Robust Distributed Hybrid Beamforming in Coordinated Multi-user Multi-cell mmWave MIMO Systems Relying on Imperfect CSI

Meesam Jafri, *Student Member, IEEE,* Amrit Anand, Suraj Srivastava, *Member, IEEE,* Aditya K. Jagannatham, *Member, IEEE,* and Lajos Hanzo, *Life Fellow, IEEE*

***Abstract*—Novel hybrid beamformer designs are conceived for a multi-user multi-cell (MUMC) mmWave system relying on base station (BS) coordination and total transmit power minimization subject to realistic signal-to-interference-plus-noise ratio (SINR) constraints at each mobile station (MS). Initially, a semidefinite relaxation (SDR)-based approach is developed for a centralized MUMC system to determine the fully digital beamformer having perfect CSI. Subsequently, a Bayesian learning (BL) technique is harnessed for decomposing the fully-digital (FD) solution into its analog and digital components for constructing a hybrid transceiver. Next, an alternating direction method of multipliers (ADMM) based distributed hybrid beamformer is designed for the same system, which requires only local CSI and limited information exchange among the BSs, thus avoiding the excessive signalling overheads required by the centralized approach. Then we further extend both the centralized and the above distributed hybrid designs to construct robust beamformers that minimize the worst-case transmit power with imperfect CSI. Our robust beamforming techniques leverage the S-lemma, which is eminently suitable for the infinitely many constraints arising from the associated CSI uncertainty. Finally, our simulation results demonstrate the improved performance of the proposed centralized and distributed methods over the system having no coordination.**

***Index Terms*—mmWave, convex optimization, multi cell, coordinated beamforming, channel state information(CSI) uncertainty, alternating direction method of multipliers (ADMM), semidefinite relaxation (SDR).**

I. Introduction

Next-generation wireless networks are designed for meeting the ever-increasing demand for high data rates. Additionally, they promise to deliver massive connectivity and, ultra-low latency communication [1]–[3]. The maturing mmWave wireless technology can play a significant role in realizing these ambitious goals due to its ability to exploit the abundance of bandwidth in the 30-300 GHz frequency range. As a result, mmWave communication has gained wide acclaim as a promising next generation technique [4]–[6]. The short wavelength of mmWave signals renders it suitable for compact multiple-input-multiple-output (MIMO) solutions in order to compensate for the high propagation losses [7]. However, conventional transceiver designs require an individual RF chain for each antenna, which can pose a significant implementational challenge in the mmWave regime due to the large number of antennas coupled with the soaring power consumption of the high-bandwidth analog-to-digital convertors [8]. To overcome this obstacle, novel hybrid beamforming architectures were conceived in [9], [10] relying on a few RF chains. The overall MIMO signal processing is divided into two stages, namely, the analog processing that operates in the RF domain to provide directional gain, and the digital baseband processing that provides spatial multiplexing [7].

Since mmWave signals suffer from high blockage and propagation losses, the base stations (BSs) are typically densely deployed for creating small cells [11]. Therefore, coordination among the BSs is an excellent option for improving the spectral efficiency, by mitigating the inter-cell interference (ICI) [12], [13]. In a typical coordinated multi-cell system, several BSs are connected to a central unit (CU) via high speed backhaul links in order to jointly optimize their transmission. However, full coordination among the BSs, also termed centralized beamforming, requires the channel state information (CSI) between each user and all the BSs to be available at the CU, which is challenging in practice [14], since it necessitates an extremely high signalling overhead on the backhaul, which increases exponentially with the number of coordinated cells [15]–[17]. Distributed beamforming, which requires only local CSI associated with limited information exchange, is an excellent solution for overcoming this challenge [18]. The key principle in the distributed beamformer design is to decouple the original optimization problem into smaller sub-problems, each of which is computationally simpler that the original problem, and requires only local CSI. Finally, one must also bear in mind, that it is challenging to obtain perfect CSI between the user and BSs in practice due to the finite training sequence length and limited feedback. The resultant practical CSI error imposes significant performance degradation on the overall system if not specifically accounted for the beamformer design process. Thus, it is essential to design robust coordinated hybrid beamformer (HBF) techniques that take the CSI uncertainty into consideration in both centralized and distributed beamforming. Hence, the review of the existing solutions is presented below.

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

A. Literature Review

Hybrid precoder design constitutes a challenging problem in mmWave MIMO systems, which has attracted substantial research attention. A single-user mmWave MIMO based transmit precoder (TPC) was designed by the authors of [19]–[21]. To elaborate, in [19], orthogonal matching pursuit (OMP) based hybrid beamforming solutions were presented for single user scenarios that intelligently exploited the sparse scattering nature of the mmWave channel, which substantially simplified this challenging problem. Furthermore, a phase pursuit algorithm was proposed in [20] wherein the hybrid beamforming problem was decomposed into several sub-problems in order to obtain the optimal solution. In [21], the authors proposed the joint hybrid transceiver based on a codebook for multiple stream communication in mmWave MIMO systems. The key limitation of all these papers is that they consider only a single user scenario. However, in a multi-user (MU) system, the multi-user interference (MUI) significantly degrades the overall system performance. Hence, the corresponding hybrid beamformer design minimizing the MUI is an additional challenge in MU systems.

In this context, the authors of [22] presented a novel technique based on the simultaneous orthogonal matching pursuit (SOMP) algorithm to obtain hybrid beamforming solutions for a MU mmWave MIMO system, which required low training and feedback overheads. Further, a evolved technique has been studied in [23], where the minimum-mean-squared-error (MMSE) principle is employed in addition to the OMP approach for attaining further capacity improvements in low-SNR scenarios. In [24], the authors minimized the mean-squared error (MSE) of the received signal by optimizing the weights of the hybrid TPCs in the presence of MUI. The authors of [25] consider multi-user mmWave system where each user is equipped with a single antenna and design the analog RF TPCs based on the transmit beam directions, where the digital processing is applied to the equivalent baseband channels.

However, the above contributions consider only a single-cell scenario, thus failing to leverage the significant gains of BS coordination in MUMC systems. In [26]–[29], the authors studied MUMC systems with the objective of designing fully-digital beamformers assuming the availability of perfect CSI. However, it is challenging to obtain perfect CSI at BSs in practice due to the limited pilot overhead. Various authors have focused their attention on robust beamformer design by taking into consideration the CSI uncertainty [30]–[33]. Chen *et al.* [30] developed a robust power control scheme under imperfect CSI by relying on the multi-objective optimization of the quality of service, optimal SINR target tracking, and minimum average power consumption in a MUMC system. The authors of [31] also considered the CSI uncertainty and successfully developed an energy efficient robust transceiver by minimizing the interference leakage to address their power control problem. The authors of [32] designed downlink beamformers for MUMC systems relying on imperfect CSI by considering different optimization criteria, viz., minimization of the total downlink transmit power subject to specific QoS constraints at each user and the maximization of the worst-case effective SINR subject to BS power constraints. In [33], the authors developed a robust multicast beamformer for multiple groups of users in an MUMC system, under the assumption of realistic channel uncertainty and achieved the target SINR of all users. Then Ararat *et al.* [34] proposed a two-stage robust distributed beamformer for MUMC systems. In the first stage, each BS obtains a local version of its coupling variables from a global variable, followed by decomposing the problem into independently solvable subproblems and a master problem. In the second stage, a gradient projection-based iterative algorithm is employed to solve the master problem. The authors of [35] developed a low-complexity algorithm for distributed robust beamformer design and power loading in a multi-cell downlink environment.

Above works mainly focused on the design of fully digital beamforming schemes, which would require an excessive number of RF chains in MUMC mmWave systems. The BS cooperation of mmWave MIMO networks was studied in [36]–[39]. Michaloliakos *et al.* in [36] presented the joint design of analog TPCs relying on the predefined beam patterns for maximizing the data rate of all the users, where each user is served by only a single BS. A novel analog beamformer was developed in [37], which is based on decomposing the beamformer as well as the data and interference path-vectors into the corresponding Kronecker products of unit-modulus phase-shift vectors. In [38], the authors employed signal-to-leakage-plus-noise-ratio (SLNR) based interference coordination in a multi-cell system for designing regularized zero-forcing hybrid TPCs for interference mitigation. The authors of [39], proposed a measurement-campaign-based framework for efficiently mitigating the outage probability, and enhance the spectral efficiency by BS cooperation. Bai *et al.* in [40] proposed a cooperative multi-user beamforming method for energy efficient transmission, which exploits the specific propagation characteristics of the mmWave MIMO channel. Furthermore, the authors of [41] proposed Interference Subspace Alignment (ISA) based centralized fully-digital and hybrid TPC schemes for their MUMC systems, while considering an interfering broadcast channel. However, the proposed scheme needs a large number of RAs at each user for achieving ISA with the aid of hybrid precoding. Note however that all the above papers only present techniques for centralized beamformer design, where the channel knowledge of all the users in the system is assumed to be available both at each BS and at the CU for joint computation of all the TPC weights. As a further development, Daniel *et al.* [42] computed the hybrid TPC in a distributed manner, where analog precoding is performed at the BSs and the baseband precoder is harnessed at the CU for joint processing. As a further advance, the authors of [43] proposed distributed coordinated non-linear hybrid TPC schemes for MUMC massive MIMO systems, where an

approximate block diagonalization (BD) technique is utilized at the RF-stage and a minimum mean square error vector perturbation technique is employed at the BB-stage. However, power optimization cannot be realize in the proposed framework under interference coordination between cells since the BSs do not exchange signals on the backhaul.

This paper conceives a technique for both centralized as well as for distributed beamformer design, while also considering realistic CSI uncertainty that is inevitable in practical MUMC systems. Our proposed hybrid beamformer designs are power efficient as well as requires low information exchange among the BSs, thus results in low backhaul signaling. To the best of our knowledge, none of the papers in the existing literature have comprehensively addressed the above challenges. Our new contributions are as follows.

B. Novel Contributions

We design efficient TPCs for MUMC mmWave MIMO systems, which minimize the total transmit power at each BS, while ensuring that the signal-to-interference-noise ratio (SINR) at each MS exceeds a desired minimum threshold.

- We commence with the centralized TPC design of an MUMC system, where the weight optimization problem is shown to be non-convex, which is difficult to solve. Therefore, a novel two-step process is devised for solving this problem. In the first step, semidefinite relaxation (SDR) [44], [45] is used for converting this to a tractable convex optimization problem for the design of the digital beamformers. In the second step, a Bayesian learning (BL) [46] based method is presented for decomposing the fully digital TPC into its RF and baseband components for implementation in a hybrid mmWave MIMO transceiver [47].
- Next, to avoid the high signalling overhead incurred by the above centralized technique, a distributed beamformer is developed based on the alternating direction method of multipliers (ADMM), which invokes the principles of dual decomposition and augmented Lagrangian cost function [48]. The key advantage of the proposed ADMM-based distributed beamformer design is that it requires only local CSI and limited information exchange between the BSs, thus making it well-suited for practical implementation.
- Next, we extend our analysis to include realistic CSI uncertainty, and derive the robust centralized beamformer that minimizes the multi-cell transmit power for the worst-case channel error. A challenging aspect of this robust beamformer design optimization problem, in addition to its non-convexity, is that there are infinitely many constraints due to the CSI uncertainty. The S-lemma [49] is successfully invoked in this scenario for converting the optimization problem to one having finite number of convex constraints, thus rendering it tractable.

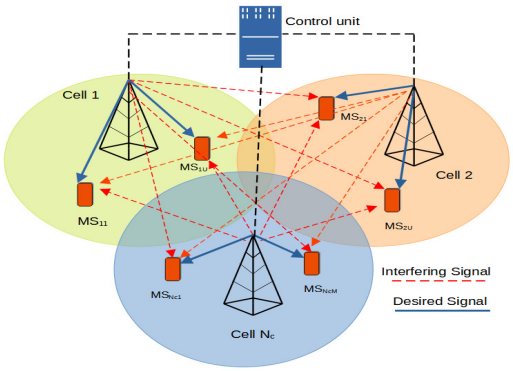


Figure 1: Coordinated DL beamforming in MUMC mmWave MIMO systems

- The above framework is once again extended to a robust distributed beamformer design relying on imperfect CSI using the ADMM technique.
- Finally, our simulation results demonstrate the feasibility of the distributed solution along with the improved power and spectral-efficiency of the coordinated beamforming solution in contrast to a scenario having no coordination. Our new contributions are boldly and explicitly contrasted to the literature in table I.

C. Organization of the paper

The rest of the paper is organized as follows. The MUMC system model together with the mmWave MIMO channel model is described in Section II, followed by the centralized beamformer design using SDR and BL-based hybrid TPC decomposition. In Section III, an ADMM-based distributed beamformer design is conceived for MUMC mmWave systems. Subsequently, in Sections IV and V, robust beamformer designs are derived for centralized and distributed MUMC scenarios, respectively, while relying on imperfect CSI. Finally, Section VI presents our simulation results for characterizing the performance of our proposed techniques, followed by our conclusions in Section VII, finally the Appendix describes beamformer design for a scenario having no coordination.

Notations: Vectors are denoted by boldfaced lowercase alphabets such as $\mathbf{a}$ and matrices are denoted by boldfaced uppercase alphabets such as $\mathbf{A}$, $\mathbb{C}^n$ and $\mathbb{R}^n$ denote the set of n-dimensional complex and real vectors, respectively, $\mathbb{R}^n_+$ represents the set of n-dimensional non-negative vectors, $\mathbf{I}_n$ and $\mathbf{0}$ denote the n-dimensional identity matrix and the all zero matrix/vector of suitable dimension, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$ and $\mathbf{A}^\dagger$ denote the transpose, Hermitian and pseudo inverse of a matrix $\mathbf{A}$. The notation $\mathbf{A} \succeq \mathbf{0}$ and $\mathbf{A} \succ \mathbf{0}$ denote that the matrix $\mathbf{A}$ is positive semidefinite and positive definite, respectively. The functions $\text{rank}(\mathbf{A})$ and $\text{Tr}(\mathbf{A})$ denote the rank and trace of a matrix $\mathbf{A}$. The quantity $\|\mathbf{a}\|$ denotes the Euclidean norm of a vector $\mathbf{a}$.

Table I: CONTRASTING OUR CONTRIBUTION TO THE EXISTING LITERATURE

	[19]	[20]	[22]	[25]	[27]	[29]	[31]	[36]	[39]	[40]	Our
mmWave communication	✓	✓	✓	✓				✓	✓	✓	✓
Hybrid architecture	✓	✓	✓	✓				✓	✓	✓	✓
Multi-cell					✓	✓	✓	✓	✓		✓
Multi-user			✓	✓	✓	✓	✓	✓	✓	✓	✓
Coordinated beamforming					✓			✓	✓	✓	✓
Centralized design					✓			✓	✓		✓
Distributed design										✓	✓
Robust design							✓				✓

II. MUMC mmWave MIMO System Model

Let us Consider the MUMC downlink (DL) system in N_c cells, wherein each cell has a single BS with M transmit antennas (TA) that serves U single-antenna users. This MUMC mmWave MIMO system is shown in Fig. 1, where the BS performs coordinated beamforming to serve multiple users. Again the BSs are connected to the CU via high-speed backhaul links. In contrast to a centralized scenario, where the CU can be assumed to have perfect knowledge of the CSI of all the users in the system, in our distributed system considered therein, the BSs are assumed to have only local CSI, i.e., the CSI of the users in its own cell, along with limited information exchange capability with the neighbouring BSs. This assumption renders the system amenable for practical implementation, but challenging to design.

A more precise analytical model for this system can be developed as follows. The signal transmitted by the nth BS, denoted as BS_n, can be described as

$$\mathbf{x}_n = \sum_{u=1}^{U} \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}s_{nu}, \qquad \forall n \in \mathcal{N}_c, \tag{1}$$

where $\mathbf{F}_{\mathrm{RF},n} \in \mathbb{C}^{M\times N_{\mathrm{RF},n}}$ denotes the RF TPC and $\mathbf{f}_{\mathrm{BB},nu} \in \mathbb{C}^{N_{\mathrm{RF},n}\times 1}$ represents the baseband TPC employed by BS_n for the transmission to MS u in the nth cell, represented as MS_{nu}. The quantity $s_{nu} \in \mathbb{C}$ is the baseband signal intended for MS_{nu} that has an average power of unity, i.e. $\mathbb{E}\{|s_{nu}|^2\} = 1$. Moreover, $N_{\mathrm{RF},n}$ represents the number of RF chains at BS_n obeying $1 \leq N_{\mathrm{RF},n} << M$. Since we consider single-stream transmission for each user, the number of RF chains $N_{\mathrm{RF},n}$ can be set to the number of users U in each cell.

Let $\mathbf{h}_{mnu} \in \mathbb{C}^M$ denote the mmWave downlink channel spanning from BS_m to MS_{nu}. The signal $y_{nu} \in \mathbb{C}$ received at MS_{nu} can be formulated as

$$\begin{aligned} y_{nu} &= \sum_{m=1}^{N_c} \mathbf{h}_{mnu}^H \mathbf{x}_m + \delta_{nu} \\ &= \mathbf{h}_{nnu}^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}s_{nu} + \sum_{i\neq u}^{U} \mathbf{h}_{nnu}^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},ni}s_{ni} \\ &+ \sum_{m\neq n}^{N_c}\sum_{i=1}^{U} \mathbf{h}_{mnu}^H \mathbf{F}_{\mathrm{RF},m}\mathbf{f}_{\mathrm{BB},mi}s_{mi} + \delta_{nu}, \end{aligned} \tag{2}$$

where δ_{nu} denotes the complex additive white Gaussian noise with zero mean and variance σ_{nu}^2. In (2), the first term represents the signal intended for MS_{nu}, the second term denotes the intra-cell interference, whereas the third term represents the inter-cell interference (ICI). Thus, the instantaneous SINR at MS_{nu} is given by

$$\mathrm{SINR}_{nu} = \frac{\left|\mathbf{h}_{nnu}^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\right|^2}{\sum\limits_{i\neq u}^{U} \left|\mathbf{h}_{nnu}^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},ni}\right|^2 + \sum\limits_{m\neq n}^{N_c}\sum\limits_{i=1}^{U} \left|\mathbf{h}_{mnu}^H \mathbf{F}_{\mathrm{RF},m}\mathbf{f}_{\mathrm{BB},mi}\right|^2 + \sigma_{nu}^2}. \tag{3}$$

Note that in order to satisfy the QoS requirement of MS_{nu}, one has to ensure that the SINR_{nu} exceeds the desired target SINR γ_{nu} i.e., $\mathrm{SINR}_{nu} \geq \gamma_{nu}$. The mmWave channel model for this system is described next.

Employing the popular geometric channel model described in [19], [8] for the mmWave regime, the channel between the BS_m and MS_{nu} can be modeled as

$$\mathbf{h}_{mnu}^H = \sqrt{\frac{M}{L}} \sum_{l=1}^{L} \alpha_{l,mnu}\mathbf{a}_T^H(\phi_l^T),$$

where L denotes the number of multipath components, and α_l represents the complex gain of the lth multipath component. The quantity $\mathbf{a}_T(\phi_l^T) \in \mathbb{C}^{M\times 1}$ denotes the array response corresponding to the angle of departure (AoD) $\phi_l^T \in [0, 2\pi]$ of the lth multipath component, which can be expressed for a uniformly spaced linear array (ULA) as $\mathbf{a}_T(\phi_l^T) = \frac{1}{\sqrt{M}}\left[1,\ e^{j\frac{2\pi}{\lambda_d}d\sin(\phi_l^T)},\ \ldots,\ e^{j\frac{2\pi}{\lambda_d}(M-1)d\sin(\phi_l^t)}\right]^T$, where λ_d and d represent the carrier wavelength and the inter-element spacing, respectively. The procedure of centralized mmWave MUMC beamformer design is detailed below.

A. Bayesian Learning-based Centralized Beamformer Design

The centralized beamformer design problem, which minimizes the total transmit power of the system while simultaneously satisfying the QoS requirement, can be formulated as

$$\begin{aligned} \min_{\{\mathbf{F}_{\mathrm{RF},n}\},\{\mathbf{F}_{\mathrm{BB},nu}\}} & \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \left\|\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\right\|^2\right) \\ \text{s.t.} \quad & \mathrm{SINR}_{nu} \geq \gamma_{nu}, \quad \forall n, u, \\ & |\mathbf{F}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{M}}, \quad \forall n, \end{aligned} \tag{4}$$

where β_n denotes the power weighting factor of BS_n and SINR_{nu} is as defined in (3). It is easy to see that the above optimization problem is non-convex in nature due to the constant magnitude constraint for the elements of each RF TPC, which renders it challenging to solve. Hence, for ensuring mathematical tractability, a two-step procedure can be conceived in order to efficiently solve the above power minimization problem. In the first step, the optimal fully-digital TPC $\mathbf{f}_{nu}$ is determined using the popular semidefinite relaxation (SDR) technique. Then Bayesian learning is leveraged subsequently for decomposing $\mathbf{f}_{nu}$ into the corresponding RF and baseband TPCs. To this end, substituting $\mathbf{f}_{nu} = \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}$ into (4), the optimization problem can be recast as

$$\begin{aligned} \min_{\{\mathbf{f}_{nu}\}} \quad & \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \left(\mathbf{f}_{nu}\mathbf{f}_{nu}^H \right) \right) \\ \text{s.t.} \quad & \mathrm{SINR}_{nu} \geq \gamma_{nu}, \quad \forall n, u, \end{aligned} \tag{5}$$

where the quantity $\|\mathbf{f}_{nu}\|^2$ has been replaced by $\mathrm{Tr}\left(\mathbf{f}_{nu}\mathbf{f}_{nu}^H\right)$. The SDR principle can now be invoked upon replacing the matrix $\mathbf{f}_{nu}\mathbf{f}_{nu}^H$ with rank-1 positive semidefinite matrix (PSD) $\mathbf{F}_{nu} \succeq \mathbf{0}$. The resultant optimization problem is given by the

$$\min_{\{\mathbf{f}_{nu}\}} \quad \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) \right) \tag{6a}$$

$$\text{s.t.} \quad \mathrm{SINR}_{nu} \geq \gamma_{nu}, \quad \forall n, u, \tag{6b}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \quad \mathrm{rank}\left(\mathbf{F}_{nu}\right) = 1, \forall n, u, \tag{6c}$$

where the quantity SINR_{nu} can be rewritten as

$$\mathrm{SINR}_{nu} = \frac{\mathrm{Tr}\left(\mathbf{h}_{nnu}^H \mathbf{F}_{nu} \mathbf{h}_{nnu}\right)}{\sum\limits_{i \neq u}^{U} \mathrm{Tr}(\mathbf{h}_{nnu}^H \mathbf{F}_{ni} \mathbf{h}_{nnu}) + \sum\limits_{m \neq n}^{N_c} \sum\limits_{i=1}^{U} \mathrm{Tr}(\mathbf{h}_{mnu}^H \mathbf{F}_{mi} \mathbf{h}_{mnu}) + \sigma_{nu}^2}. \tag{7}$$

Relaxing the unity rank constraint in (6c) and expanding the quantity SINR_{nu} for each n, u, we have

$$\min_{\{\mathbf{f}_{nu}\}} \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) \right) \tag{8a}$$

$$\text{s.t} \quad \frac{1}{\gamma_{nu}} \mathrm{Tr}\Big(\mathbf{H}_{nnu}\mathbf{F}_{nu}\Big) - \sum_{j \neq u}^{U} \mathrm{Tr}\Big(\mathbf{H}_{nnu}\mathbf{F}_{nj}\Big) \tag{8b}$$

$$\geq \sum_{m \neq n}^{N_c} \sum_{i=1}^{U} \mathrm{Tr}\Big(\mathbf{H}_{mnu}\mathbf{F}_{mi}\Big) + \sigma_{nu}^2, \forall n, u, \tag{8c}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \forall n, u, \tag{8d}$$

where $\mathbf{H}_{mnu} = \mathbf{h}_{mnu}\mathbf{h}_{mnu}^H$. As a result of the above manipulations, it can be seen that the ensuing optimization problem in (8) is convex and can be efficiently solved using widely available tools such as CVX [50]. The optimal digital beamformer $\mathbf{f}_{nu}$ can be finally determined as the eigenvector corresponding to the largest eigenvalue of $\mathbf{F}_{nu}^*$,

Algorithm 1: BL-based hybrid TPC design

1 **Input:** Concatenated optimal digital TPC matrix $\mathbf{F}_{n,\mathrm{opt}}$, dictionary matrix $\mathbf{A}_T$, RF chains $N_{\mathrm{RF},n}$, variance of approximation error σ_e^2, stopping parameters υ and $n_{\max}$;

2 **Initialization:** $\hat{\gamma}_i^{(0)} = 1$, $\forall 1 \leq i \leq G \Rightarrow \hat{\mathbf{\Gamma}}^{(0)}$, set counter $k = 0$ and $\mathbf{\Gamma}^{(-1)} = \mathbf{0}$;

3 **while** $\left(\left\| \hat{\boldsymbol{\gamma}}^{(k)} - \hat{\boldsymbol{\gamma}}^{(k-1)} \right\|_2^2 > \upsilon \text{ and } k < n_{\max} \right)$ **do**

i **E-step:** Evaluate the *a posteriori* covariance and mean

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

$$\tilde{\mathbf{F}}_{\mathrm{BB},n}^{(k)} = \frac{1}{\sigma_e^2}\mathbf{\Gamma}^{(k)}\mathbf{A}_T^H\mathbf{F}_{n,\mathrm{opt}}$$

ii **M-step:** Update the hyperparameters as

for $i = 1, 2, \ldots, G$

$$\hat{\gamma}_i^{(k)} = \mathbf{\Sigma}^{(k)}(i,i) + \frac{1}{U}\sum_{j=1}^{U} \left| \tilde{\mathbf{F}}_{\mathrm{BB},n}^{(k)}(i,j) \right|^2$$

end for

end

4 **Output:** Obtain $\mathbf{F}_{\mathrm{BB},n}^*$ and $\mathbf{F}_{\mathrm{RF},n}^*$ using (46) and (47).

i.e., $\mathbf{f}_{nu,\mathrm{opt}} = \sqrt{\lambda}_{\max}\hat{\mathbf{f}}_{nu}$, where $\hat{\mathbf{f}}_{nu}$ represents the unit-norm eigenvector corresponding to $\lambda_{\max}$. The state-of-the-art BL principle can now be exploited for decomposing the optimal fully digital TPC $\mathbf{f}_{nu,\mathrm{opt}}$ into the constituent RF and BB precoders, as shown in Appendix. A succinct description of the various steps in the BL-based hybrid TPC design procedure is presented in Algorithm 1. The above centralized beamforming solution, although performs well, it is based on the assumption that the CU has perfect knowledge of the CSI of all the MSs, i.e., requires global CSI. However, this may be challenging to achieve in practice as it requires excessive backhaul capacity, which increases substantially with the number of coordinated cells. Therefore, the next section develops an ADMM-based distributed beamforming solution, which employs only the local CSI available at each BS, thus leading to significant savings in terms of the backhaul cost and complexity.

III. ADMM BASED DISTRIBUTED MULTICELL COORDINATED BEAMFORMING

ADMM is an advanced optimization method that combines the idea of dual decomposition and of the augmented Lagrangian method, which is ideally suited for distributed optimization due to its excellent convergence properties [51]. In order to develop our ADMM-based distributed design, the centralized beamformer design problem of (8) formulated for the MUMC system can be recast by defining the following

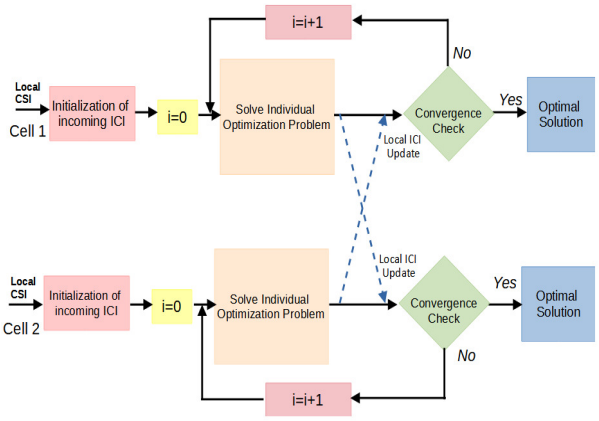


Figure 2: Flow-chart of the distributed beamformer design for a two-cell scenario

auxiliary variables:

$$p_n = \sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}), \quad \forall n, \tag{9a}$$

$$q_{mnu} = \sum_{i=1}^{U} \mathrm{Tr}\Big(\mathbf{H}_{mnu}\mathbf{F}_{mi}\Big), \quad \forall m,n,u, \tag{9b}$$

$$Q_{nu} = \sum_{m\neq n}^{N_c} q_{mnu}, \forall n,u. \tag{9c}$$

The quantity p_n denotes the transmission power corresponding to BS_n, q_{mnu} represents the inter-BS interference power impinging from BS_m upon MS_{nu} and Q_{nu} represents the total inter BS interference power arriving from the neighbouring BSs to MS_{nu}. Using (9), the centralized TPC design problem of (8) can be reformulated as

$$\min_{\{\mathbf{f}_{nu}\}} \quad \sum_{n=1}^{N_c} \beta_n p_n \tag{10a}$$

$$\text{s.t.} \quad p_n = \sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}), \quad \forall n, \tag{10b}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \qquad \forall n,u, \tag{10c}$$

$$\frac{1}{\gamma_{nu}}\mathrm{Tr}\Big(\mathbf{H}_{nnu}\mathbf{F}_{nu}\Big) - \sum_{j\neq u}^{U}\mathrm{Tr}\Big(\mathbf{H}_{nnu}\mathbf{F}_{nj}\Big) \geq Q_{nu} + \sigma_{nu}^2, \forall n,u, \tag{10d}$$

$$q_{mnu} = \sum_{i=1}^{U} \mathrm{Tr}\Big(\mathbf{H}_{mnu}\mathbf{F}_{mi}\Big), \quad \forall m,n, \tag{10e}$$

$$Q_{nu} = \sum_{m\neq n}^{N_c} q_{mnu} \geq 0 \quad \forall n,u,m\neq n. \tag{10f}$$

It can be readily observed from the SINR constraint (10d), that each MS_{nu} in cell n experiences only the sum of the interference power Q_{nu} of (10f) from all other cells. Interestingly, it can be observed that the constraint in Eq.(10e) remains unchanged upon interchanging the subindices m and n. Hence, the constraints in (10b) to (10f) can be decomposed into N_c independent convex sets by interchanging the subindices m and n in (10e) as follows:

$$\mathcal{C}_n = \Bigg\{ \big(\{\mathbf{F}_{nu}\}_{n,u}, p_n, \{Q_{nu}\}_u, \{q_{nmu}\}_{m,u}\big) \Big| \\ p_n = \sum_{u=1}^{U}\mathrm{Tr}(\mathbf{F}_{nu}), \\ q_{nmu} = \sum_{i=1}^{U}\mathrm{Tr}(\mathbf{H}_{nmu}\mathbf{F}_{ni}), \quad \forall m\neq n, \ \forall u, \\ \frac{1}{\gamma_{nu}}\mathrm{Tr}(\mathbf{H}_{nnu}\mathbf{F}_{nu}) - \sum_{j\neq u}^{U}\mathrm{Tr}(\mathbf{H}_{nnu}\mathbf{F}_{nj}) \geq Q_{nu} + \sigma_{nu}^2, \\ \mathbf{F}_{nu}\succeq\mathbf{0}, \ Q_{nu}\geq 0, \ \forall u \Bigg\}, \forall n \in N_c, u. \tag{11}$$

Additionally, we define the new variables:

$$\mathbf{q} = \big[[q_{121},\ldots,q_{12U}],\ldots,[q_{N_c(N_c-1)1},\ldots,q_{N_c(N_c-1)U}]\big]^T \\ \in \mathbb{R}^{N_c(N_c-1)U}, \tag{12}$$

$$\mathbf{q}_n = \big[[Q_{n1},\ldots,Q_{nU}],[q_{n11},\ldots,q_{n1U}],\ldots, \\ \ldots,[q_{nN_c1},\ldots,q_{nN_cU}]\big]^T \in \mathbb{R}_+^{N_cU}, \ \forall n, \tag{13}$$

where the vector $\mathbf{q}$ collects all the global ICI variables and $\mathbf{q}_n$ collects the local ICI variables from all the other cells, i.e. $\{Q_{nu}\}_{u=1}^{U}$ and $\{q_{nmu}\}_{m,u}$ for $m \in N_c \setminus \{n\}$. Here, it is interesting to observe that $\mathbf{q}_n$ represents the total interference experienced by the BS_n and it also denotes the total interference experienced by other cells inflicted by BS_n. Furthermore, we have $\mathbf{q}_n = \mathbf{J}_n\mathbf{q}$, where $\mathbf{J}_n \in \{0,1\}^{N_cU\times N_c(N_c-1)U}$ denotes the linear mapping matrix. Therefore, Problem (10) can be reformulated as:

$$\min_{\{\mathbf{F}_{nu},\mathbf{q}_n,p_n,\mathbf{q}\}} \sum_{n=1}^{N_c}\beta_n p_n \tag{14a}$$

$$\text{s.t.} \quad (\{\mathbf{F}_{nu}\}_u, \mathbf{q}_n, p_n) \in \mathcal{C}_n, \quad \forall n, \tag{14b}$$

$$\mathbf{q}_n = \mathbf{J}_n\mathbf{q}, \quad \forall n. \tag{14c}$$

Upon applying the ADMM technique to (14), one obtains

$$\min_{\substack{\{\mathbf{F}_{nu},\mathbf{q}_n,\\ p_n,\rho_n,\mathbf{q}\}}} \Bigg\{ \sum_{n=1}^{N_c}\beta_n p_n + \frac{c}{2}\sum_{n=1}^{N_c}\|\mathbf{J}_n\mathbf{q}-\mathbf{q}_n\|^2 \\ + \frac{c}{2}\sum_{n=1}^{N_c}(\rho_n - p_n)^2 \Bigg\} \tag{15a}$$

$$\text{s.t.} \quad (\{\mathbf{F}_{nu}\}_u, \mathbf{q}_n, p_n) \in \mathcal{C}_n, \ \forall n, \tag{15b}$$

$$\mathbf{q}_n = \mathbf{J}_n\mathbf{q} \quad \forall n, \tag{15c}$$

$$p_n = \rho_n \ \forall n, \tag{15d}$$

Note that $\rho_n \geq 0, \forall n$ represents the slack variables, which are employed for imposing the penalty term $\frac{c}{2}\sum_{n=1}^{N_c}(\rho_n - p_n)^2$ on the objective (15). Thus, Problem (15) is equivalent to problem (14). One can now define the following quantities

to establish a correspondence between (15) and the canonical form of ADMM as

$$\mathbf{x} = \left[\mathbf{q}^T, \rho_1, \ldots, \rho_{N_c}\right]^T, \mathbf{z} = -\left[\mathbf{q}_1^T, \quad \ldots, \mathbf{q}_{N_c}^T, p_1, \ldots, p_{N_c}\right]^T,$$
$$F(\mathbf{x}) = 0, \quad G(\mathbf{z}) = \sum_{n=1}^{N_c} \beta_n p_n,$$
$$\mathbf{A} = \begin{bmatrix} \mathbf{J} & \mathbf{0}, \\ \mathbf{0} & \mathbf{I}_{N_c} \end{bmatrix}, \mathbf{B} = \mathbf{I}, \mathbf{y} = \mathbf{0}$$
$$\boldsymbol{\xi} = \left[\boldsymbol{\nu}_1^T, \ldots, \boldsymbol{\nu}_{N_c}^T, \mu_1, \ldots, \mu_{N_c}\right]^T,$$
$$\mathcal{S}_1 = \mathbb{R}^{N_c(N_c-1)U+N_c},$$
$$\mathcal{S}_2 = \Big\{ \left[\mathbf{q}_1^T, \ldots, \mathbf{q}_{N_c}^T, p_1, \ldots, p_{N_c}\right]^T \Big| \left(\{\mathbf{F}_{nu}\}_u, \mathbf{q}_n, p_n\right) \in \mathcal{C}_n, n \in N_c \Big\}, \tag{16}$$

where $\mathbf{J} = [\mathbf{J}_1^T, \cdots, \mathbf{J}_{N_c}^T]^T$, $\boldsymbol{\nu} \in \mathbb{R}^{N_c U}$ and $\mu_n \in \mathbb{R}, \forall n$, are the dual variables associated with the constraints (15c) and (15d), respectively. Using the ADMM principle, the optimization problem in (15) can be recast as

$$\min_{\substack{\{\mathbf{F}_{nu}\}_u, \{\lambda_{nmu}\}_{m,u}, \\ \mathbf{q}_n, p_n, n=1,\ldots,N_c}} \sum_{n=1}^{N_c} \Bigg\{ \beta_n p_n + \frac{c}{2} \left\| \mathbf{J}_n \mathbf{q}^{(i)} - \mathbf{q}_n \right\|^2 + \frac{c}{2} \left(\rho_n^{(i)} - p_n \right)^2 - \boldsymbol{\nu}_n^{(i)^T} \mathbf{q}_n - \mu_n^{(i)} p_n \Bigg\} \tag{17a}$$

$$\text{s.t.} \quad \left(\left\{\mathbf{F}_{nu}\right\}_u, \mathbf{q}_n, p_n\right) \in \mathcal{C}_n, \forall n. \tag{17b}$$

The decoupled problem formulated for our distributed beamformer design for the nth cell of the MUMC system can now be expressed as

$$\left\{\mathbf{q}_n^{(i+1)}, p_n^{(i+1)}\right\} = \arg\min \Bigg\{ \beta_n p_n + \frac{c^{(i)}}{2} \left\| \mathbf{J}_n \mathbf{q}^{(i)} - \mathbf{q}_n \right\|^2 + \frac{c^{(i)}}{2} \left(\rho_n^{(i)} - p_n \right)^2 - \boldsymbol{\nu}_n^{(i)^T} \mathbf{q}_n - \mu_n^{(i)} p_n \Bigg\},$$
$$\text{s.t.} \quad \left(\left\{\mathbf{F}_{nu}\right\}_u, \mathbf{q}_n, p_n\right) \in \mathcal{C}_n. \tag{18}$$

The above Problem (18) is convex for each cell and can be readily solved using a suitable convex solver. The dual variables $\boldsymbol{\nu}_n$ and μ_n can be updated as

$$\boldsymbol{\nu}_n^{(i+1)} = \boldsymbol{\nu}_n^{(i)} + c^{(i)} \left(\mathbf{J}_n \mathbf{q}^{(i+1)} - \mathbf{q}_n^{(i+1)} \right), \quad \forall n, \tag{19a}$$

$$\mu_n^{(i+1)} = \mu_n^{(i)} + c^{(i)} \left(\rho_n^{(i+1)} - p_n^{(i+1)} \right), \quad \forall n, \tag{19b}$$

where the intermediate problems of updating the variables $\mathbf{q}^{(i+1)}$ and $\rho_n^{(i+1)}$ are given by

$$\mathbf{q}^{(i+1)} = \underset{\mathbf{q} \in \mathbb{R}^{N_c(N_c-1)U}}{\arg\min} \frac{c^{(i)}}{2} \sum_{n=1}^{N_c} \|\mathbf{J}_n \mathbf{q} - \mathbf{q}_n^{(i+1)}\|^2 + \sum_{n=1}^{N_c} \boldsymbol{\nu}_n^{(i)^T} \mathbf{J}_n \mathbf{q}, \tag{20}$$

$$\left\{\rho_n^{(i+1)}\right\}_{n=1}^{N_c} = \underset{\substack{\rho_n \in \mathbb{R}, \\ n=1,\ldots,N_c}}{\arg\min} \frac{c^{(i)}}{2} \sum_{n=1}^{N_c} \left(\rho_n - p_n^{(i+1)} \right)^2 + \sum_{n=1}^{N_c} \mu_n^{(i)} \rho_n. \tag{21}$$

Since the Problems (20) and (21) are convex quadratic in nature, their closed form solutions can be determined as shown below

$$\mathbf{q}^{(i+1)} = \mathbf{J}^\dagger \left(\tilde{\mathbf{q}}^{(i+1)} - \frac{1}{c} \tilde{\boldsymbol{\nu}}^{(i)} \right), \tag{22a}$$

$$\rho_n^{(i+1)} = p_n^{(i+1)} - \frac{1}{c} \mu_n^{(i)}, \quad \forall n, \tag{22b}$$

where $\tilde{\mathbf{q}}^{(i+1)} = \left[\left(\mathbf{q}_1^{(i+1)}\right)^T, \ldots, \left(\mathbf{q}_{N_c}^{(i+1)}\right)^T \right]^T$ and $\tilde{\boldsymbol{\nu}}^{(i)} = \left[\left(\boldsymbol{\nu}_1^{(i)}\right)^T, \ldots, \left(\boldsymbol{\nu}_{N_c}^{(i)}\right)^T \right]^T$.

The proposed ADMM-based distributed beamformer design for the MUMC system is summarized next. The ADMM steps in (18), (19), and (22) can be evaluated independently at each BS in a distributed manner using purely local CSI, which defined as that of the users in its own cell. Subsequently, each BS sends the updated local information $\{\mathbf{q}_n\}$ to all the other BSs via the CU. Using $\{\mathbf{q}_n\}$, the public ICI variable $\mathbf{q}$ is calculated iteratively from equation (22a) at each BS, which is further exploited for updating $\boldsymbol{\nu}_n^{(i+1)}$ using (19a) at each BS. Furthermore, each BS can independently update the dual variables $\rho_n^{(i+1)}$ and $\mu_n^{(i+1)}$. These steps are summarized in Algorithm 2. As a visual aid, the flow chart of a simple two-cell scenario is shown in Fig. 2. Finally, the corresponding hybrid TPC design can once again be obtained using the BL-technique derived in Section II.

Algorithm 2: ADMM based distributed TPC for MUMC Systems

1. **Initialization**: Set $i = 0$ and initialize $\{\boldsymbol{\nu}_n^{(i)}, \mu_n^{(i)}, \mathbf{q}^{(i)}, \rho_n^{(i)}\}$ with zeros for each cell, choose a penalty parameter $c > 0$;
2. **while** *(stopping criteria is satisfied)* **do**
3. solve the beamformer design problem corresponding to each cell using equation (18) to obtain the local ICI iterate $\{\mathbf{q}_n^{(i+1)}\}$ and $\{p_n^{(i+1)}\}$;
4. inform the local ICI $\{\mathbf{q}_n^{(i+1)}\}$ to all the neighbouring BSs;
5. update the public ICI $\{\mathbf{q}^{(i+1)}\}$ and $\{\rho_n^{(i+1)}\}$ using (22a) and (22b) respectively;
6. update the dual variables $\{\boldsymbol{\nu}_n\}$ and $\{\mu_n\}$ using (19a) and (19b) respectively;
7. $i \leftarrow i + 1$;

IV. Robust Beamformer Design with Imperfect CSI

The previous sections described both our centralized and distributed beamformers designed for an MUMC system having perfect CSI. However, in practice, this is challenging due to CSI impairments arising from channel estimation error, quantization and feedback. The imperfect CSI degrades the performance of the beamformer design. To overcome this problem, in this section, we conceive techniques for robust MUMC beamformer design with imperfect CSI. Let $\hat{\mathbf{h}}_{mnu} \in \mathbb{C}^M$, $\forall m, n$ denote the available CSI. The true CSI $\mathbf{h}_{mnu}$ may then be modelled as

$$\mathbf{h}_{mnu} = \hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu}, \tag{23}$$

where the quantity $\mathbf{\Delta}_{mnu} \in \mathbb{C}^M$ denotes the CSI error that can be modeled as

$$\mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1. \tag{24}$$

This is termed as the ellipsoidal uncertainty model and $\mathbf{Q}_{mnu}$ denotes a positive definite (PD) matrix that characterizes the CSI uncertainty ellipsoid, as described in [49]. When $\mathbf{Q}_{mnu} = r_{mnu}^{-2}\mathbf{I}_M$, where $r_{mnu}^2 > 0$, (24) reduces to the popular spherical model $\|\mathbf{\Delta}_{mnu}\|^2 \leq r_{mnu}^2$.

A. Centralized MUMC robust beamformer design

The robust MUMC beamformer design problem for minimizing the total transmit power while satisfying the SINR constraint at each MS for the worst-case channel information can be formulated as

$$\min_{\{\mathbf{F}_{\mathrm{RF},n}\},\{\mathbf{F}_{\mathrm{BB},nu}\}} \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \|\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\|^2 \right) \tag{25a}$$

$$\text{s.t.} \quad \mathrm{SINR}_{nu} \geq \gamma_{nu},$$

$$\forall\ \mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1, \forall m, n, \tag{25b}$$

$$|\mathbf{F}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{M}} \quad \forall n. \tag{25c}$$

The corresponding SINR_{nu} expression is given by

$$\mathrm{SINR}_{nu} = \frac{\left| \left(\hat{\mathbf{h}}_{nnu} + \mathbf{\Delta}_{nnu} \right)^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu} \right|^2}{\left\{ \sum_{i\neq u}^{U} \left| \left(\hat{\mathbf{h}}_{nnu} + \mathbf{\Delta}_{nnu} \right)^H \mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},ni} \right|^2 + \sum_{m\neq n}^{N_c} \sum_{i=1}^{U} \left| \left(\hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu} \right)^H \mathbf{F}_{\mathrm{RF},m}\mathbf{f}_{\mathrm{BB},mi} \right|^2 + \sigma_{nu}^2 \right\}}, \tag{26}$$

Similar to Section II-A, in order to render the optimization problem tractable, substituting $\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu} = \mathbf{f}_{nu}$ as the fully digital TPC, the SINR constraint of each MS can be written as

$$\left(\hat{\mathbf{h}}_{nnu}^H + \mathbf{\Delta}_{nnu}^H\right) \left(\frac{1}{\gamma_{nu}} \mathbf{f}_{nu}\mathbf{f}_{nu}^H - \sum_{i\neq u}^{U} \mathbf{f}_{ni}\mathbf{f}_{ni}^H \right) \left(\hat{\mathbf{h}}_{nnu} + \mathbf{\Delta}_{nnu}\right)$$
$$\geq \sum_{m\neq n}^{N_c} \left(\hat{\mathbf{h}}_{mnu}^H + \mathbf{\Delta}_{mnu}^H\right) \left(\sum_{i=1}^{U} \mathbf{f}_{mi}\mathbf{f}_{mi}^H \right) \left(\hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu}\right) + \sigma_{nu}^2,$$
$$\forall\ \mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1, \forall n, u. \tag{27}$$

Then SDR can once again be employed upon replacing the matrix $\mathbf{f}_{nu}\mathbf{f}_{nu}^H$ by the PSD matrix $\mathbf{F}_{nu} \succeq \mathbf{0}$ and relaxing the rank constraint to obtain the problem

$$\min_{\{\mathbf{F}_{nu}\}} \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) \right) \tag{28a}$$

$$\text{s.t.} \left(\hat{\mathbf{h}}_{nnu}^H + \mathbf{\Delta}_{nnu}^H\right) \left(\frac{1}{\gamma_{nu}} \mathbf{F}_{nu} - \sum_{i\neq u}^{U} \mathbf{F}_{ni} \right) \left(\hat{\mathbf{h}}_{nnu} + \mathbf{\Delta}_{nnu}\right)$$
$$\geq \sum_{m\neq n}^{N_c} \left(\hat{\mathbf{h}}_{mnu}^H + \mathbf{\Delta}_{mnu}^H\right) \left(\sum_{i=1}^{U} \mathbf{F}_{mi} \right) \left(\hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu}\right) + \sigma_{nu}^2,$$
$$\forall\ \mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1, \quad \forall n, u, \tag{28b}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}. \tag{28c}$$

Although the relaxed problem in (28) above is now convex, it remains computationally challenging to solve due to of infinitely many SINR constraints (28b). Interestingly, these infinitely many constrains can be converted into a finite number of constrains with aid of the S-lemma [49]. To this end, the SINR constraint of (28b) can be rewritten as

$$\min_{\mathbf{\Delta}_{nnu}^H \mathbf{Q}_{nnu} \mathbf{\Delta}_{nnu} \leq 1} \left(\hat{\mathbf{h}}_{nnu}^H + \mathbf{\Delta}_{nnu}^H\right) \left(\frac{1}{\gamma_{nu}} \mathbf{F}_{nu} - \sum_{i\neq u}^{U} \mathbf{F}_{ni} \right) \times \left(\hat{\mathbf{h}}_{nnu} + \mathbf{\Delta}_{nnu}\right)$$
$$\geq \sum_{m\neq n}^{N_c} \left\{ \max_{\mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1} \left(\hat{\mathbf{h}}_{mnu}^H + \mathbf{\Delta}_{mnu}^H\right) \times \left(\sum_{i=1}^{U} \mathbf{F}_{mi} \right) \left(\hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu}\right) \right\} + \sigma_{nu}^2. \tag{29}$$

Note that, the right hand side term for each m represents the worst-case ICI power arising from BS_m to MS_{nu}, $\forall\ m \in \mathcal{N}_c \setminus \{n\}$. Hence, introducing the slack variable q_{mnu} as

$$q_{mnu} = \max_{\mathbf{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \mathbf{\Delta}_{mnu} \leq 1} \left(\hat{\mathbf{h}}_{mnu}^H + \mathbf{\Delta}_{mnu}^H\right) \left(\sum_{i=1}^{U} \mathbf{F}_{mi} \right) \times \left(\hat{\mathbf{h}}_{mnu} + \mathbf{\Delta}_{mnu}\right), \tag{30}$$

the problem (28) can be reformulated as

$$\min_{\{\mathbf{F}_{nu}\}} \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) \right) \tag{31a}$$

$$\text{s.t.} \left(\hat{\mathbf{h}}_{nnu}^H + \boldsymbol{\Delta}_{nnu}^H\right) \left(\frac{1}{\gamma_{nk}} \mathbf{F}_{nu} - \sum_{i \neq u}^{U} \mathbf{F}_{ni} \right) \left(\hat{\mathbf{h}}_{nnu} + \boldsymbol{\Delta}_{nnu}\right)$$

$$\geq \sum_{m \neq n}^{N_c} q_{mnu} + \sigma_{nu}^2, \ \forall \ \boldsymbol{\Delta}_{nnu}^H \mathbf{Q}_{nnu} \boldsymbol{\Delta}_{nnu} \leq 1, \tag{31b}$$

$$\left(\hat{\mathbf{h}}_{mnu}^H + \boldsymbol{\Delta}_{mnu}^H\right) \left(\sum_{i=1}^{U} \mathbf{F}_{mi} \right) \left(\hat{\mathbf{h}}_{mnu} + \boldsymbol{\Delta}_{mnu}\right)$$

$$\leq q_{mnu}, \forall \ \boldsymbol{\Delta}_{mnu}^H \mathbf{Q}_{mnu} \boldsymbol{\Delta}_{mnu} \leq 1, \forall m \neq n, u, \tag{31c}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}. \tag{31d}$$

In order to apply the S-lemma [49], the constraints in (31b) and (31c) can be reformulated as the positive semi-definite (PSD) matrices

$$\begin{aligned} &\boldsymbol{\Phi}_{nu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, \{q_{mnu}\}_m, \lambda_{nnu} \right) \\ &\triangleq \begin{bmatrix} \mathbf{I} \\ \hat{\mathbf{h}}_{nnu}^H \end{bmatrix} \left(\frac{1}{\gamma_{nu}} \mathbf{F}_{nu} - \sum_{i \neq u}^{U} \mathbf{F}_{ni} \right) \begin{bmatrix} \mathbf{I} & \hat{\mathbf{h}}_{nnu} \end{bmatrix} \\ &+ \begin{bmatrix} \lambda_{nnu} \mathbf{Q}_{nnu} & \mathbf{0} \\ \mathbf{0} & -\sigma_{nu}^2 - \sum\limits_{m \neq n}^{N_c} q_{mnu} - \lambda_{nnu} \end{bmatrix}, \end{aligned} \tag{32}$$

$$\begin{aligned} &\boldsymbol{\Psi}_{mnu} \left(\{\mathbf{F}_{mi}\}_{i=1}^{U}, q_{mnu}, \lambda_{mnu} \right) \\ &\triangleq \begin{bmatrix} \mathbf{I} \\ \hat{\mathbf{h}}_{mnu}^H \end{bmatrix} \left(- \sum_{i=1}^{U} \mathbf{F}_{mi} \right) \begin{bmatrix} \mathbf{I} & \hat{\mathbf{h}}_{mnu} \end{bmatrix} \\ &+ \begin{bmatrix} \lambda_{mnu} \mathbf{Q}_{mnu} & \mathbf{0} \\ \mathbf{0} & q_{mnu} - \lambda_{mnu} \end{bmatrix}. \end{aligned} \tag{33}$$

Upon using the above quantities, the optimization Problem (31) can be recast as:

$$\min_{\{\mathbf{F}_{nu}\}, \{\lambda_{mnu}\}, \{q_{mnu}\}} \sum_{n=1}^{N_c} \beta_n \left(\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) \right) \tag{34a}$$

$$\text{s.t.} \ \boldsymbol{\Phi}_{nu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, \{q_{mnu}\}_m, \lambda_{nnu} \right) \succeq \mathbf{0}, \tag{34b}$$

$$\boldsymbol{\Psi}_{mnu} \left(\{\mathbf{F}_{mi}\}_{i=1}^{U}, q_{mnu}, \lambda_{mnu} \right) \succeq \mathbf{0}, \forall m \neq n, \tag{34c}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \tag{34d}$$

$$\lambda_{mnu} \geq 0, \forall m. \tag{34e}$$

Observe that the Problem in (34) is a convex semi-definite program (SDP) that can be solved efficiently and the optimal solution $\mathbf{f}_{nu,\mathrm{opt}}$ can be obtained as the dominant unit-norm eigenvector of the matrix $\mathbf{F}_{nu}$, as described in Section II. Finally, the corresponding hybrid TPC design can be obtained using the BL-technique derived in Appendix 1. The corresponding ADMM-based robust distributed mmWave TPC design relying on imperfect CSI is described next.

V. Robust Distributed MUMC mmWave Beamformer Design

We commence by defining the following auxiliary variables, similar to the perfect CSI model of Section II-A:

$$p_n = \sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}), \quad Q_{nu} = \sum_{m \neq n}^{N_c} q_{mnu} \ \ \forall n, u. \tag{35}$$

Recall that the variable p_n denotes the transmit power corresponding to BS_n and q_{mnu} represents the worst-case ICI power emanating from the BS_m to MS_{nu}, whereas Q_{nu} represents the total ICI power arriving from the neighbouring BSs to MS_{nu}. Using the auxiliary variables of (35), the Problem (34) can be expressed as

$$\min_{\substack{\{\mathbf{F}_{nu} \succeq \mathbf{0}\}, \{p_n\}, \\ \{\lambda_{mnu} \geq 0\}, \{q_{mnu}\}}} \sum_{n=1}^{N_c} \beta_n p_n \tag{36a}$$

$$\text{s.t.} \ \ \boldsymbol{\Phi}_{nu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, Q_{nu}, \lambda_{nnu} \right) \succeq \mathbf{0}, \forall n, u, \tag{36b}$$

$$\boldsymbol{\Psi}_{nmu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, q_{nmu}, \lambda_{nmu} \right) \succeq \mathbf{0}, \forall m \neq n, u, \tag{36c}$$

$$\sum_{u=1}^{U} \mathrm{Tr}(\mathbf{F}_{nu}) = p_n, \forall n, \tag{36d}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \forall n, u, \tag{36e}$$

$$\lambda_{mnu} \geq 0, \forall m \neq n, u. \tag{36f}$$

The constraints (36b)-(36f) can be decomposed into N_c independent convex sets as follows

$$\begin{aligned} \mathcal{C}_n = \Bigg\{ &\left(\{\mathbf{F}_{nu}\}_u, \{\lambda_{nmu}\}_{m,u}, \{Q_{nu}\}_u, \{q_{nmu}\}_{m,u}, p_n \right) | \\ &\boldsymbol{\Phi}_{nu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, Q_{nu}, \lambda_{nnu} \right) \succeq \mathbf{0}, \forall u, \\ &\boldsymbol{\Psi}_{nmu} \left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, q_{nmu}, \lambda_{nmu} \right) \succeq \mathbf{0}, \forall m \neq n, u, \\ &\lambda_{nmu} \geq 0, \quad \forall m, n, u, \\ &\sum_{u=1}^{U} \mathrm{Tr}\left(\mathbf{F}_{nu}\right) = p_n, \\ &\mathbf{F}_{nu} \succeq \mathbf{0}, \quad \forall u, \\ &Q_{nu} \geq 0, \quad \forall u \Bigg\}, \forall n. \end{aligned} \tag{37}$$

Finally, the optimization Problem (36) can be reformulated as

$$\min_{\substack{\{\mathbf{F}_{nu}\}, \{\lambda_{nmu}\}, \\ \{\mathbf{q}_n\}, \{p_n\}, \mathbf{q}}} \sum_{n=1}^{N_c} \beta_n p_n \tag{38a}$$

$$\text{s.t.} \quad \left(\{\mathbf{F}_{nu}\}_k, \{\lambda_{nmu}\}_{m,u}, \mathbf{q}_n, p_n \right) \in \mathcal{C}_n, \forall n, \tag{38b}$$

$$\mathbf{q}_n = \mathbf{J}_n \mathbf{q}, \quad \forall n, \tag{38c}$$

which is similar to Problem (14). Hence, Algorithm 2 can be readily applied for finding the optimal distributed robust TPC solution. The corresponding hybrid TPC can be

obtained using the BL-technique derived in Section II. Using Algorithm 2, each BS iteratively approaches the optimal beamforming solution until the relevant ICI information $\left(\mathbf{q}_n^{(i+1)}\right)$ for BS_n is extracted from the ICI information $\left(\mathbf{q}^{(i+1)}\right)$ i.e. $\mathbf{J}_n\mathbf{q}^{(i+1)} = \mathbf{q}_n^{(i+1)}$ for all n. It is important to note that the quantities $\{\mathbf{F}_{nu}\}$ and $\{\lambda_{nmu}\}$ obtained in Step-4 of Algorithm 2 may not be feasible for the primal Problem (34), since the ADMM algorithm operates in the dual domain, which does not guarantee that the constraint $\mathbf{J}_n\mathbf{q}^{(i+1)} = \mathbf{q}_n^{(i+1)}$ holds true before convergence. However, each BS may perform one more optimization as

$$\begin{aligned} \min_{\{\mathbf{F}_{nu}\succeq\mathbf{0}\}_u,\{\lambda_{mnu}\geq 0\}_{n,u}} & \sum_{n=1}^{N_c}\beta_n\left(\sum_{u=1}^{U}\mathrm{Tr}(\mathbf{F}_{nu})\right) \\ \text{s.t.} \quad & \mathbf{\Phi}_{nu}\left(\{\mathbf{F}_{ni}\}_{i=1}^{U},\{q_{mnu}\}_m,\lambda_{nnu}\right)\succeq\mathbf{0}, \\ & \mathbf{\Psi}_{nmu}\left(\{\mathbf{F}_{ni}\}_{i=1}^{U},q_{nmu},\lambda_{nmu}\right)\succeq\mathbf{0},\forall m\neq n, \end{aligned} \tag{39}$$

using the tentatively consented ICI power vector $\mathbf{q}^{(i+1)}$. If the optimization Problem (37) yields feasible solutions for all the BSs, it follows that the computed quantities $\{\mathbf{F}_{nu}\}$ and $\{\lambda_{nmu}\}$ are feasible for the SDR Problem (34). Furthermore, additional iterations are needed for Algorithm 2 to converge if at least one of the BSs declares the infeasibility of (39), since it might not have reached a reasonable consensus concerning the global ICI $\mathbf{q}^{(i+1)}$. The signalling overhead required can be determined as follows. At each iteration of Algorithm 2, each BS update transmits its local ICI variable $\mathbf{q}_n$ to all other neighbouring BSs. Thus, Algorithm 2 requires each BS to transmit N_cU real values, i.e. $\mathbf{q}_n$, to the other (N_c-1) BSs, which leads to a total signaling overhead order of $(N_c-1)N_cU$. 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 moved to our technical report [52]. It can be observed that the complexity of Algorithm-2, which designs the fully-digital TPC is $\mathcal{O}\left(M^3\right)$. Next, the fully-digital TPC is decomposed into its constituent RF and BB precoders using Algorithm-1. This step incurs a complexity of order $\mathcal{O}\left(G^3\right)$, which can be attributed to the matrix inversion of size $[G\times G]$ in Eq.(43). Since $G >> M$, the overall complexity of the coordinated hybrid TPC can be closely approximated by $\mathcal{O}\left(G^3\right)$.

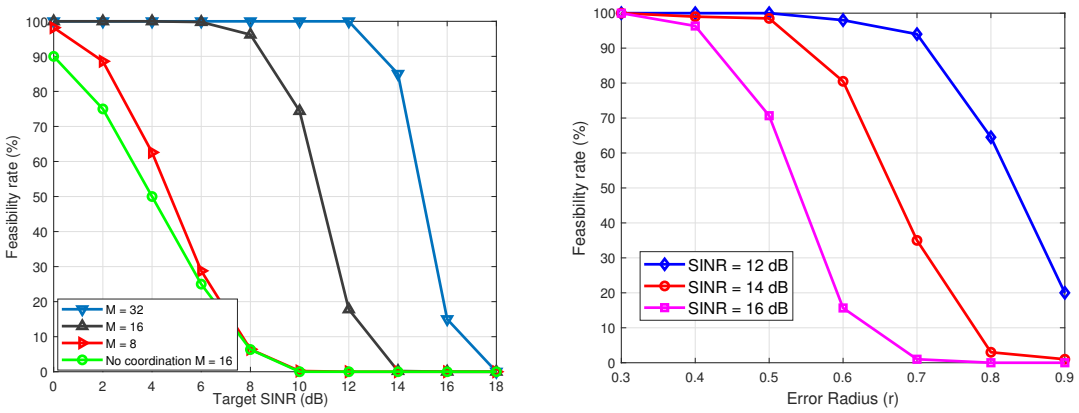
VI. Simulation Results

This section presents simulation results for characterizing the performance of the proposed hybrid beamformer design techniques for both centralized and distributed approaches while considering both perfect as well as imperfect CSI. Consider a MUMC network consisting of two cells, where the BS of each cell has $M\in\{8,16,32\}$ TAs serving the $U=2$ single-antenna users. The number of RF chains $N_{\mathrm{RF},n}$ at each BS is set to the number of users. These parameters are in line with those of other closely related works such as [33]–[35], [37], [38]. It is assumed that the large-scale fading has been perfectly compensated by power control at the BS, and hence only the effect of small-scale fading has been considered in this work, similar to [33], [37], [53], [54], and the multipath gains $\alpha_{l,mnu}$ are considered to be zero-mean complex Gaussian with unit variance. The inter-antenna spacing of the ULA at the BS is fixed as $\frac{\lambda_d}{2}$. The target SINRs of all the MSs are set equal, i.e., $\gamma_{nu}=\gamma$, $\forall n,u$. The power priority weight β_n is set to 1, $\forall n$. The noise power σ_n^2 is considered to be 0.1. For the BL-based TPC design procedure, the feasible AoD space Φ_T is comprised of $G=64$ angular grid points. Typically, the grid-size G is set to twice the number of antennas [19], as also considered in our simulation results, for achieving sufficiently high angular resolution. Furthermore, the stopping parameter (υ) and the maximum number of EM iterations $(n_{\max})$ are set to 10^{-5} and 50, respectively. We consider the spherical error model for CSI errors i.e., $\mathbf{Q}_{mnu}=r_{mnu}^{-2}\mathbf{I}_M$, where $r_{mnu}^2>0$ represents the radius of the circular uncertainty region. Table-II shows the detailed parameters considered in our simulations for centralized and distributed beamformer design.

Table II: Simulation parameters for an MUMC system

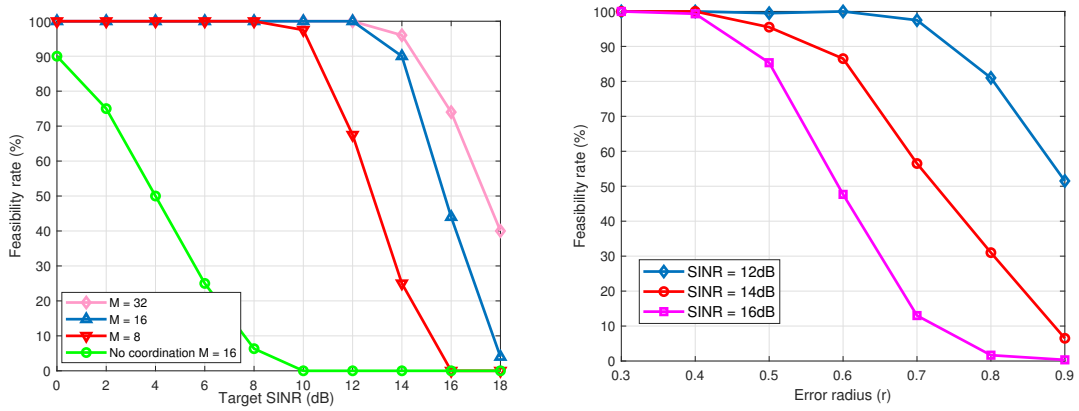
Parameters	**Values**
No. of Cells	2
TAs per BS	{8,16,32}
RAs per user	1
Users per cell	2
Noise power, σ_n^2	0.1
Power priority weight, β_n	1
Angular grid point, G	64
Stopping parameter, υ	10^{-5}
EM iterations, η_{max}	50

Figures 3a and 3b plot the feasibility rate versus the target SINR (γ) parameterized by the number of BS antennas M and versus the error radius r parameterized by the target SINR, respectively, for the SDR-based centralized beamformer design explicitly, the feasibility rate is defined as the percentage of the number of successful computations of the corresponding quantities, namely of the average transmit power and beamformer weights using the algorithms proposed in (8), (17), (34) and (38) as a fraction of the total number of channel realizations. It can be observed from Fig. 3a that as expected, the feasibility rate decreases for higher target SINR values. This is due to the fact that the increased SINR requirements of the users makes it difficult to obtain a feasible solution satisfying the SINR constraint in (5). In addition, the proposed coordinated beamformer design has substantially improved feasibility beyond that of the robust beamformer design operating without coordination. This explicitly demonstrates the enhanced capability of coordinated beamforming as a benefit of the improved degrees of freedom provided by multiple BSs. Future research should be dedicated to deriving an upper bound on the feasibility of the proposed algorithms. Interestingly, it can also be seen that the feasibility rate increases upon increasing the number



(a) (b)

Figure 3: Centralized design: (a) Feasibility rate versus target SINR for $r = 0.6$; (b) Feasibility rate versus error radius for $M = 16$.

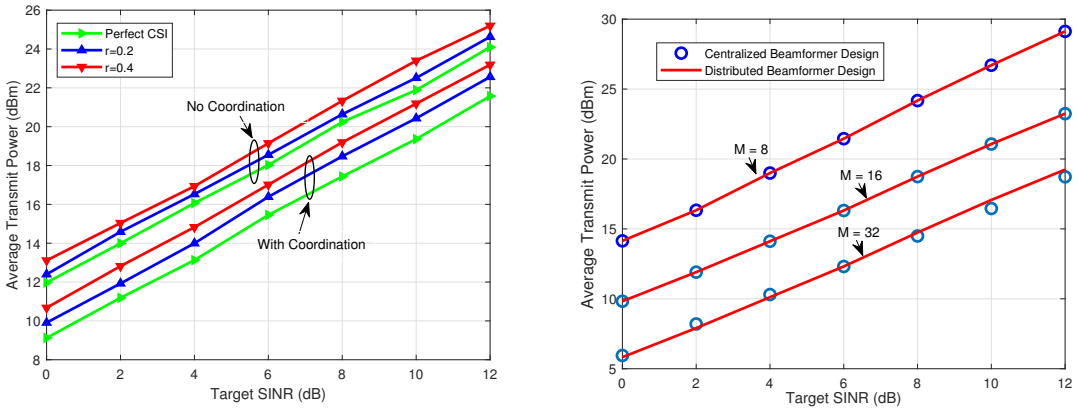


(a) (b)

Figure 4: Distributed design: (a) Feasibility rate versus target SINR for $r = 0.6$; (b) Feasibility rate versus error radius for $M = 16$.

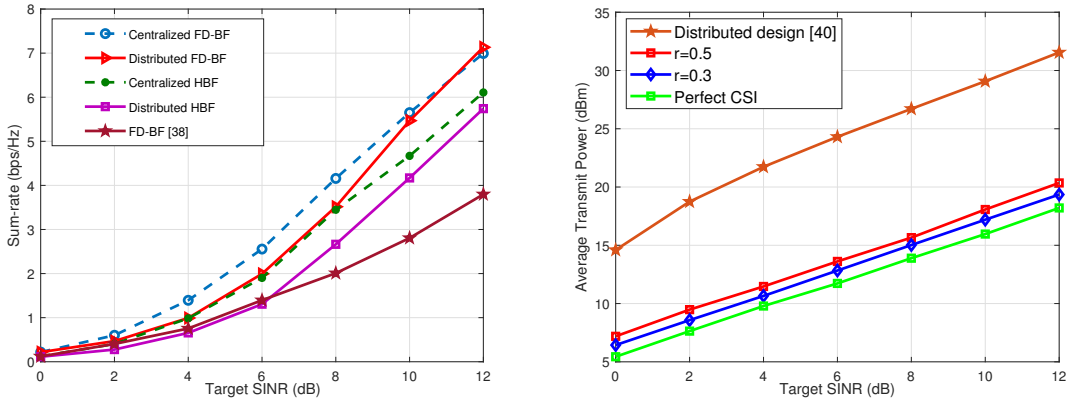
of BS antennas, which can be attributed to the fact that the MUMC MIMO system and the proposed design technique is able to efficiently exploit the resultant increased array gain. It can also be seen that the feasibility rate of the proposed algorithm decreases upon increasing the CSI uncertainty error radius r, as shown in Fig. 3b, since the worst-case SINR constraint is challenging to satisfy for large values of r. However, the proposed algorithm still achieves a reasonably high feasibility rate in the presence of large CSI errors. Fig. 4a and 4b plot the feasibility rate versus the target SINR and error radius, respectively for the ADMM-based distributed design. The plots exhibit a similar trend to that seen for the centralized design. Additionally, it can be observed from Fig. 3 and 4 that the feasibility rates for the centralized and distributed solutions are almost identical. This implies that the distributed beamformer design is as efficient as the centralized one, despite dispensing with global CSI knowledge, which makes it ideally suited for practical implementation.

Fig. 5a shows the power efficiency of the beamformer determined using the centralized SDR method in contrast to the scenario having no coordination. One can observe that the average transmit power increases upon increasing the target SINR, since a high target SINR requires high transmit power to satisfy the SINR constraint. It can also be seen that the multi-cell coordinated hybrid beamformer (MCCH-BF) design has a higher power efficiency than the multi-cell no coordination hybrid beamformer (MCNC-BF) design, described in Appendix A, at the same target SINR requirements. The average transmit power of the MCCH-BF design is seen to be approximately 3 dBm lower than that of the MCNC-BF design. The transmit power reduction can be attributed to the precise knowledge of the interference power arising from the neighbouring cells in the MUMC mmWave system, which is exploited by the MCCH-BF design. On the other hand, the MCNC-BF method does not rely on any prior knowledge of the interference power, hence requiring a high



(a) (b)

Figure 5: (a) Average transmit power versus target SINR for $M = 16$; (b) Average transmit power versus target SINR for centralized and distributed beamformer design for $r = 0.3$.



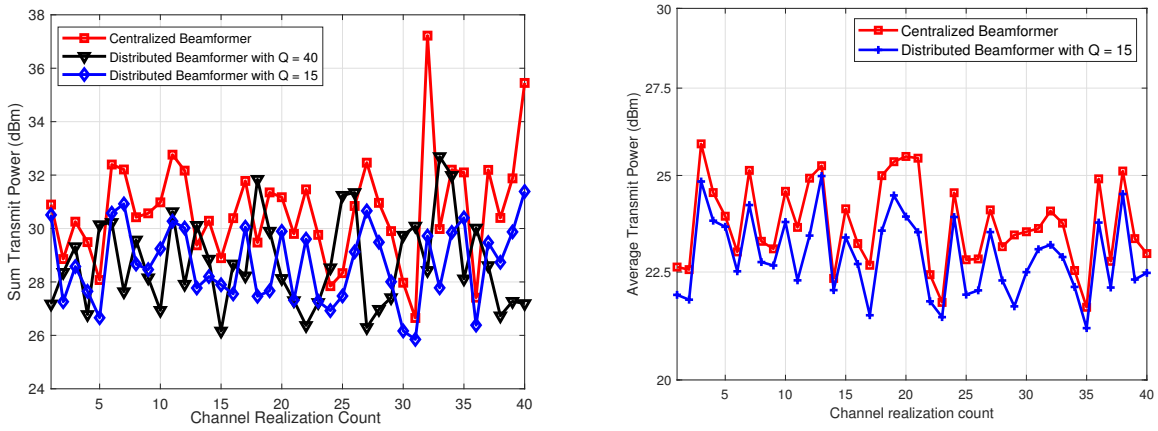
(a) (b)

Figure 6: (a) Sum-rate versus target SINR for $r = 0.3$; (b) Average transmit power versus target SINR

transmit power. Although the average transmit power is seen to increase with the uncertainty error radius r, which seems counter-intuitive, this is justified, since the BSs require a high transmit power in the presence of high CSI errors to meet the QoS constraint. It is very close to that of the perfect CSI scenario, which illustrates the efficiency of the proposed designs.

Fig. 5b shows the power-efficiency of the proposed SDR-based centralized and ADMM-based distributed beamformer designs versus the target SINR for different values of the number of transmit antennas M. As discussed earlier, the transmit power increases upon increasing the target SINR values. However, it can be observed that the transmit power required decreases upon increasing the number of TAs at the BS, thanks to the increased array gain of larger antenna arrays. This further justifies the importance of a large number of antennas in the mmWave regime for improved power efficiency. Furthermore, it is also important to note that the average transmit powers of the centralized and distributed designs are close to each other. This is due to the fact that the distributed beamformer design achieves a performance similar to that of the centralized design in fewer iterations, as also demonstrated in Fig. 7a.

Fig. 6a compares the sum-rate of the BL-based hybrid beamformer design to that of the ideal fully digital beamformer. The centralized FD-BF represents the fully-digital centralized beamformer obtained in Section II, whereas the centralized HBF denotes its corresponding hybrid beamformer counterpart derived via Algorithm-1 of Section II. Similarly, the distributed FD-BF and the distributed HBF represent the fully-digital and hybrid designs, respectively. It can be observed that the BL-based design achieves a performance close to that of the fully digital design, despite having a significantly lower number of RF chains. This is attributed to the fact that the mmWave MIMO channel has a significantly lower number of multipath components, which is readily exploited by our hybrid transceiver design. Therefore, the ideal fully digital beamformer can be tightly



(a) (b)

Figure 7: Sum power comparison between SDR-based centralized and ADMM-based distributed beamformer design: (a) Equal target SINR values for each user; (b) Different target SINR values for each user.

approximated using only a few transmit array response vectors. Furthermore, we have also compared the proposed algorithms to the fully digital beamformer of [38] in Fig. 6a. It can be observed that the proposed methods exhibit a higher sum-rate than that of [38] as the latter fails to efficiently mitigate the inter-cell interference. Fig. 6b shows a performance comparison of the proposed distributed design method to that of [40]. Observe in Fig. 6b that the distributed design of [40] requires a higher transmit power than the proposed method. This is due to the fact that the algorithm of [40] does not share the ICI information amongst the distributed antenna units, which in turn causes severe interference. Fig. 7a illustrates the sum transmit power of the ADMM-based distributed beamformer for both $Q = 15$ and $Q = 40$ iterations considering equal target SINR values for all the users i.e., $\gamma_{nu} = 10$dB $\forall n, u$. Observe that $Q = 15$ iterations are sufficient for approaching the performance of the centralized method. Hence, the proposed distributed design converges to the optimal beamformer solution in a few iterations, while requiring only local CSI along with limited information exchange, which are the key advantages of the distributed design. Furthermore, Fig. 7b shows the robustness of the proposed distributed beamformer design at the target SINR values of $\gamma_{11} = 8$dB, $\gamma_{12} = 10$dB, $\gamma_{21} = 6$dB and $\gamma_{22} = 12$dB for the users. Observed that the proposed distributed beamformer design once again achieves a performance close to the centralized beamformer in $Q = 15$ iterations.

VII. Conclusions

Robust coordinated hybrid beamforming solutions have been designed for MUMC mmWave MIMO systems. First, centralized beamformers were obtained using the SDR and BL-based hybrid TPC decomposition techniques. Next, ADMM-based distributed beamformers were developed, using only the local CSI available at each BS for the users of its own cell. Subsequently, robust centralized and ADMM-based distributed designs were developed for practical scenarios having imperfect CSI. Our simulation results demonstrate that the proposed centralized and distributed beamformer designs yield an improved performance in comparison to a beamformer having no coordination. It was also seen that the proposed distributed design has a high power and spectral-efficiency, which is close to that of the centralized solution. The distributed design framework offers comparable performance to that of the centralized one without a high signalling overhead, and it is thus ideally suited for practical implementation.

Future research may explore different variants of the ADMM algorithm to obtain faster convergence and a lower backhaul signaling overhead for designing the distributed hybrid beamformer. Furthermore, another interesting aspect can potentially be to design an asynchronous distributed beamformer that is robust to the transmission delay and packet losses of the backhaul network.

Appendix A
BL-based hybrid TPC design

Let the concatenated fully digital TPC corresponding to all users at the BS_n be constructed as $\mathbf{F}_{n,\mathrm{opt}} = [\mathbf{f}_{n1,\mathrm{opt}}, \mathbf{f}_{n2,\mathrm{opt}}, \ldots, \mathbf{f}_{nU,\mathrm{opt}}] \in \mathbb{C}^{M\times U}$. The approximation problem of jointly designing the baseband and RF TPCs $\mathbf{F}_{\mathrm{BB},n} \in \mathbb{C}^{N_{\mathrm{RF},n}\times U}$ and $\mathbf{F}_{\mathrm{RF},n} \in \mathbb{C}^{M\times N_{\mathrm{RF},n}}$, respectively, at BS_n, can be formulated as

$$\left(\mathbf{F}^*_{\mathrm{BB},n}, \mathbf{F}^*_{\mathrm{RF},n}\right) = \underset{\mathbf{F}^*_{\mathrm{BB},n}, \mathbf{F}^*_{\mathrm{RF},n}}{\arg\min} \|\mathbf{F}_{n,\mathrm{opt}} - \mathbf{F}_{\mathrm{RF},n}\mathbf{F}_{\mathrm{BB},n}\|^2_{\mathrm{F}}$$
$$\text{s.t.} \qquad |\mathbf{F}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{M}}. \qquad (40)$$

The problem above can be simplified as follows. Let the quantized transmit array response dictionary matrix $\mathbf{A}_T$ be constructed as

$$\mathbf{A}_T \triangleq [\mathbf{a}_T(\phi_1), \mathbf{a}_T(\phi_2), \cdots, \mathbf{a}_T(\phi_G)] \in \mathbb{C}^{M\times G},$$

where the AoD-set $\Phi_T = \{\phi_g,\ \forall 1 \leq g \leq G\}$ spans the angular range $[0,\ \pi]$ with $\cos(\phi_g) = \frac{2}{G}(g-1) - 1$, and G represents the grid size [7]. Note that since the elements of the dictionary matrix of the transmit array response satisfy the constant magnitude constraint (40), the columns of the RF TPC can be suitably chosen from the dictionary matrix $\mathbf{A}_T$. Therefore, the equivalent coordinated hybrid TPC design problem of the MUMC mmWave system can be formulated as

$$\begin{aligned} &\arg\min_{\tilde{\mathbf{F}}_{\mathrm{BB},n}} \left\|\mathbf{F}_{n,\mathrm{opt}} - \mathbf{A}_T\tilde{\mathbf{F}}_{\mathrm{BB},n}\right\|_{\mathrm{F}}^2 \\ &\text{s.t. } \left\|\mathrm{diag}\left(\tilde{\mathbf{F}}_{\mathrm{BB},n}\tilde{\mathbf{F}}_{\mathrm{BB},n}^H\right)\right\|_0 = N_{\mathrm{RF},n}, \end{aligned} \tag{41}$$

where $\tilde{\mathbf{F}}_{\mathrm{BB},n} \in \mathbb{C}^{G\times U}$ denotes the intermediate baseband TPC. The constraint in (41) arises due to the fact that $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ may only contain $N_{\mathrm{RF},n}$ non-zero rows, due to the presence of only $N_{\mathrm{RF},n}$ RF chains. The BL-based TPC design procedure begins by assigning a parameterized Gaussian prior to the matrix $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ as shown below

$$\begin{aligned} p\left(\tilde{\mathbf{F}}_{\mathrm{BB},n};\boldsymbol{\Gamma}\right) &= \prod_{i=1}^{G} p\left(\tilde{\mathbf{F}}_{\mathrm{BB},n}(i,:);\gamma_i\right) \\ &= \prod_{i=1}^{G} \frac{1}{\pi\gamma_i}\exp\left(-\frac{\left\|\tilde{\mathbf{F}}_{\mathrm{BB},n}(i,:)\right\|^2}{\gamma_i}\right), \end{aligned} \tag{42}$$

where $\boldsymbol{\Gamma} = \mathrm{diag}(\gamma_1,\ldots,\gamma_G) \in \mathbb{R}^{G\times G}$ denotes the diagonal matrix of hyperparameters. Based on the prior assignment above, it can be readily observed that the ith row of the matrix $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ is assigned the hyperparameter γ_i, which enforces row sparsity, as seen in the constraint (41). The posterior density of the matrix $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ can be formulated as $p\left(\tilde{\mathbf{F}}_{\mathrm{BB},n} \mid \mathbf{F}_{n,\mathrm{opt}};\boldsymbol{\Gamma}\right) \sim \mathcal{CN}(\boldsymbol{\mathcal{F}},\boldsymbol{\Sigma})$, where the *a posteriori* matrix $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{G\times U}$ and the associated covariance matrix $\boldsymbol{\Gamma} \in \mathbb{C}^{G\times G}$ are expressed as

$$\boldsymbol{\mathcal{F}} = \frac{1}{\sigma_e^2}\boldsymbol{\Sigma}\mathbf{A}_T^H\mathbf{F}_{\mathrm{opt},n} \quad \text{and} \quad \boldsymbol{\Sigma} = \left(\frac{1}{\sigma_e^2}\mathbf{A}_T^H\mathbf{A}_T + \boldsymbol{\Gamma}^{-1}\right)^{-1}. \tag{43}$$

The quantity σ_e^2 above denotes the variance of the approximation error. It can be observed that the minimum mean square error (MMSE) estimate, i.e. the *a posteriori* mean $\boldsymbol{\mathcal{F}}$, depends on the hyperparameter matrix $\boldsymbol{\Gamma}$. Furthermore, as $\gamma_i \to 0$, the ith row of the matrix $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ denoted by $\tilde{\mathbf{F}}_{\mathrm{BB},n}(i,:) \to 0$, tends to $\mathbf{0}$. Thus, the estimation of $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ is equivalent to that of the associated hyperparameter vector $\boldsymbol{\gamma} = [\gamma_1,\ldots,\gamma_G]^T$. Since the pertinent optimization maximizing the likelihood estimate of $\boldsymbol{\gamma}$ by maximizing the Bayesian evidence $p(\mathbf{F}_{\mathrm{opt}};\boldsymbol{\Gamma})$ is non-convex, it is difficult to solve. Hence, the BL aided hybrid TPC design relies on a low-complexity iterative expectation-maximization (EM) framework.

To begin with, let $\hat{\boldsymbol{\Gamma}}^{(k-1)}$ denote the estimate of the hyperparameter vector $\boldsymbol{\Gamma}$ obtained from the $(k-1)$th iteration. The EM procedure involves two key steps. In the first step, known as the expectation step (E-step), the average log-likelihood function $\mathcal{L}\left(\boldsymbol{\Gamma} \mid \hat{\boldsymbol{\Gamma}}^{(k-1)}\right)$ of the hyperparameters is evaluated as

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

The subsequent maximization step (M-step), maximizes the average log-likelihood with respect to the hyperparameter vector $\boldsymbol{\gamma}$ as

$$\begin{aligned} &\hat{\boldsymbol{\gamma}}^{(k)} = \\ &\arg\max_{\boldsymbol{\gamma}} \mathbb{E}\left\{\log p\left(\mathbf{F}_{n,\mathrm{opt}} \mid \tilde{\mathbf{F}}_{\mathrm{BB},n}\right) + \log p\left(\tilde{\mathbf{F}}_{\mathrm{BB},n};\boldsymbol{\Gamma}\right)\right\}. \end{aligned} \tag{44}$$

It can be readily observed that the first term within the expectation operator of Eq. (44) is independent of the hyperparameter $\boldsymbol{\gamma}$, and it can thus be ignored in the subsequent M-step. Therefore, the equivalent optimization problem is given by

$$\begin{aligned} \hat{\boldsymbol{\gamma}}^{(k)} &= \arg\max_{\boldsymbol{\gamma}} \mathbb{E}_{\tilde{\mathbf{F}}_{\mathrm{BB},n}|\mathbf{F}_{n,\mathrm{opt}},\hat{\boldsymbol{\Gamma}}^{(k-1)}}\left\{\log p\left(\tilde{\mathbf{F}}_{\mathrm{BB},n};\boldsymbol{\Gamma}\right)\right\} \\ &= \arg\max_{\boldsymbol{\gamma}} \sum_{i=1}^{G}\left[-\log(\gamma_i) - \frac{\mathbb{E}\left(\left\|\tilde{\mathbf{F}}_{\mathrm{BB},n}(i,:)\right\|_2^2\right)}{\gamma_i}\right], \\ &= \arg\max_{\boldsymbol{\gamma}} \sum_{i=1}^{G} -\log(\gamma_i) - \frac{\left\|\boldsymbol{\mathcal{F}}^{(k)}(i,:)\right\|^2 + U\boldsymbol{\Sigma}_{(i,i)}^{(k)}}{\gamma_i}, \end{aligned} \tag{45}$$

where $\boldsymbol{\mathcal{F}}^{(k)}$ and $\boldsymbol{\Sigma}^{(k)}$ are obtained from (43) by setting $\boldsymbol{\Gamma} = \hat{\boldsymbol{\Gamma}}^{(k-1)}$. The optimal value of $\hat{\gamma}_i^{(k)}$ can be evaluated by calculating the gradient of the objective function in (45) with respect to $\boldsymbol{\gamma}$ and setting it to zero. This yields the update equation for each hyperparameter as

$$\hat{\gamma}_i^{(k)} = \frac{1}{U}\left\|\boldsymbol{\mathcal{F}}^{(k)}(:,i)\right\|^2 + \boldsymbol{\Sigma}_{(i,i)}^{(k)}.$$

After the convergence of the EM procedure, the RF and baseband TPCs are obtained as follows. Let $\mathcal{S}$ denote the set of indices of the $N_{\mathrm{RF},n}$ hyperparameters having the largest magnitude. The concatenated optimal baseband precoder matrix $\mathbf{F}_{\mathrm{BB},n}^*$ can be obtained from $\tilde{\mathbf{F}}_{\mathrm{BB},n}$ as

$$\mathbf{F}_{\mathrm{BB},n}^* = \tilde{\mathbf{F}}_{\mathrm{BB},n}(\mathcal{S},:). \tag{46}$$

Similarly, the optimal RF TPC $\mathbf{F}_{\mathrm{RF},n}^*$ can be extracted from the transmit array response dictionary matrix $\mathbf{A}_T$ as

$$\mathbf{F}_{\mathrm{RF},n}^* = \mathbf{A}_T(:,\mathcal{S}). \tag{47}$$

APPENDIX B
MULTI-CELL NO-COORDINATION BEAMFORMER (MCNC-BF)

The beamformer design procedure relying on no coordination among the BSs for the general scenario having realistic CSI uncertainty is presented below. In such a scenario, BS_n actively considers the interferences arriving from all other cells as noise and designs the beamformer by solving the following optimization problem:

$$\min_{\{\mathbf{F}_{\mathrm{RF},n},\mathbf{F}_{\mathrm{BB},nu}\}} \sum_{u=1}^{U} \|\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\|^2 \tag{48a}$$

$$\text{s.t.} \quad \frac{\left|\mathbf{h}_{nnu}^H\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\right|^2}{\sum\limits_{i\neq u}^{U}|\mathbf{h}_{nnu}^H\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},ni}|^2+\sum\limits_{m\neq n}^{M}\zeta_{mnu}+\sigma_{nu}^2} \geq \gamma_{nu},$$

$$\forall\ \mathbf{\Delta}_{nnu}^H\mathbf{Q}_{nnu}\mathbf{\Delta}_{nnu} \leq 1, \forall\ u, \tag{48b}$$

$$\sum_{u=1}^{U}\left|\mathbf{h}_{nml}^H\mathbf{F}_{\mathrm{RF},n}\mathbf{f}_{\mathrm{BB},nu}\right|^2 \leq \zeta_{nml},$$

$$\forall\ \mathbf{\Delta}_{nml}^H\mathbf{Q}_{nml}\mathbf{\Delta}_{nml} \leq 1, l = 1,2,..,U, m \neq n, \tag{48c}$$

$$|\mathbf{F}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{M}}, \forall i,j. \tag{48d}$$

The quantity $\mathbf{h}_{nmu} = \hat{\mathbf{h}}_{nmu} + \mathbf{\Delta}_{nmu}$ denotes the channel vector contaminated by CSI error and $\zeta_{mnu} > 0$ represents the maximum interference tolerance arriving from BS_m to MS_{nu}. It can be observed that when there is no coordination among the BSs, each BS solves its beamformer design problem in a decentralized fashion by considering the interferences impinging from different cells as noise. Following the procedure described in Section IV, the corresponding problem of designing the fully digital beamformer can be recast as

$$\min_{\{\mathbf{F}_{nu}\}_{u=1}^{U}} \left(\sum_{u=1}^{U}\mathrm{Tr}(\mathbf{F}_{nu})\right) \tag{49a}$$

$$\text{s.t.} \left(\hat{\mathbf{h}}_{nnu}^H+\mathbf{\Delta}_{nnu}^H\right)\left(\frac{1}{\gamma_{nu}}\mathbf{F}_{nu}-\sum_{i\neq u}^{U}\mathbf{F}_{ni}\right)(\hat{\mathbf{h}}_{nnu}+\mathbf{\Delta}_{nnu})$$

$$\geq \sum_{m\neq n}^{N_c}\zeta_{mnu}+\sigma_{nu}^2, \forall\ \mathbf{\Delta}_{nnu}^H\mathbf{Q}_{nnu}\mathbf{\Delta}_{nnu} \leq 1, \quad \forall u, \tag{49b}$$

$$\sum_{i=1}^{U}(\hat{\mathbf{h}}_{nml}^H + \mathbf{\Delta}_{nml}^H)\mathbf{F}_{ni}(\hat{\mathbf{h}}_{nml} + \mathbf{\Delta}_{nml}) \leq \zeta_{nml}$$

$$\forall\ \mathbf{\Delta}_{nml}^H\mathbf{Q}_{nml}\mathbf{\Delta}_{nml} \leq 1, \quad \forall u, m \neq n, \tag{49c}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}. \tag{49d}$$

Upon applying the S-lemma described in Section IV, the beamformer design problem corresponding to BS_n can be formulated as

$$\min_{\{\mathbf{F}_{nu}\}_{u=1}^{U}} \sum_{u=1}^{U}\mathrm{Tr}(\mathbf{F}_{nu}) \tag{50a}$$

$$\text{s.t.} \ \mathbf{\Phi}_{nu}\left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, \{\zeta_{mnu}\}_{m\neq n}^{N_c}, \lambda_{nnu}\right) \succeq \mathbf{0}, \forall u, \tag{50b}$$

$$\mathbf{\Psi}_{nmu}\left(\{\mathbf{F}_{ni}\}_{i=1}^{U}, \zeta_{nmu}, \lambda_{nmu}\right) \succeq \mathbf{0}, \forall u, m \neq n, \tag{50c}$$

$$\mathbf{F}_{nu} \succeq \mathbf{0}, \quad \lambda_{nmu} \geq 0, \quad \forall u, \tag{50d}$$

which can be readily solved using a standard convex solver.

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] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1850–1859, 2012.

[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] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?" *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, 2014.

[5] L. Li, X. Niu, Y. Chai, L. Chen, T. Zhang, D. Cheng, H. Xia, J. Wang, T. Cui, and X. You, "The path to 5G: mmwave aspects," *Journal of Communications and Information Networks*, vol. 1, no. 2, pp. 1–18, 2016.

[6] L. Bai, L. Zhu, X. Zhang, W. Zhang, and Q. Yu, "Multi-satellite relay transmission in 5G: Concepts, techniques, and challenges," *IEEE Network*, vol. 32, no. 5, pp. 38–44, 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] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE communications magazine*, vol. 49, no. 6, pp. 101–107, 2011.

[10] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[11] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, 2014.

[12] S. A. Jafar, G. J. Foschini, and A. J. Goldsmith, "Phantomnet: Exploring optimal multicellular multiple antenna systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 5, pp. 1–14, 2004.

[13] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No. 01CH37202)*, vol. 3. IEEE, 2001, pp. 1745–1749.

[14] D. J. Love, R. W. Heath, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, 2008.

[15] P. Marsch and G. Fettweis, "On multicell cooperative transmission in backhaul-constrained cellular systems," *Annals of Telecommunications-Annales Des téléCommunications*, vol. 63, no. 5, pp. 253–269, 2008.

[16] S. Yang, T. Lv, R. G. Maunder, and L. Hanzo, "Distributed probabilistic-data-association-based soft reception employing base station cooperation in MIMO-aided multiuser multicell systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3532–3538, 2011.

[17] L.-N. Tran, M. F. Hanif, A. Tolli, and M. Juntti, "Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 872–875, 2012.

[18] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 235–251, 2010.

[19] 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.

[20] 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.

[21] M. Li, Z. Wang, X. Tian, and Q. Liu, "Joint hybrid precoder and combiner design for multi-stream transmission in mmwave MIMO systems," *IET Communications*, vol. 11, no. 17, pp. 2596–2604, 2017.

[22] 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.

[23] D. H. Nguyen, L. B. Le, and T. Le-Ngoc, "Hybrid MMSE precoding for mmwave multiuser MIMO systems," in *2016 IEEE international conference on communications (ICC)*. IEEE, 2016, pp. 1–6.

[24] V. Venkateswaran and A.-J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4131–4143, 2010.

[25] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in *2013 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2013, pp. 3679–3684.

[26] 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.

[27] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, 2012.

[28] S. Schwarz and M. Rupp, "Exploring coordinated multipoint beamforming strategies for 5G cellular," *IEEE Access*, vol. 2, pp. 930–946, 2014.

[29] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1711–1721, 2007.

[30] W.-Y. Chen, B.-S. Chen, and W.-T. Chen, "Multiobjective beamforming power control for robust SINR target tracking and power efficiency in multicell MU-MIMO wireless system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6200–6214, 2020.

[31] X. Xie, H. Yang, and A. V. Vasilakos, "Robust transceiver design based on interference alignment for multi-user multi-cell MIMO networks with channel uncertainty," *IEEE Access*, vol. 5, pp. 5121–5134, 2017.

[32] D. Ponukumati, F. Gao, M. Bode, and X. Liao, "Multicell downlink beamforming with imperfect channel knowledge at both transceiver sides," *IEEE communications letters*, vol. 15, no. 10, pp. 1075–1077, 2011.

[33] B. Liu, Y. Cheng, and Q. Zhou, "Robust rank-two beamforming for multicell multigroup multicast," *IET Communications*, vol. 10, no. 3, pp. 283–291, 2016.

[34] A. Shaverdian and M. R. Nakhai, "Robust distributed beamforming with interference coordination in downlink cellular networks," *IEEE transactions on communications*, vol. 62, no. 7, pp. 2411–2421, 2014.

[35] M. Medra and T. N. Davidson, "Low-complexity robust multi-cell MISO downlink precoder design," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 3546–3550.

[36] 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.

[37] 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.

[38] 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.

[39] G. R. MacCartney and T. S. Rappaport, "Millimeter-wave base station diversity for 5G coordinated multipoint (CoMP) applications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3395–3410, 2019.

[40] 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. 12 394–12 397, 2018.

[41] L. Zhang, L. Gui, X. Mo, and M. Qi, "Interference subspace alignment-based precoding design for multi-cell multi-user systems," *IEEE Transactions on Broadcasting*, vol. 67, no. 1, pp. 106–118, 2020.

[42] D. Castanheira, P. Lopes, A. Silva, and A. Gameiro, "Hybrid beamforming designs for massive MIMO millimeter-wave heterogeneous systems," *IEEE Access*, vol. 5, pp. 21 806–21 817, 2017.

[43] R. Mai and T. Le-Ngoc, "Nonlinear hybrid precoding for coordinated multi-cell massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2459–2471, 2019.

[44] Z.-Q. Luo, T.-H. Chang, D. Palomar, and Y. Eldar, "SDP relaxation of homogeneous quadratic optimization: approximation," *Convex Optimization in Signal Processing and Communications*, p. 117, 2010.

[45] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, 2010.

[46] 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.

[47] M. Majumder, H. Saxena, S. Srivastava, and A. K. Jagannatham, "Optimal bit allocation-based hybrid precoder-combiner design techniques for mmwave MIMO-OFDM systems," *IEEE Access*, vol. 9, pp. 54 109–54 125, 2021.

[48] D. P. Bertsekas, "Nonlinear programming," *Journal of the Operational Research Society*, vol. 48, no. 3, pp. 334–334, 1997.

[49] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[50] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.

[51] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, no. 5, pp. 303–320, 1969.

[52] M. Jafri, A. Anand, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Technical report: Robust distributed hybrid beamforming in coordinated multi-user multi-cell mmWave MIMO systems relying on imperfect CSI," *IIT Kanpur, Tech. Rep.*, 2022 [Online]. [Online]. Available: https://www.iitk.ac.in/mwn/documents/MWNLAB_TR_MCMU.pdf

[53] G.-W. Hsu, B. Liu, H.-H. Wang, and H.-J. Su, "Joint beamforming for multicell multigroup multicast with per-cell power constraints," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4044–4058, 2016.

[54] H. Pennanen, A. Tölli, and M. Latva-aho, "Decentralized robust beamforming for coordinated multi-cell MISO networks," *IEEE Signal Processing Letters*, vol. 21, no. 3, pp. 334–338, 2014.