseband TPC and RC \mathbf{f}_{opt} and \mathbf{w}_{opt} , respectively, assuming a fully digital architecture. This step only requires the knowledge of the mmWave MIMO channel matrices $\mathbf{H}_{u,m}$, rather than their component AoAs/AoDs and path gains. Subsequently, a BL-based algorithm decomposes the fully digital precoders/combiners \mathbf{f}_{opt} and \mathbf{w}_{opt} into their RF and baseband components. The signal $\mathbf{r}_u \in \mathbb{C}^{N_R \times 1}$ received by the u th user can be expressed as

$$\mathbf{r}_u = \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{f}_{u,m} x_u + \sum_{\substack{i=1 \\ i \neq u}}^U \sum_{m=1}^M \mathbf{H}_{u,m} \mathbf{f}_{i,m} x_i + \boldsymbol{\delta}_u. \quad (24)$$

Let the concatenated fully digital TPC and channel matrix for the u th user be defined as $\mathbf{f}_u = [\mathbf{f}_{u,1}^T, \mathbf{f}_{u,2}^T, \dots, \mathbf{f}_{u,M}^T]^T$, and $\mathbf{H}_u = [\mathbf{H}_{u,1}, \mathbf{H}_{u,2}, \dots, \mathbf{H}_{u,M}]$, respectively. Employing the fully digital RC $\mathbf{w}_u \in \mathbb{C}^{N_R \times 1}$, the signal $\tilde{r}_u = \mathbf{w}_u^H \mathbf{r}_u$ received by the u th user can be written as

$$\tilde{r}_u = \mathbf{w}_u^H \mathbf{H}_u \mathbf{f}_u x_u + \sum_{\substack{i=1 \\ i \neq u}}^U \mathbf{w}_u^H \mathbf{H}_u \mathbf{f}_i x_i + \underbrace{\mathbf{w}_u^H \boldsymbol{\delta}_u}_{\tilde{\delta}_u}. \quad (25)$$

The SINR of the u th user is given by

$$\text{SINR}_u = \frac{|\mathbf{w}_u^H \mathbf{H}_u \mathbf{f}_u|^2}{\sum_{i=1, i \neq u}^U |\mathbf{w}_u^H \mathbf{H}_u \mathbf{f}_i|^2 + \sigma_\delta^2}. \quad (26)$$

Upon employing the proposed OSHB-U algorithm, the optimal fully digital TPC and RC are obtained as $\mathbf{f}_u = \mathbf{A}_{u-1}^\perp \boldsymbol{\nu}_u$ and $\mathbf{w}_u = (\mathbf{R}_{\tau_u}^{-1/2})^H \tilde{\mathbf{w}}_u$ respectively, where $\boldsymbol{\nu}_u$ and $\tilde{\mathbf{w}}_u$ are the principal right and left singular vectors, respectively, of the matrix $\mathbf{R}_{\tau_u}^{-1/2} \mathbf{H}_u \mathbf{A}_{u-1}^\perp \in \mathbb{R}^{N_R \times (NN_T - u + 1)}$. The columns of the matrix $\mathbf{A}_{u-1}^\perp \in \mathbb{C}^{MN_T \times (MN_T - u + 1)}$ represent a basis orthonormal to the null-space of the matrix $\mathbf{A}_{u-1} = \left[(\mathbf{w}_1^H \mathbf{H}_1)^T, \dots, (\mathbf{w}_{u-1}^H \mathbf{H}_{u-1})^T \right]^T \in \mathbb{C}^{(u-1) \times MN_T}$. Furthermore, the concatenated fully digital TPC $\mathbf{F}_{\text{opt}} \in \mathbb{C}^{MN_T \times U}$ corresponding to all users is given by $\mathbf{F}_{\text{opt}} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_U]$. Now, the decomposition problem of jointly designing the baseband TPC $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{\text{RF},m} \times U}$ and RF TPC $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{MN_T \times \sum_{m=1}^M N_{\text{RF},m}}$ can be formulated as

$$\begin{aligned} (\mathbf{F}_{\text{BB}}^*, \mathbf{F}_{\text{RF}}^*) &= \arg \min_{\mathbf{F}_{\text{BB}}, \mathbf{F}_{\text{RF}}} \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 \\ \text{s.t.} \quad &|\mathbf{F}_{\text{RF}}(i, j)| = \frac{1}{\sqrt{N_T}}, \end{aligned} \quad (27)$$

where $\mathbf{F}_{\text{BB}} = [\mathbf{f}_{\text{BB},1}, \mathbf{f}_{\text{BB},2}, \dots, \mathbf{f}_{\text{BB},U}]$ denotes the concatenated baseband TPC and $\mathbf{F}_{\text{RF}} = \text{blkdiag}(\mathbf{F}_{\text{RF},1}, \dots, \mathbf{F}_{\text{RF},m})$ represents the block-diagonal matrix of RF TPCs corresponding to all the APs. The constant magnitude constraint in (27) is due to the analog phase shifter based architecture, which results in a non-convex optimization problem. Towards solving this, one can first define a dictionary matrix of the feasible transmit array response vectors as $\mathbf{G}_T = [\mathbf{a}_T(\phi_1), \mathbf{a}_T(\phi_2), \dots, \mathbf{a}_T(\phi_S)] \in \mathbb{C}^{N_T \times S}$, where the set of AoDs $\{\phi_s, \forall 1 \leq s \leq S\}$ spans the angular range $[0, \pi]$, obeying $\cos(\phi_s) = \frac{2}{S}(s-1) - 1$, and the quantity S represents the angular grid size [7]. Note that the elements of the dictionary matrix satisfy the constant magnitude constraint of (27). Hence, the columns of the RF TPC \mathbf{F}_{RF} can be readily chosen from the dictionary matrix \mathbf{G}_T . The equivalent problem for hybrid precoder design in the CFMU mmWave MIMO system can be formulated as

$$\begin{aligned} \arg \min_{\tilde{\mathbf{F}}_{\text{BB}}} &\|\mathbf{F}_{\text{opt}} - \mathbf{G}_T \tilde{\mathbf{F}}_{\text{BB}}\|_{\text{F}} \\ \text{s.t.} \quad &\left\| \text{diag}(\tilde{\mathbf{F}}_{\text{BB}} \tilde{\mathbf{F}}_{\text{BB}}^H) \right\|_0 = \sum_{m=1}^M N_{\text{RF},m}, \end{aligned} \quad (28)$$

where $\tilde{\mathbf{F}}_{\text{BB}} \in \mathbb{C}^{S \times U}$ denotes the intermediate baseband TPC. The constraint in (28) arises due to the fact that there can only be $\sum_{m=1}^M N_{\text{RF},m}$ non-zero rows in the matrix $\tilde{\mathbf{F}}_{\text{BB}}$, since there are only $N_{\text{RF},m}$ RF chains at each AP. This leads to a simultaneous sparse structure in the matrix $\tilde{\mathbf{F}}_{\text{BB}}$. The procedure of obtaining the proposed BL-based hybrid TPC corresponding to the optimization problem (28) is discussed next.

The BL method proposed for our hybrid TPC design assigns the parameterized Gaussian prior $p(\tilde{\mathbf{F}}_{\text{BB}}; \boldsymbol{\Gamma})$ to the

baseband TPC matrix $\tilde{\mathbf{F}}_{\text{BB}}$, where we have

$$\begin{aligned} p(\tilde{\mathbf{F}}_{\text{BB}}; \boldsymbol{\Gamma}) &= \prod_{i=1}^S p(\tilde{\mathbf{F}}_{\text{BB}}(i, :); \gamma_i) \\ &= \prod_{i=1}^S \frac{1}{\pi \gamma_i} \exp\left(-\frac{\|\tilde{\mathbf{F}}_{\text{BB}}(i, :)\|_2^2}{\gamma_i}\right), \end{aligned} \quad (29)$$

and $\boldsymbol{\Gamma} = \text{diag}(\gamma_1, \dots, \gamma_S) \in \mathbb{R}^{S \times S}$ denotes the diagonal matrix of hyper-parameters. One can observe that the i th row of the matrix $\tilde{\mathbf{F}}_{\text{BB}}$ is assigned the hyperparameter γ_i , which enforces the row sparsity, as given in the constraint (28). The MMSE estimator of the matrix $\tilde{\mathbf{F}}_{\text{BB}}$, represented by the matrix $\boldsymbol{\Phi} \in \mathbb{C}^{S \times U}$ and the associated covariance matrix $\boldsymbol{\Pi} \in \mathbb{C}^{S \times S}$ are given as

$$\boldsymbol{\Phi} = \frac{1}{\sigma_e^2} \boldsymbol{\Pi} \mathbf{G}_T^H \mathbf{F}_{\text{opt},n} \quad \text{and} \quad \boldsymbol{\Pi} = \left(\frac{1}{\sigma_e^2} \mathbf{G}_T^H \mathbf{G}_T + \boldsymbol{\Gamma}^{-1} \right)^{-1}, \quad (30)$$

where σ_e^2 denotes the variance of the approximation error. It can be observed that the MMSE estimate $\boldsymbol{\Phi}$ depends on the hyperparameter matrix $\boldsymbol{\Gamma}$. Furthermore, as $\gamma_i \rightarrow 0$, the i th row of the matrix $\tilde{\mathbf{F}}_{\text{BB}}$ denoted by $\tilde{\mathbf{F}}_{\text{BB}}(i, :)$ $\rightarrow 0$. Thus, the estimation of $\tilde{\mathbf{F}}_{\text{BB}}$ is equivalent to the estimation of the associated hyperparameter vector $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_S]^T$. To find the maximal likelihood estimate of $\boldsymbol{\gamma}$, the Bayesian evidence $p(\mathbf{F}_{\text{opt}}; \boldsymbol{\Gamma})$ can be maximized. However, the pertinent optimization problem becomes non-convex, and it is hence difficult to solve. Therefore, the BL technique harnessed for hybrid TPC design in CFMU systems relies on a low-complexity iterative expectation-maximization (EM) framework, which guarantees likelihood maximization in each iteration, and convergence to a local optimum.

To begin with, let $\hat{\boldsymbol{\Gamma}}^{(k-1)}$ denote the estimate of the hyperparameter matrix $\boldsymbol{\Gamma}$ obtained from the $(k-1)$ st iteration. The EM procedure involves two key steps. In the first step, known as the expectation step (E-step), we evaluate the average log-likelihood function $\mathcal{L}(\boldsymbol{\Gamma} | \hat{\boldsymbol{\Gamma}}^{(k-1)})$ of the hyper-parameters as follows

$$\mathcal{L}(\boldsymbol{\Gamma} | \hat{\boldsymbol{\Gamma}}^{(k-1)}) = \mathbb{E}_{\tilde{\mathbf{F}}_{\text{BB}} | \mathbf{F}_{\text{opt}}; \hat{\boldsymbol{\Gamma}}^{(k-1)}} \left\{ \log p(\mathbf{F}_{\text{opt}}, \tilde{\mathbf{F}}_{\text{BB}}; \boldsymbol{\Gamma}) \right\}.$$

The second step, which is known as the maximization step (M-step), maximizes the average log-likelihood computed above with respect to the hyperparameter vector $\boldsymbol{\gamma}$ as

$$\hat{\boldsymbol{\gamma}}^{(k)} = \arg \max_{\boldsymbol{\gamma}} \mathbb{E} \left\{ \log p(\mathbf{F}_{\text{opt}} | \tilde{\mathbf{F}}_{\text{BB}}) + \log p(\tilde{\mathbf{F}}_{\text{BB}}; \boldsymbol{\Gamma}) \right\}. \quad (31)$$

It can be readily observed from (31) that the expression $p(\mathbf{F}_{\text{opt}} | \tilde{\mathbf{F}}_{\text{BB}})$ is independent of the hyperparameter $\boldsymbol{\gamma}$ and

can be disregarded in the successive M-step. Therefore, the equivalent optimization problem is formulated as

$$\begin{aligned}
\hat{\gamma}^{(k)} &= \arg \max_{\gamma} \mathbb{E}_{\tilde{\mathbf{F}}_{\text{BB}} | \mathbf{F}_{\text{opt}}, \hat{\Gamma}^{(k-1)}} \left\{ \log p \left(\tilde{\mathbf{F}}_{\text{BB}}; \Gamma \right) \right\} \\
&\equiv \arg \max_{\gamma} \sum_{i=1}^S \left[-\log(\gamma_i) - \frac{\mathbb{E} \left(\left\| \tilde{\mathbf{F}}_{\text{BB}}(i, :) \right\|_2^2 \right)}{\gamma_i} \right], \\
&\equiv \arg \max_{\gamma} \sum_{i=1}^S -\log(\gamma_i) - \frac{\left\| \Phi^{(k)}(i, :) \right\|_2^2 + U \Pi_{(i,i)}^{(k)}}{\gamma_i},
\end{aligned} \tag{32}$$

where $\Phi^{(k)}$ and $\Pi^{(k)}$ are obtained from (30) by setting $\Gamma = \hat{\Gamma}^{(k-1)}$. The optimal value of $\hat{\gamma}_i^{(k)}$ can be evaluated by equating the gradient of the objective function in (32) with respect to γ to zero, which yields the update equation for each hyperparameter as

$$\hat{\gamma}_i^{(k)} = \frac{1}{U} \left\| \Phi^{(k)}(:, i) \right\|_2^2 + \Pi_{(i,i)}^{(k)}.$$

Upon convergence of the EM procedure, the RF and baseband TPCs are obtained as follows. Let \mathcal{S} denote the set of indices for the $\left(\sum_{m=1}^M N_{\text{RF},m} \right)$ hyper-parameters having the largest magnitudes. The concatenated optimal baseband TPC matrix \mathbf{F}_{BB}^* can be extracted from $\tilde{\mathbf{F}}_{\text{BB}}$ as

$$\mathbf{F}_{\text{BB}}^* = \tilde{\mathbf{F}}_{\text{BB}}(\mathcal{S}, :). \tag{33}$$

Similarly, the optimal RF TPC \mathbf{F}_{RF}^* can be obtained from the dictionary matrix of the transmit array response \mathbf{G}_T as

$$\mathbf{F}_{\text{RF}}^* = \mathbf{G}_T(:, \mathcal{S}). \tag{34}$$

Algorithm 2 presents the step-by-step procedure of the BL-based technique for designing the hybrid TPC. One can follow a similar procedure to obtain the digital baseband and RF RCs from the fully digital design. The complexity analysis of the proposed two-stage distributed hybrid beamformer design is summarized next. Due to lack of space, the detailed derivations for the computational complexities of Algorithms 1 and 2 have been relegated to our technical report [61]. It can be observed that the complexity of Algorithm-1, which designs the fully-digital transmit precoder (TPC), is $\mathcal{O} \left((U-1)^2 M N_T \right)$. Next, the fully-digital TPC is decomposed into its constituent radio frequency (RF) and baseband (BB) precoders using Algorithm-2. This step incurs a complexity of order $\mathcal{O} \left(S^3 \right)$, which can be attributed to the matrix inversion of size $[S \times S]$ in Eq. 29. Since, $S \gg (U-1)^2 M N_T$, the overall complexity of the coordinated hybrid TPC can be closely approximated by $\mathcal{O} \left(S^3 \right)$. Furthermore, the global convergence of the BL algorithm's cost function is obtained at the sparsest solution, which guarantees the sparsest representation of the digital precoder [62]. Further, as a benefit of the EM-algorithm, the proposed BL technique is guaranteed to converge to a fixed point of the log-likelihood function, regardless of its initialization. This leads to the robust performance of the BL

algorithm and makes it ideally suited for hybrid precoder/combiner design in mmWave MIMO networks.

Algorithm 2: BL-based hybrid TPC design

- 1 **Input:** Concatenated optimal digital precoder matrix \mathbf{F}_{opt} , dictionary matrix \mathbf{G}_T , RF chains $\sum_{m=1}^M N_{\text{RF},m}$, variance of approximation error σ_e^2 , stopping parameters ϵ and k_{max} ;
 - 2 **Initialization:** $\hat{\gamma}_i^{(0)} = 1, \forall 1 \leq i \leq S \rightarrow \hat{\Gamma}^{(0)} = \mathbf{I}_S$. Set counter $k = 0$ and $\Gamma^{(-1)} = \mathbf{0}$;
 - 3 **while** $\left(\left\| \hat{\gamma}^{(k)} - \hat{\gamma}^{(k-1)} \right\|^2 > \epsilon \text{ and } k < k_{\text{max}} \right)$ **do**
 - 4 **E-step:** Evaluate the *a posteriori* covariance and mean

$$\Gamma^{(k)} = \left(\frac{1}{\sigma_e^2} \mathbf{G}_T^H \mathbf{G}_T + \left(\hat{\Gamma}^{(k-1)} \right)^{-1} \right)^{-1}$$

$$\tilde{\mathbf{F}}_{\text{BB}}^{(k)} = \frac{1}{\sigma_e^2} \Pi^{(k)} \mathbf{G}_T^H \mathbf{F}_{\text{opt}}$$
 - M-step:** Evaluate the hyperparameter estimate

$$\hat{\gamma}_i^{(k)} = \Pi_{(i,i)}^{(k)} + \frac{1}{U} \sum_{j=1}^U \left| \tilde{\mathbf{F}}_{\text{BB},n}(i, j) \right|^2$$
 - end**
 - 5 **end**
 - 6 **Output:** Obtain \mathbf{F}_{BB}^* and \mathbf{F}_{RF}^* using (33) and (34).
-

VII. SUCCESSIVE MULTIUSER UPLINK HYBRID BEAMFORMING (SCUHF)

Consider now the cooperative CFMU mmWave MIMO uplink system with U users and M cooperative APs. The u th user has N_R TAs and $N_{\text{RF},u}$ RF chains, whereas the m th AP is equipped with N_T RAs and $N_{\text{RF},m}$ RF chains. The symbol transmitted by user u can be expressed as $\mathbf{s}_u = \mathbf{F}_{\text{RF},u} \mathbf{f}_{\text{BB},u} x_u$. The signal received at the m th AP can be expressed as

$$\mathbf{r}_m = \sum_{u=1}^U \mathbf{H}_{u,m} \mathbf{s}_u + \delta_m = \sum_{u=1}^U \mathbf{H}_{u,m} \mathbf{F}_{\text{RF},u} \mathbf{f}_{\text{BB},u} x_u + \delta_m,$$

where $\mathbf{F}_{\text{RF},u} \in \mathbb{C}^{N_R \times N_{\text{RF},u}}$ represents the RF TPC and $\mathbf{f}_{\text{BB},u} \in \mathbb{C}^{N_{\text{RF},u} \times 1}$ denotes the baseband TPC of the user u . Furthermore, $\mathbf{H}_{u,m} \in \mathbb{C}^{N_T \times N_R}$ denotes the channel matrix between the m th AP and the user u . Therefore, the stacked signal $\mathbf{r} = [\mathbf{r}_1^H, \mathbf{r}_2^H, \dots, \mathbf{r}_M^H]^H \in \mathbb{C}^{M N_T \times 1}$ received at all the cooperating APs can be written as

$$\mathbf{r} = \sum_{u=1}^U \mathbf{H}_u \mathbf{F}_{\text{RF},u} \mathbf{f}_{\text{BB},u} x_u + \delta, \tag{35}$$

where $\mathbf{H}_u = [\mathbf{H}_{u,1}^H, \mathbf{H}_{u,2}^H, \dots, \mathbf{H}_{u,M}^H]^H \in \mathbb{C}^{M N_T \times N_R}$ denotes the aggregate channel matrix between all the APs

and the u th user. Also, $\delta \in \mathbb{C}^{MN_T \times 1}$ denotes the complex AWGN with covariance matrix $\mathbb{E}[\delta\delta^H] = \sigma_\delta^2 \mathbf{I}_{MN_T}$. The estimated received symbol corresponding to user u after employing the hybrid RC $\mathbf{w}_u = \mathbf{W}_{\text{RF}} \mathbf{w}_{\text{BB},u}$ at the cooperative APs is given by

$$\begin{aligned} \hat{x}_u &= \mathbf{w}_u^H \mathbf{r} = \mathbf{w}_u^H \mathbf{H}_u \mathbf{s}_u + \sum_{i=1, i \neq u}^U \mathbf{w}_u^H \mathbf{H}_i \mathbf{s}_i + \underbrace{\mathbf{w}_u^H \delta}_{\tilde{\delta}} \\ &= \mathbf{w}_{\text{BB},u}^H \mathbf{W}_{\text{RF}}^H \mathbf{H}_u \mathbf{F}_{\text{RF},u} \mathbf{f}_{\text{BB},u} x_u \\ &+ \sum_{i=1, i \neq u}^U \mathbf{w}_{\text{BB},u}^H \mathbf{W}_{\text{RF}}^H \mathbf{H}_i \mathbf{F}_{\text{RF},i} \mathbf{f}_{\text{BB},i} x_i + \tilde{\delta}. \end{aligned} \quad (36)$$

Let us define the effective channel matrix $\mathbf{H}_{\text{eff},u}$ between the user u and the APs as $\mathbf{H}_{\text{eff},u} = \mathbf{W}_{\text{RF}}^H \mathbf{H}_u \mathbf{F}_{\text{RF},u} \in \mathbb{C}^{MN_{\text{RF}},m \times N_{\text{RF},u}}$. From (36), the effective system model for the estimated symbol \hat{x}_u is given by

$$\hat{x}_u = \mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},u} x_u + \sum_{i=1, i \neq u}^U \mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},i} \mathbf{f}_{\text{BB},i} x_i + \tilde{\delta}. \quad (37)$$

The baseband RC designed for the u th user is given by $\mathbf{w}_{\text{BB},u} = \mathbf{H}_{\text{eff},u} \mathbf{f}_{\text{BB},u}$. It can be observed from (37) that the term $\mathbf{f}_{\text{BB},u}^H (\mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u}) \mathbf{f}_{\text{BB},u}$ represents the signal power of the u th user in the estimate \hat{x}_u and the term $\sum_{i=1, i \neq u}^U |\mathbf{f}_{\text{BB},u}^H \mathbf{H}_{\text{eff},i}^H \mathbf{H}_{\text{eff},i} \mathbf{f}_{\text{BB},i}|^2$ denotes the interference part. In the cooperative uplink scenario, the objective is to design the uplink hybrid TPCs for ensuring that they maximize the SINR of each user in the system. Hence, the pertinent optimization problem is given by

$$\begin{aligned} &\underset{\mathbf{f}_{\text{BB},u}}{\text{maximize}} \quad \mathbf{f}_{\text{BB},u}^H (\mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u}) \mathbf{f}_{\text{BB},u} \\ &\text{s.t.} \quad (\mathbf{f}_{\text{BB},i}^H \mathbf{H}_{\text{eff},i}^H \mathbf{H}_{\text{eff},u}) \mathbf{f}_{\text{BB},u} = 0, \quad i = 1, 2, \dots, u-1, \\ &\quad \|\mathbf{W}_{\text{RF}} \mathbf{w}_{\text{BB},u}\|_2^2 \leq P. \end{aligned} \quad (38)$$

From the first constraint in (38), it can be readily observed that if the baseband TPC $\mathbf{f}_{\text{BB},u}$ is designed so that any user i , $i < u$, does not face interference due to transmission of user u , then the same holds vice versa, i.e., the transmission of user i does not create any interference for user u , $u > i$. Hence, one can deduce that the interference cancellation process is symmetric in nature. Furthermore, it can also be deduced that in the successive CFMU mmWave uplink system there is no interference among the scheduled users, which is in contrast to the successive downlink scenario. The solution to the optimization problem in (38) can be achieved by solving the following problem:

$$\begin{aligned} &\underset{\mathbf{b}_u}{\text{maximize}} \quad \mathbf{b}_u^H \left[(\mathbf{M}_{u-1}^\perp)^H \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \right] \mathbf{b}_u \\ &\text{s.t.} \quad \|\mathbf{b}_u\|_2^2 \leq P, \end{aligned} \quad (39)$$

where $\mathbf{b}_u \in \mathbb{C}^{(N_{\text{RF},u-u+1}) \times 1}$, and $\mathbf{M}_{u-1}^\perp \in \mathbb{C}^{N_{\text{RF},u} \times (N_{\text{RF},u-u+1})}$ denotes a basis for the null space of

the matrix $\mathbf{M}_{u-1} \in \mathbb{C}^{(u-1) \times N_{\text{RF},u}}$, which is given as

$$\mathbf{M}_{u-1} = \begin{bmatrix} \mathbf{f}_{\text{BB},1}^H \mathbf{H}_{\text{eff},1}^H \\ \mathbf{f}_{\text{BB},2}^H \mathbf{H}_{\text{eff},2}^H \\ \vdots \\ \mathbf{f}_{\text{BB},u-1}^H \mathbf{H}_{\text{eff},u-1}^H \end{bmatrix} \mathbf{H}_{\text{eff},u},$$

and can be obtained from the SVD of \mathbf{M}_{u-1} . Therefore, the uplink baseband TPC designed for the u th user is given as $\mathbf{f}_{\text{BB},u} = \sqrt{P} \mathbf{M}_{u-1}^\perp \mathbf{b}_u^*$, where \mathbf{b}_u^* denotes the dominant eigenvector of the matrix $(\mathbf{M}_{u-1}^\perp)^H \mathbf{H}_{\text{eff},u}^H \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp$. Furthermore, the rate R_u of the u th user is given as $R_u = \log_2(1 + \|\mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \mathbf{b}_u^*\|_2^2)$, and the capacity C_{sum} in the uplink scenario is given as $C_{\text{sum}} = \sum_{u=1}^U R_u$. Hence, one can notice that the proposed SCUHBF scheme maximizes each user's SINR, while successfully cancelling the IUI.

VIII. SIMULATION RESULTS

In this section, we provide simulation results for illustrating the performance of the proposed HTBFs in various scenarios. In the cell-free multi-user network considered, the M APs and U users are considered to be uniformly distributed at random within a square coverage area of size $200 \times 200 \text{ m}^2$. The mmWave MIMO channel $\mathbf{H}_{u,m}$ between each user and the AP has been modelled using (2) with $L = 6$ multipath components. The AoA/AoDs of the channels between the APs and users are generated with uniform distribution over $[0 \ 2\pi]$. Furthermore, the multipath gains $\alpha_{u,m}^l$ are assumed to be complex Gaussian distributed as $\mathcal{CN}(0, 1)$. For the BL-based precoder and combiner design, the initial value of all the hyper-parameters is set to $\hat{\gamma}_i^{(0)} = 1, \forall 1 \leq i \leq S$, with the maximum number of EM iterations and the stopping parameter set as $k_{\text{max}} = 50$ and $\epsilon = 10^{-6}$, respectively. We consider 3000 such channel realizations for plotting the capacity R_{sum} versus the total power P_T (dB) available at the APs.

A. Beamforming for broadcast scenarios: HBBF

For a broadcast scenario, a (16×8) elements mmWave MIMO system is considered, where we have $N_T = 16$ TAs at AP and $N_R = 8$ RAs at user, $M = 4$ APs and $U = 8$ users. For HTBF, we assume that the number of RF chains at each AP is equal to the number of users i.e. $N_{\text{RF},m} = U$, and number of RF chains at each user is set equal to the total number of cooperative APs, i.e. $N_{\text{RF},m} = M$. Figure 2a illustrates the performance of the HBBF and HBBF-P schemes described in Section III. It can be observed that the hybrid design of the proposed HBBF schemes closely approaches the performance of an ideal fully-digital beamformer (FDBF), where the number of RF chains is equal to the number of antennas. Furthermore, it can also be observed that the fairness-based HBBF-P scheme associated with per-AP power constraints closely follows the performance of the HBBF, given a pooled power constraint. Thus, the proposed HBBF-aided hybrid TPC design results in the optimal solution, and it is also well-suited for transmission in broadcast CFMU cooperative mmWave

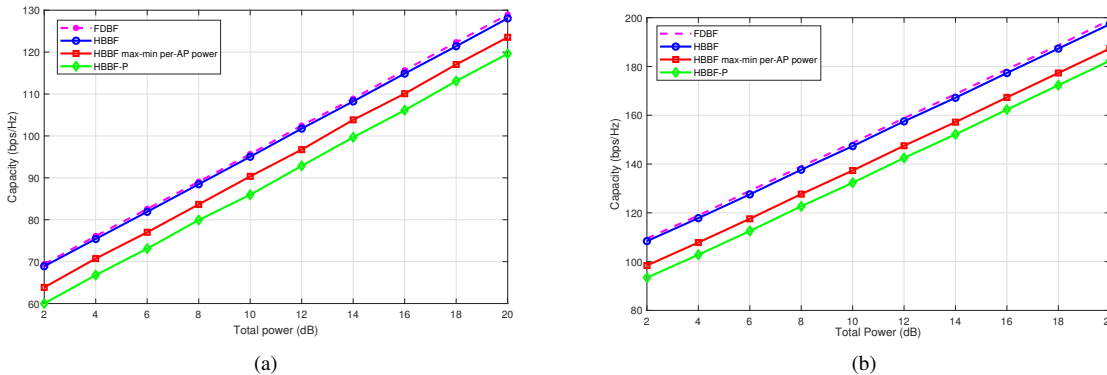


Figure 2: Capacity comparison of various schemes for a broadcast scenario with (a) $M = 4, N_T = 16, N_R = 8, N_{RF,m} = 4,$ and $U = 10$, (b) $M = 4, N_T = 64, N_R = 8, N_{RF,m} = 4,$ and $U = 15$.

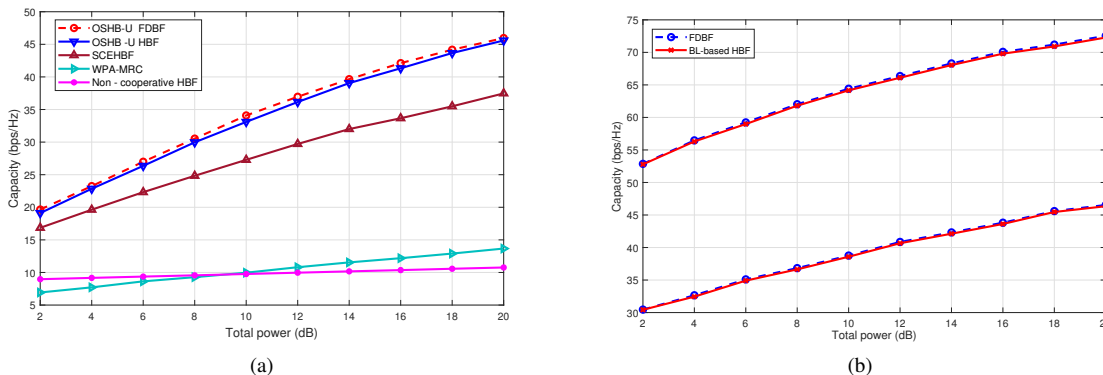


Figure 3: Unicast Scenario: (a) Capacity performance of the OSHB-U with various schemes for $N_T = 16, N_R = 8, N_{RF,m} = 4,$ and $U = 4$, (b) capacity comparison of the fully-digital design and BL-based hybrid beamformer design for $N_T = 32$ and $N_R = 8$ antennas.

MIMO scenarios. Fig. 2b illustrates the performance of the HBBF scheme in a large-scale antenna array regime. It can be observed that the capacity increases upon increasing the number of TAs at the AP, thanks to the array gain of the large antenna arrays. This further establishes the importance of having a large number of antennas in the mmWave regime for improved power efficiency.

B. Beamforming schemes for downlink and uplink unicast scenarios: OSHB-U and SCUHBF

Next, we consider both downlink and uplink CFMU mmWave MIMO unicast scenarios in the face of interuser interference and compare the capacity performance of various schemes. Fig. 3a shows the capacity performance of the OSHB-U in a downlink system along with those of the other schemes for $M = 4$ APs and $U = 4$ users with $N_T = 16, N_R = 8, N_{RF,m} = 4$ and $N_{RF,u} = M$. It can be observed that the OSHB-U scheme performs similarly to the ideal fully-digital beamformer. The proposed OSHB-U scheme achieves an improved performance in terms of capacity as a benefit of the SIC and the maximization of the SINR by the MVDR-based beamformer. For multi-cell cooperative scenarios, one can also employ the sub-optimal successive constrained eigen hybrid beamforming

(SCEHBF) [63] scheme based on MRC, which was originally derived for a sub-6 GHz system and has been suitably extended for mmWave for fair comparison. Its performance is poor, since it disregards the interference inflicted by the previously scheduled users. Furthermore, the simplistic white power allocation (WPA) [63] scheme using MRC results in a significant performance degradation because it fails to cancel the interference.

Fig. 3b compares the capacity of the BL-based HTBF design with that of the ideal fully-digital beamformer. The BL-based HBF design does not require the full knowledge of AoAs/AoDs and their path gains for designing the baseband and RF precoders. It can be observed that the BL-based design attains a performance close to that of the fully-digital design with fewer RF chains. This is attributed to the fact that the mmWave MIMO channel is comprised of fewer multipath components, hence resulting in a low-rank channel. Therefore, the ideal FDBF can be approximated using a linear combination of only a few transmit array response vectors.

Fig. 4a plots the capacity versus the number of users considering both the uniform planar array (UPA) and the uniform linear array (ULA) structures. Interestingly, it can be

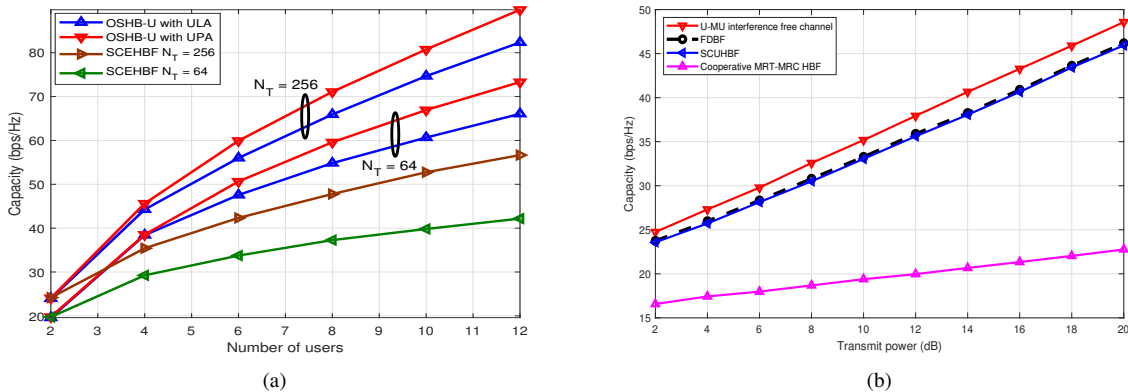


Figure 4: Unicast scenario: (a) Capacity versus number of users with $M = 4$, (b) Uplink scenario: capacity versus total power with $M = 4$, $N_T = 16$, $N_R = 8$, $G = 2$ and $U = 4$.

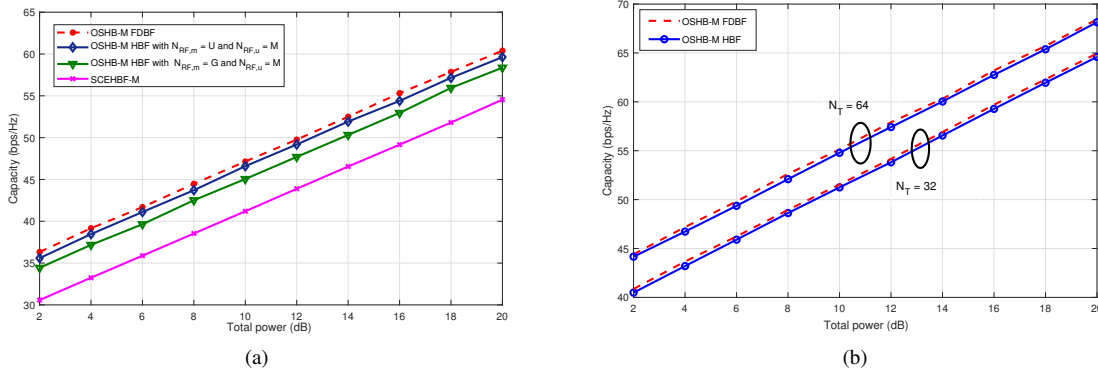


Figure 5: Unicast Scenario: (a) Capacity versus total power for a multicast scenario with $M = 4$, $N_T = 16$, $N_R = 8$, $G = 2$ and $U_g = 4$, (b) capacity versus total power for a multicast scenario $M = 4$, $N_R = 8$, $G = 2$ and $U_g = 4$.

seen that the capacity increases upon increasing the number of AP antennas, which can be attributed to the fact that the CFMU mmWave MIMO system and the proposed design technique is capable of efficiently exploiting the resultant increased array gain. Furthermore, one can also observe that the proposed OSHBF-U algorithm outperforms the SCEHBF, when serving different number of users. To be specific, as the number of served users increases, the gap between the proposed and benchmark algorithms progressively increases, which further confirms the stability and effectiveness of our proposed algorithm. Fig. 4b plots the capacity versus transmit power for the uplink scenario. For this analysis, $U = 4$ active users are equipped with $N_R = 8$ TAs, and assumed to be distributed in $M = 4$ cooperating cells with each AP having $N_T = 16$ RAs. It can be observed that the proposed SCUHBF scheme performs close to the capacity of U interference-free point-to-point mmWave MIMO systems having dimensions of $\sum_{m=1}^M N_{RF,m} \times N_{RF,u}$, and it also closely matches the performance of the ideal FD design. Furthermore, it is also observed that the capacity achieved by the proposed SCUHBF scheme for a cooperative uplink scenario is much higher than that of employing maximum ratio transmit (MRT) beamforming at the users and MRC at

the APs.

C. Beamforming for multicast scenarios: OSHB-M

Fig. 5a compares the capacity performance of the OSHB-M scheme in the face of inter-group interference for different values of the number of RF chains, $M = 4$ APs and $U_g = 4$ users having $N_T = 16$, $N_R = 8$ antennas, and $G = 2$ user groups. It can be observed that the performance of the proposed OSHB-M scheme closely follows that of the ideal fully-digital TBF design, but with a much lower number of RF chains at the APs. Next, Fig. 5b demonstrates the large-scale antenna benefits of the mmWave regime. It can be observed that the performance of the OSHB-M schemes improve significantly upon increasing the number of antennas at the APs.

IX. CONCLUSION

Optimal hybrid beamformers for broadcast, unicast and multicast CFMU mmWave MIMO systems are designed in the face of interuser/intergroup interference. Initially, for a broadcast scenario, the optimal beamforming solutions were derived by considering a total power constraint as well as per-AP power constraints. Subsequently, OSHB-based

techniques relying on the optimal MVDR principle were proposed for unicast and multicast scenarios, which efficiently cancel the resultant interuser and intergroup interference. Next, BL-based joint hybrid transceiver was also design which did not require the knowledge of the AoAs/AoDs and the gains of the various multipath components of the mmWave MIMO channels. Finally, we also derived the SCUHBF-based cooperative uplink HBF, which maximizes the SINR, hence achieving high data rates. Our simulation results demonstrated the improved spectral efficiency as well as interference cancellation capability of the proposed HBBF, OSHB and SCUHBF techniques. The proposed hybrid designs have a low-complexity, and can simultaneously support an increased number of users in comparison to the existing methods, which makes them eminently suitable for practical mmWave CFMU MIMO systems.

APPENDIX A MVDR BEAMFORMER

Considering the system model of (17), the signal \tilde{y}_u received at user u after RF combining can be written as

$$\tilde{y}_u = \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u x_u + \boldsymbol{\tau}_u,$$

where $\boldsymbol{\tau}_u$ represents the interference-plus-noise component associated with the covariance matrix $\mathbf{C}_{\boldsymbol{\tau}_u}$. After baseband combining at user u , the signal estimate \hat{x}_u is given as

$$\hat{x}_u = \mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u x_u + \mathbf{w}_{\text{BB},u}^H \boldsymbol{\tau}_u.$$

Therefore, the SINR at user u is given by

$$\text{SINR}_u = \frac{|\mathbf{w}_{\text{BB},u}^H \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u|^2}{\mathbf{w}_{\text{BB},u}^H \mathbf{C}_{\boldsymbol{\tau}_u} \mathbf{w}_{\text{BB},u}}.$$

The MVDR beamformer is the optimal beamformer that maximizes the SINR [59], and it is given as $\mathbf{w}_{\text{BB},u} = \mathbf{C}_{\boldsymbol{\tau}_u}^{-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u$. Furthermore, the maximum value of the SINR is determined as

$$\begin{aligned} \text{SINR}_{\text{max},u} &= \frac{|\boldsymbol{\nu}_u^H \mathbf{M}_{u-1}^{\perp H} \mathbf{H}_{\text{eff},u}^H \mathbf{C}_{\boldsymbol{\tau}_u}^{-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u|^2}{\boldsymbol{\nu}_u^H \mathbf{M}_{u-1}^{\perp H} \mathbf{H}_{\text{eff},u}^H \mathbf{C}_{\boldsymbol{\tau}_u}^{-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u} \\ &= \boldsymbol{\nu}_u^H \mathbf{M}_{k-1}^{\perp H} \mathbf{H}_{\text{eff},u}^H \mathbf{C}_{\boldsymbol{\tau}_u}^{-1} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp \boldsymbol{\nu}_u. \end{aligned}$$

Observe from the above expression that the optimal vector $\boldsymbol{\nu}_u$ that maximizes the $\text{SINR}_{\text{max},u}$ is the principal right singular vector of $\mathbf{C}_{\boldsymbol{\tau}_u}^{-1/2} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp$. Therefore, the optimal MVDR receive beamformer $\mathbf{w}_{\text{BB},u}$ is given by $\mathbf{w}_{\text{BB},u} = (\mathbf{C}_{\boldsymbol{\tau}_u}^{-1/2})^H \tilde{\mathbf{w}}_{\text{BB},u}$, where $\tilde{\mathbf{w}}_{\text{BB},u}$ is the dominant left singular vector of the matrix $\mathbf{C}_{\boldsymbol{\tau}_u}^{-1/2} \mathbf{H}_{\text{eff},u} \mathbf{M}_{u-1}^\perp$. This completes the proof.

REFERENCES

- [1] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2017.
- [2] W. Huang, Y. Huang, R. Zhao, S. He, and L. Yang, "Wideband millimeter wave communication: Single carrier based hybrid precoding with sparse optimization," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9696–9710, 2018.
- [3] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3437–3458, 2015.
- [4] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1394–1404, 2019.
- [5] Y. Chen, D. Chen, T. Jiang, and L. Hanzo, "Millimeter-wave massive MIMO systems relying on generalized sub-array-connected hybrid precoding," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8940–8950, 2019.
- [6] B. P. Sahoo, C.-C. Chou, C.-W. Weng, and H.-Y. Wei, "Enabling millimeter-wave 5G networks for massive IoT applications: A closer look at the issues impacting millimeter-waves in consumer devices under the 5G framework," *IEEE Consumer Electronics Magazine*, vol. 8, no. 1, pp. 49–54, 2018.
- [7] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.
- [8] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, 2016.
- [9] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE journal on selected areas in communications*, vol. 28, no. 9, pp. 1380–1408, 2010.
- [10] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*. IEEE, 2007, pp. 83–87.
- [11] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, 2011.
- [12] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [13] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.
- [14] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 611–652, 2021.
- [15] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [16] O. El 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.
- [17] Y. Wang and W. Zou, "Low complexity hybrid precoder design for millimeter wave MIMO systems," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1259–1262, 2019.
- [18] J. Noh, T. Kim, J.-Y. Seol, and C. Lee, "Zero-forcing based hybrid beamforming for multi-user millimeter wave systems," *IET Communications*, vol. 10, no. 18, pp. 2670–2677, 2016.
- [19] 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.
- [20] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin, "Performance impact of LoS and NLoS transmissions in dense cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2365–2380, 2015.
- [21] X. Zhang and J. G. Andrews, "Downlink cellular network analysis with multi-slope path loss models," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1881–1894, 2015.
- [22] M. Hong, R. Sun, H. Baligh, and Z.-Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 226–240, 2013.

- [23] R. Zhang and L. Hanzo, "Cooperative downlink multicell preprocessing relying on reduced-rate back-haul data exchange," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 539–545, 2010.
- [24] J. Li, E. Björnson, T. Svensson, T. Eriksson, and M. Debbah, "Joint precoding and load balancing optimization for energy-efficient heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5810–5822, 2015.
- [25] Y. Cheng, M. Pesavento, and A. Philipp, "Joint network optimization and downlink beamforming for CoMP transmissions using mixed integer conic programming," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3972–3987, 2013.
- [26] H. J. Yang, W.-Y. Shin, B. C. Jung, C. Suh, and A. Paulraj, "Opportunistic downlink interference alignment for multi-cell MIMO networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1533–1548, 2017.
- [27] A. Michaloliakos, W.-C. Ao, and K. Psounis, "Joint user-beam selection for hybrid beamforming in asynchronously coordinated multi-cell networks," in *2016 Information Theory and Applications Workshop (ITA)*. IEEE, 2016, pp. 1–10.
- [28] G. Zhu, K. Huang, V. K. Lau, B. Xia, X. Li, and S. Zhang, "Hybrid beamforming via the kronecker decomposition for the millimeter-wave massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2097–2114, 2017.
- [29] S. Sun, T. S. Rappaport, M. Shafi, and H. Tataria, "Analytical framework of hybrid beamforming in multi-cell millimeter-wave systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7528–7543, 2018.
- [30] L. Bai, T. Li, Q. Yu, J. Choi, and W. Zhang, "Cooperative multiuser beamforming in mmwave distributed antenna systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12394–12397, 2018.
- [31] C. Fang, B. Makki, J. Li, and T. Svensson, "Hybrid precoding in cooperative millimeter wave networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 5373–5388, 2021.
- [32] J. Kim, S.-H. Park, O. Simeone, I. Lee, and S. S. Shitz, "Joint design of fronthauling and hybrid beamforming for downlink c-ran systems," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4423–4434, 2019.
- [33] Q. Hou, S. He, Y. Huang, H. Wang, and L. Yang, "Joint user scheduling and hybrid precoding design for mimo c-ran," in *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*. IEEE, 2017, pp. 1–6.
- [34] G. Femenias and F. Riera-Palou, "Cell-free millimeter-wave massive MIMO systems with limited fronthaul capacity," *IEEE Access*, vol. 7, pp. 44 596–44 612, 2019.
- [35] C. Feng, W. Shen, J. An, and L. Hanzo, "Weighted sum rate maximization of the mmwave cell-free MIMO downlink relying on hybrid precoding," *IEEE Transactions on Wireless Communications*, 2021.
- [36] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmwave massive MIMO through deep learning," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10325–10329, 2019.
- [37] D. D. Souza, M. M. Freitas, G. S. Borges, A. M. Cavalcante, D. B. da Costa, and J. C. Costa, "Effective channel blind estimation in cell-free massive MIMO networks," *IEEE Wireless Communications Letters*, vol. 11, no. 3, pp. 468–472, 2021.
- [38] M. Guo and M. C. Gursoy, "Joint activity detection and channel estimation in cell-free massive MIMO networks with massive connectivity," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 317–331, 2021.
- [39] H. Song, T. Goldstein, X. You, C. Zhang, O. Tirkkonen, and C. Studer, "Joint channel estimation and data detection in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, 2021.
- [40] X. Wang, A. Ashikhmin, Z. Dong, and C. Zhai, "Two-stage channel estimation approach for cell-free IoT with massive random access," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1428–1440, 2022.
- [41] A. C. Gilbert and J. A. Tropp, "Applications of sparse approximation in communications," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005*. IEEE, 2005, pp. 1000–1004.
- [42] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Transactions on Communications*, vol. 60, no. 12, pp. 3705–3716, 2012.
- [43] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using lse and sparse message passing for mmwave MIMO systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 245–259, 2018.
- [44] S. Srivastava, C. S. K. Patro, A. K. Jagannatham, and L. Hanzo, "Sparse, group-sparse, and online bayesian learning aided channel estimation for doubly-selective mmwave hybrid mimo ofdm systems," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5843–5858, 2021.
- [45] G. Dartmann, X. Gong, and G. Ascheid, "Low complexity cooperative multicast beamforming in heterogeneous networks," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*. IEEE, 2013, pp. 1–6.
- [46] C. Guo, Y. Cui, D. W. K. Ng, and Z. Liu, "Multi-quality multicast beamforming with scalable video coding," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5662–5677, 2018.
- [47] M. Dong and Q. Wang, "Multi-group multicast beamforming: Optimal structure and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3738–3753, 2020.
- [48] J. Choi, "Iterative methods for physical-layer multicast beamforming," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5185–5196, 2015.
- [49] M. Dai and B. Clerckx, "Hybrid precoding for physical layer multicasting," *IEEE Communications Letters*, vol. 20, no. 2, pp. 228–231, 2015.
- [50] J. Wang, H. Xu, B. Zhu, L. Fan, and A. Zhou, "Hybrid beamforming design for mmwave joint unicast and multicast transmission," *IEEE Communications Letters*, vol. 22, no. 10, pp. 2012–2015, 2018.
- [51] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*. Simon & Schuster, Inc., 1992.
- [52] H. v. Trees *et al.*, "Optimum array processing," 2002.
- [53] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2015.
- [54] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 2.1," 2014.
- [55] D. P. Palomar and Y. C. Eldar, *Convex optimization in signal processing and communications*. Cambridge University Press, 2010.
- [56] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.
- [57] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.
- [58] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, 2004.
- [59] H. L. Van Trees, "Modulation theory, optimum array processing (part iv)," 2002.
- [60] M. Sadeghi, E. Björnson, E. G. Larsson, C. Yuen, and T. Marzetta, "Joint unicast and multi-group multicast transmission in massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6375–6388, 2018.
- [61] M. Jafri, S. Srivastava, V. Naveen, A. K. Jagannatham, and L. Hanzo, "Technical report: Cooperative hybrid transmit beamforming in cell-free mmwave MIMO networks," *IIT Kanpur, Tech. Rep.*, 2022 [Online]. [Online]. Available: https://www.iitk.ac.in/mwn/documents/MWNLAB_TR_CFMU.pdf
- [62] D. P. Wipf and B. D. Rao, "Sparse bayesian learning for basis selection," *IEEE Transactions on Signal processing*, vol. 52, no. 8, pp. 2153–2164, 2004.
- [63] N. K. D. Venkateswara, N. Tandon, and A. K. Jagannatham, "Cooperative multi-cell beamforming for MIMO unicast/multicast broadband H.264 scalable video networks," in *2013 IEEE International Conference on Multimedia and Expo (ICME)*, 2013, pp. 1–6.