Hybrid Precoding in Cooperative Millimeter Wave Networks

Chao Fang, Behrooz Makki, Senior Member, IEEE, Jingya Li and Tommy Svensson, Senior Member, IEEE,

Abstract

In this paper, we study the performance of cooperative millimeter wave (mmWave) networks using various precoding schemes. In our proposed method, multiple base stations (BSs) can jointly serve users and the BSs associated with no users are put into sleep mode to reduce the power consumption. Considering both the hardware power and the RF transmit power of the BSs, we propose a hybrid precoding scheme which minimizes the sum power consumption of the BSs subject to per-user spectral efficiency constraints as well as the per BS peak power constraints. In order to obtain tractable and near-optimal hybrid precoding solutions, we reformulate the analog precoding part as an equal-gain transmission problem and the digital precoding part as a relaxed convex semidefinite program. We present the results for both fully- and partially-connected hybrid precoding (FHP and PHP, respectively) schemes and show that, depending on the parameter settings, the power consumption of the PHP may be dominated by the RF transmit power and it may result in a larger power consumption than the FHP. For the cases with 2 BSs and 4 users, implementation of the FHP and the PHP in cooperative networks reduces the required RF transmit power, compared to the case in a non-cooperative network, by 71% and 65%, respectively.

Index Terms

Millimeter Wave, hybrid beamforming, cooperative networks, energy efficiency, user association.

C. Fang and T. Svensson are with the Dept. of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden (e-mail: {fchao, tommy.svensson}@chalmers.se).

B. Makki and J. Li are with Ericsson Research, Ericsson AB, Gothenburg, Sweden (e-mail: {behrooz.makki, jingya.li}@ericsson.com).
I. INTRODUCTION

5G networks are seen as the next revolution in wireless communications, promising high bandwidth, high energy efficiency, wide coverage, high reliability and low latency [1]. Particularly, compared to 4G, future networks are expected to support a wide range of use cases, including enhanced mobile broadband with $10^3$ times higher user data rate, massive machine type communications with $10^4$ times more connected low-cost devices, and critical machine type communications with ultra reliability and low latency in the order of milliseconds [2]. In order to meet the requirements, different key technologies are currently being considered in 5G, among which network densification [3], millimeter wave (mmWave) and massive multiple-input multiple-output (MIMO) [4] are of particular interest.

Due to the ever-growing mobile data demand, conventional networks operating on the sub-6 GHz spectrum are becoming overcrowded. On the other hand, the mmWave spectrum, in the range of 6-100 GHz, remains largely unused. Therefore, one of the key features of 5G is to utilize the large bandwidth at carrier frequency $> 6$ GHz and to provide gigabytes-per-second data rate. However, one of the main drawbacks of mmWave frequency signals is the high path loss, which may result in a shorter communication range. With the roll out of small BSs, which only cover an area of few hundreds of meters [3], [5], [6] and the adoption of large antenna arrays that can provide high array gain and directional transmissions, mmWave-enabled networks can become practical. Other challenges in mmWave cellular systems include channel modeling and precoder design. Channel measurements have been conducted on various carrier frequencies in various scenarios [7], [8], in order to validate simulated data with field measurements and identify the requirements of key system design parameters. As mmWave signals are sensible to blocking, the difference between line-of-sight (LOS) and non-line-of-sight (NLOS) paths need to be considered in the path loss estimation. Environment-dependent path loss exponents and LOS probabilities were proposed in recent studies to improve the accuracy of the path loss models for mmWave communications [9]–[12].

In order to compensate the high path loss and maintain a sufficient link budget, precoding with large number of antennas is essential for mmWave BSs to serve multiple users with multiple data streams simultaneously. However, the conventional fully digital precoding (FDP) architecture, which requires a complete radio-frequency (RF) chain and a digital-to-analog converter (DAC) or analog-to-digital converter (ADC) per antenna, may be infeasible as the cost of hardware
increases rapidly with the number of antennas. To reduce the cost as well as the complexity of the hardware, the hybrid beamforming architectures, which use less RF chains and DACs/ADCs than the number of antennas, have been proposed [13]–[24].

Depending on the extent of connections between the RF chains and antennas, the hybrid precoding architecture can be categorized as the fully-connected hybrid precoding (FHP) and the partially-connected hybrid precoding (PHP) architecture. The FHP architecture requires each antenna to be connected to all RF chains via phase shifters (PSs) and has been shown to achieve a spectral efficiency close to that of the FDP in a single BS setup [13]–[16]. Even though the FHP reduces the power consumption by using less RF chains, it may still need a large number of PSs as the number of antennas grows, which poses other challenges such as high PS power consumption, insertion loss, and wiring complexity [17]. For this reason, by allowing each RF chain to be connected to a part of the antennas, the PHP architecture is proposed to further reduce the number of PSs, and various PHP schemes have been studied to optimize the spectral efficiency [17]–[19] and achieve better energy efficiency than the FHP architecture [20]–[24].

In multi-cell scenarios, coordination between BSs is required to alleviate inter-BS interference and reduce service outage. If multi-user channel state information (CSI) is shared among the BSs, power allocation and beam directions can be coordinated for interference management. Furthermore, if user data is available to all BSs, a full cooperation allows a user to receive multiple data streams from different BSs. In this way, the precoders can be jointly designed utilizing the global CSI information of the network. For the FDP, many cooperation schemes have been proposed to optimize the system performance in multi-cell networks [25]–[28]. Also, coordinated multi-cell systems based on hybrid precoding have been studied in terms of spectral efficiency optimization [29], [30]. Considering predefined beam patterns and serving each user from only one BS, analog precoders have been jointly selected to maximize the users’ data rates [29]. In [30], the performance of interference coordination based on the signal-to-leakage-plus-noise-ratio and regularized zero-forcing hybrid precoding methods were studied in terms of spectral efficiency. In addition, a measurement-campaign-based study for cooperative mmWave BSs is shown to significantly reduce the outage and improve spectral efficiency [31].

Despite the advantage of multi-cell cooperation in terms of high spectral efficiency, there is a lack of studies on optimizing the energy efficiency of hybrid-precoding-based cooperative multi-cell mmWave networks. Studying such problems is of interest, because one of the main motivations for hybrid precoding is to reduce the power consumption without much performance
loss. In order to accurately quantify how power savings of specific architectures improve the energy efficiency of mmWave systems, power models that consist of hardware and system-level power consumption have been used [21], [32], [33].

In this paper, we propose a hybrid precoding algorithm for minimizing the sum power consumption of the BSs in a cooperative mmWave network subject to per-user rate constraints and per-BS maximum power constraints. The contributions of this paper are as follows. We first present a fairly realistic power model including the BS sleep mode which saves power when no users are associated with a BS. The power model takes into account both the RF transmit power and the power consumption of hardware components such as the RF chains, the PSs and the DACs. We further propose to decouple the joint hybrid precoding problem into an analog precoding problem with a solution based on equal gain transmission and a digital precoding problem which is in the form of a relaxed semidefinite program. For both the FHP and PHP architectures, our proposed analog and digital precoders jointly associate users to the BSs with the optimal BS sleeping strategy and allow multiple streams from different BSs to a single user. Furthermore, we study the value of cooperation in terms of the power consumption, the probability of infeasible solutions and the probability of joint transmission. Finally, we study the effect of different parameters such as the number of antennas and the LOS probability on the network performance.

As opposed to [13]–[24], we study the system performance in cooperative mmWave multicell scenarios and consider a fairly realistic model for the power consumption. Moreover, our optimization problem formulation and the proposed hybrid precoding scheme are different from those considering the FDP [25]–[28]. In [33], we performed initial studies on the performance of hybrid precoding schemes. Compared to [33], our current work presents analytical performance analysis of both the FHP and the PHP in cooperative networks, derives optimality conditions for user association, and considers a more accurate power consumption model. Finally, our results on the effect of cooperation and comparison of different schemes in rate-constrained conditions have not been presented before. The differences in the problem formulation and the power model makes our analytical/simulation results and conclusions completely different from the ones in the literature, e.g., [29], [30], [32].

Simulation results verify that our proposed FHP gives close performance to the FDP, in terms of beam patterns and RF transmit power. As opposed to previous studies which inherently assume that the energy efficiency is improved by using the PHP with reduced number of PSs, we show
that if the power consumption is dominated by the RF transmit power, the power consumption of the PHP may be higher than that of the FHP. Furthermore, simulation results confirm the value of cooperation, as it reduces the sum power consumption of the BSs and the probability of infeasible precoding solutions. For example, when the network changes from 1 BS to 2 BSs with cooperation, for a per-user spectral efficiency of 4 bit/s/Hz and 4 users, the sum RF transmit power is reduced by 71% and 65% for the FHP and the PHP schemes, respectively.

*Notations:* We use bold lower-case letters like $\mathbf{d}$ for vectors and upper-case bold letters like $\mathbf{R}$ for matrices. Then, $\mathbf{R}^T$, $\mathbf{R}^H$, $\mathbf{R}^{(i,j)}$ and $||\mathbf{R}||_F$ denote the transpose, the Hermitian, the $(i,j)$-th entry of $\mathbf{R}$ and the Frobenius norm of $\mathbf{R}$, respectively. $||\mathbf{d}||$ denotes the Euclidean norm and $\text{Tr}(\cdot)$ is the trace of a square matrix. Finally, $\mathbb{C}^n$ represents the set of $n$-tuples of complex numbers represented as column vectors and $\mathbb{C}^{m\times n}$ denotes the set of complex $m \times n$ matrices.

### II. System Model

We consider a multi-cell multi-user mmWave system with $K$ users and $M$ BSs of, possibly, different types which are characterized by different transmission power limits and hardware. The number of antennas and the number of RF chains at a BS $m$, $0 < m \leq M$, are denoted by $N_m$ and $L_m$, respectively. For tractable analysis, a user $k$, $0 < k \leq K$, is assumed to be equipped with a single antenna or an array with each antenna element receiving beams from a certain angle range. The BSs use hybrid precoding, which satisfies $N_m > L_m \geq K$, to transmit to the users. For $N_m = L_m$, the case of the FDP is considered.

In the hybrid beamforming architecture, RF chains are interconnected to antennas via PSs. Depending on the extent of interconnections and the number of PSs needed, the hybrid precoding architectures can be further categorized to the FHP and the PHP. As shown in Fig. 1(a), the FHP architecture requires each RF chain to be connected to all antennas via PSs. Thus, the main drawback is that the number of PSs grows fast with the number of antennas, leading to high hardware power consumption and complexity. On the other hand, the PHP architecture reduces the number of PSs and the interconnections between the RF chains and the antennas by connecting each RF chain to an antenna subarray, within which each antenna is connected to one PS. For the PHP, we additionally assume that $N_m/L_m$, $\forall m$, is an integer such that each RF chain is connected to an antenna subarray and they do not overlap, as shown in Fig. 1(b). Hence, the number of PSs required in the FHP is $L_m N_m$ and it is reduced to $N_m$ in the PHP. For the FDP, because $N_m = L_m$, PSs are not needed to connect the antennas and RF chains.
Fig. 1: Hybrid precoding architectures.
The transmit symbol $x_{k,m}$ to user $k$ at BS $m$ is first precoded by a baseband digital precoder $d_{k,m} \in \mathbb{C}^{L_m \times 1}$ then followed by an analog precoder $R_m \in \mathbb{C}^{N_m \times L_m}$, such that the precoded signals at particular angles have strong power while causing less interference to other users. The digital precoder has full control over both the amplitude and the phase of the signal, while the analog precoder is enabled by PSs and can only change the phase of the signal. For the FHP, the entries in the analog precoder have the constant amplitude constraint $\left| R_m^{(i,j)} \right| = \frac{1}{\sqrt{L_m N_m}}$, $\forall m$. For the PHP, each RF chain is connected to part of the PSs, therefore the analog precoder has the form of a block diagonal matrix given by $R_m = \text{diag}[\hat{r}_{1,m}, \cdots, \hat{r}_{L_m,m}]$. Here, $\hat{r}_{i,m} \in \mathbb{C}^{N_m/L_m \times 1}$, $1 \leq i \leq L_m$, with constraints $\left| \hat{r}_{i,m}^{(j)} \right| = 1/\sqrt{L_m}$, $\forall m$.

A. Spectral Efficiency

To achieve spatial multiplexing and cooperative transmissions, a user $k$ is assumed to be able to receive useful signals from multiple BSs, where the data symbols $x_{k,m}$ from different BSs are assumed to be mutually independent [26], [34]. In this way, the user $k$ implements successive interference cancellation to decode the data streams sequentially. The stronger signals are decoded first and are then subtracted from the received signals to decode the weaker signals. Considering quasi-static channels [35] which are assumed to be known by the BSs, the composite signal received by user $k$ can be written as

$$y_k = \sum_{m=1}^{M} h_{k,m}^H w_{k,m} x_{k,m} + \sum_{m=1}^{M} \sum_{k' \neq k, k'=1}^{K} h_{k,m}^H w_{k',m} x_{k',m} + n_k.$$  \hspace{1cm} (1)

Here, $w_{k,m} = R_m d_{k,m}$ is the overall precoder, $h_{k,m} \in \mathbb{C}^{N_m \times 1}$ is the channel vector, $x_{k,m}$ is the data symbol with $\mathbb{E}[x_{k,m} x_{k,m}^H] = 1$, $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the complex Gaussian noise with variance $\sigma_k^2$, and the second term is the interference. In the case of FDB, the precoder is reduced to $w_{k,m} = d_{k,m}$ by setting $R_m = I$ and $N_m = L_m$.

Assuming Gaussian signaling, the achievable spectral efficiency for user $k$ is given by

$$\Gamma_k = \log_2 \left( 1 + \frac{\sum_{m=1}^{M} \left| h_{k,m}^H w_{k,m} \right|^2}{I_k + \sigma_k^2} \right),$$ \hspace{1cm} (2)

where

$$I_k = \sum_{m=1}^{M} \sum_{k' \neq k, k'=1}^{K} w_{k',m}^H h_{k,m} h_{k',m}^H w_{k',m}$$ \hspace{1cm} (3)

is the interference power.
B. Power Consumption Model

In this subsection, we present the power consumption model for the FDP, the FHP, and the PHP. The model includes several main power consuming hardware components, which helps us to have fair comparison of the power consumption among different precoding architectures.

We consider that the power consumption of a BS consists of the hardware power consumption and the RF transmit power. The RF transmit power at BS $m$ is given by

$$P_{tx}^{m} = \sum_{k=1}^{K} ||w_{k,m}||^2,$$  \hspace{1cm} (4)

and the hardware power consumption is given by

$$P_{hw}^{m} = \left(\frac{1}{\eta_{m}} - 1 + \Delta_{m}\right)P_{tx}^{m} + N_{m}^{PS}P_{PS} + L_{m}(P_{DAC} + P_{RF}) \frac{1 - \Delta_{m}}{1 - \Delta_{m}}.$$  \hspace{1cm} (5)

Here, $\eta_{m}$ denotes the power amplifier efficiency and $\Delta_{m}$ is a power loss factor which accounts for the extra power dissipated at various non-transmission related parts such as power supply loss and active cooling [36]. Also, $P_{PS}, P_{DAC}$ and $P_{RF}$ denote the power consumption of the PSs, DACs and RF chains, respectively, and $N_{m}^{PS}$ is the number of PSs. The number of PSs needed by each architecture is given by the FDP: $N_{m}^{PS} = 0$, the FHP: $N_{m}^{PS} = L_{m}N_{m}$, and the PHP: $N_{m}^{PS} = N_{m}$.

To further reduce the power consumption, we assume that, when there are no users associated to a BS, this BS can be put into sleep mode by deactivating some of the hardware. At the sleep mode, a BS consumes a proportion of the hardware power, which is denoted by $aP_{hw}^{m}$, where $a \in [0, 1]$ is the sleep-mode scalar whose magnitude depends on the number of hardware components deactivated during the sleep mode and affects the delay needed to reactive the BS [37]. Including the sleep mode, the general power consumption of BS $m$ is given by

$$P_{m} = \begin{cases} P_{tx}^{m} + P_{hw}^{m}, & \text{active}, \\ aP_{hw}^{m}, & \text{if } P_{tx}^{m} = 0, \text{ sleep.} \end{cases}$$  \hspace{1cm} (6)

The above power consumption model aims to enable more accurate power consumption comparisons among various precoding architectures than previously proposed models, e.g., [21], [32], [33]. The reason is that, for the active mode, the hardware power $P_{hw}^{m}$ is dependent on the RF transmit power $P_{tx}^{m}$, which models the load dependency of the power amplifier and power losses. By separating the hardware power and the RF transmit power, the impact of the hardware difference and the precoders can be better determined.
III. COOPERATIVE HYBRID BEAMFORMING

Considering the hardware power consumption and the sleep mode, for each of the considered precoding architectures, i.e., FDP, FHP and PHP, our objective is to find the analog and digital precoders \((R_m, d_{k,m}, \forall m, k)\) such that the sum power consumption is minimized. The problem can be summarized as

\[
P : \min_{\{R_m, d_{k,m}\}} \sum_{m=1}^{M} b_m P_m \tag{7}
\]

subject to

\[
\Gamma_k \geq \tau_k, \quad \forall k \tag{8}
\]

\[
\sum_{k=1}^{K} ||R_m d_{k,m}||^2 \leq P_{\text{max}, m}, \quad \forall m \tag{9}
\]

\[
|R_m^{(i,j)}| = \frac{1}{\sqrt{N_m L_m}}, \quad \forall m, i, j, \quad \text{for FHP}, \tag{10}
\]

\[
R_m = \text{diag}[\hat{r}_{1,m}, \ldots, \hat{r}_{L_m,m}], \quad |\hat{r}_{i,m}^{(j)}| = 1/\sqrt{L_m}, \quad \forall m, j, \quad \text{for PHP}. \tag{11}
\]

where \(b_m\) is the weighting parameter for balancing the load of BS \(m\), \((8)\) is the minimum spectral efficiency constraint for each user with target spectral efficiency \(\tau_k\), and \((9)\) is the maximum RF transmit power constraint on each BS. For the hybrid precoding architectures, the additional constraint \((10)\) and \((11)\) needs to be considered for the FHP and PHP, respectively.

The solution to problem \(P\) gives the analog precoder for each BS and the digital precoder for each BS-user pair. It also reflects the user association strategy, as the set of associated users of BS \(m\) is given by

\[
K_m = \{k | 0 \leq k \leq K, ||R_m d_{k,m}||^2 > 0\}, \tag{12}
\]

and the set of serving BSs of UE \(k\) is given by

\[
M_k = \{m | 0 < m \leq M, ||R_m d_{k,m}||^2 > 0\}. \tag{13}
\]

Solving problem \(P\) is challenging, because it is a non-convex optimization problem due to the analog precoder constraints \((10)\) and \((11)\). Moreover, the design parameters involved is large due to the fact that we have to jointly design precoders for each user-BS pair. Therefore, for the single-BS setup, the hybrid precoding optimization is often solved by sub-optimal methods that decouple the analog and digital precoding processes and use iterative update over the analog and digital precoders to approach the performance of the FDP \([13]–[18]\). However, such iterative
optimization methods may not be suitable for cooperative multi-cell and multi-user systems due to the fairly large amount of design parameters and feedback from all BSs, therefore, causing high computation complexity and delay.

In order to make the problem tractable, we propose a sub-optimal solution by decoupling the optimization problem $P$ into an analog precoding problem, which only depends on the channel information, and a digital precoding problem that minimizes the sum power consumption $P(7)$ conditioned on the analog precoders. The method avoids complex iterative optimization and, as shown in Section IV, achieves close spectral efficiency for a given RF transmit power compared to FDP.

A. Analog Precoding

Define the array gain between user $k$ and BS $m$ as $g_{m,k} = |h_{k,m}^H R_m d_{k,m}|^2$, we seek an analog precoder $R_m$ such that the array gain is maximized. In order to decouple the analog and digital precoders, we first give an upper bound of the array gain, then we find analog precoders maximizing the upper bound, which is independent of the digital precoders.

For the FHP, using Cauchi-Schwarz inequality, the upper bound of the array gain is given by

$$|h_{k,m}^H R_m d_{k,m}|^2 \leq \left( |h_{k,m}^H r_{k,m}|^2 + ||h_{k,m}^H (R_m)_{-k}||_F^2 \right) ||d_{k,m}||^2,$$

(14)

where $r_{k,m}$ is the $k$-th column of $R_m$ and $(R_m)_{-k}$ is a matrix after removing the $k$-th column from $R_m$.

Treating $||h_{k,m}^H (R_m)_{-k}||_F^2 ||d_{k,m}||^2$ as the interference power from other users, we simplify the array gain maximization by only focusing on the useful signal power $|h_{k,m}^H r_{k,m}|^2 ||d_{k,m}||^2$ without considering interference management. Hence, conditioned on the digital precoders, the maximization of the upper bound in (14) over $r_{k,m}$ gives the following analog precoding problem

$$P_1 : \max_{\{r_{k,m}\}} |h_{k,m}^H r_{k,m}|^2$$

(15)

s.t. $|r_{k,m}^{(i)}| = \frac{1}{\sqrt{L_m N_m}}, \quad 0 < i \leq N_m.$

(16)

The maximization problem $P_1$ is equivalent to the equal gain transmission problem \[^38\] which has the analytical solution given by

$$r_{k,m}^{(i)} = \frac{1}{\sqrt{L_m N_m}} e^{j(\xi + \angle h_{k,m}^{(i)})},$$

(17)
where $\xi \in (0, 2\pi]$ is an arbitrary phase and $\angle h_{k,m}^{(i)}$ is the phase angle of $h_{k,m}^{(i)}$. We notice that the equal gain transmission requires $K = L_m$, i.e., the number of serving users is equal to the number of RF chains at BS $m$. Here, we assume that each BS uses analog precoding to serve all users, however, after the digital precoding step, a BS may not transmit to all users, as the final users associated to BS $m$ will be determined by both precoders according to $K_m = \{k|0 \leq k \leq K, ||R_m d_{k,m}||^2 > 0\}$.

For the PHP, we additionally require that $R_m = \text{diag} [\hat{r}_{1,m}, \ldots, \hat{r}_{L_m,m}]$ with $|\hat{r}_{i,m}^{(j)}| = 1/\sqrt{L_m}$. Denoting the $k$-th column of $R_m$ by $r_{k,m}$, following the same upper bound maximization steps, Problem $\mathcal{P}_1$ can be rewritten as

$$\mathcal{P}_2 : \max_{\{r_{k,m}\}} |\hat{h}_{k,m}^H r_{k,m}|^2$$

s.t. $|r_{k,m}^{(i)}| = 1/\sqrt{L_m}$, \( (k - 1)N_m L_m < i \leq kN_m L_m \),

where $\hat{h}_{k,m}^H = h_{k,m}^H G_{k,m}$ and

$$G_{k,m} = \begin{bmatrix} 0_{(k-1)N_m L_m \times (k-1)N_m L_m}, I_{N_m L_m \times N_m L_m}, 0_{(L_m-k)N_m L_m \times (L_m-k)N_m L_m} \end{bmatrix}. \quad (20)$$

Problem $\mathcal{P}_2$ has the same form as Problem $\mathcal{P}_1$, therefore, the analytical solution is given by

$$r_{k,m}^{(i)} = \frac{1}{\sqrt{L_m}} e^{j(\xi + \angle h_{k,m}^{(k-1)N_m L_m + i})}, \quad 0 < i \leq N_m L_m. \quad (21)$$

Although our analog precoding solution based on the equal gain transmission neglects the interference power and digital precoding, it maximizes the signal-to-noise ratio in a single-BS analog beamforming system. The analog precoders only require the channel phase information for each user and aim to maximize the upper bound of the array gain. The interference coordination and the overall array gain enhancement will be addressed by the digital precoding. As shown in our simulations, for a given RF transmit power, this low complexity analog precoding solution and the proposed non-iterative optimization approach can achieve close spectral efficiency compared to FDP.

**B. Digital Precoding**

To achieve our objective in minimizing the sum power consumption as shown in Problem $\mathcal{P}$, next, we reformulate it to a convex semidefinite program conditioned on the analog precoders and obtain the sub-optimal digital precoders in terms of the sum power consumption.
We first rewrite the objective function (7) in its quadratic form. Define $D_k = d_k d_k^H$ with $d_k \triangleq \begin{bmatrix} d_{k,1}^T, \ldots, d_{k,M}^T \end{bmatrix}^T$ and a block diagonal matrix

$$\hat{R} = \text{diag} \left( b_1 R_1^H R_1, \ldots, b_M R_M^H R_M \right),$$

the sum power consumption of all BSs can be written as

$$\sum_{k=1}^{K} \text{Tr} \left( \hat{R} D_k \right) + \kappa(z_i),$$

where

$$\kappa(z_i) = \sum_{m=1}^{M} b_m (z_m + a(1 - z_m)) P_{hw,m}$$

and $z_m \in \{0, 1\}$ is a sleep mode indicator for BS $m$. If $z_m = 1$, the BS $m$ is in active mode, otherwise it is in the sleep mode. Here, $z_i$ is one unique sleeping mode indicator vector out of the $2^M - 1$ possible BS sleeping mode combinations.

Next, we express the spectral efficiency and the peak power constraints in more compact forms. Define the block diagonal matrices

$$\hat{H}_k = \text{diag} \left( R_1^H h_{k,1} h_{k,1}^H R_1, \ldots, R_M^H h_{k,M} h_{k,M}^H R_M \right),$$

where $h_{k,m}^H R_m$ is the effective channel of user $k$ and BS $m$ after analog precoding. Thus, (2) can be rewritten as

$$\Gamma_k = \log_2 \left( 1 + \frac{d_k^H \hat{H}_k d_k}{\sum_{k' \neq k, k'=1}^{K} d_{k'}^H \hat{H}_{k'} d_{k'} + \sigma_k^2} \right).$$

To satisfy a minimum spectral efficiency $\tau_k$, (26) can be reformed to the following quadratic inequalities by using the cyclic property of the trace operation:

$$\text{Tr} \left( D_k \hat{H}_k \right) - (2^{\tau_k} - 1) \sum_{k' \neq k, k'=1}^{K} \text{Tr}(D_{k'} \hat{H}_{k'}) \geq (2^{\tau_k} - 1) \sigma_k^2, \quad \forall k.$$

In order to rewrite the maximum transmit power constraints to the quadratic form, we first define

$$Q_{i,m} = \begin{cases} R_m^H R_m, & \text{if } i = m \\ 0_{L_m \times L_m}, & \text{otherwise} \end{cases}$$

and $Q_m = \text{diag} (Q_{1,m}, \ldots, Q_{M,m})$, then (9) can be expressed as

$$\sum_{k=1}^{K} \text{Tr} \left( Q_m D_k \right) \leq z_m P_{max,m}, \quad \forall m.$$
Combining (23), (27), and (29), for given $R_m$ and $z_i$, we have the following digital precoder optimization problem

$$\mathcal{P}_3 : \min_{\{D_k\}} \sum_{k=1}^{K} \text{Tr}(\hat{R} D_k) + \kappa(z_i)$$

subject to

$$\text{Tr}(D_k \hat{H}_k) \geq (2^{\tau_k} - 1) \sum_{k' \neq k, k' = 1}^{K} \text{Tr}(D_{k'} \hat{H}_{k'}) + (2^{\tau_k} - 1) \sigma_k^2, \quad \forall k,$$

$$\sum_{k=1}^{K} \text{Tr}(Q_m D_k) \leq z_m P_{\text{max},m}, \quad \forall m.$$  \hspace{1cm} (32)

Problem $\mathcal{P}_3$ is a convex semidefinite problem if we apply the semidefinite relaxation by replacing the constraint $\text{rank}(D_k) = 1$ with $D_k \succeq 0$. Then, it can be solved efficiently using convex optimization tools and a rank-1 solution always exists given that the solution is feasible [26]. The solution that minimizes the sum power consumption considering the BS sleep/active modes is given by solving $\mathcal{P}_3$ for all possible $z_i$. Moreover, the user association is found by checking the user-BS pairs with non-zero $||R_m d_{k,m}||^2$.

The digital precoding problem is conditioned on the analog precoders given by (17) or (21) and is independent of the hybrid precoding architectures. Although the proposed analog precoding does not guarantee the spectral efficiency and may cause some interference, the phase and amplitude are further adjusted via the digital precoding. The final sub-optimal solution in terms of the sum power consumption is based on the effective channel after analog precoding and ensures that the spectral efficiency and power constraints for all users and BSs are satisfied.

The overall proposed hybrid precoding for cooperative mmWave networks is described as follows and in Algorithm 1. In the algorithm, $z_0$, $P_{\text{tx}}^*$ and $P^*$ denote the optimal BS sleep indicator in terms of minimizing the sum power, the sum RF transmit power and the sum power consumption for $z_0$, respectively. We first find $R_m$ for all BSs based on (17) or (21), depending on the considered precoding setup. Conditioned on $R_m$ and $z_i$, the relaxed semidefinite program in $\mathcal{P}_3$ can be solved efficiently via convex optimization tools. The optimal BS sleep/active mode pattern $z_0$ is given by choosing $z_i$ with the minimum sum power consumption. Then, the digital precoder based on $D_k^*(z_0)$ minimizes the sum power consumption. The solution $D_k^*(z_0)$ is rank-1 and the stacked digital precoders $d_k$ are readily available via singular value decomposition.
Algorithm 1 Hybrid Precoding for Cooperative MmWave Networks

Require: $h_{k,m}, \forall k,m$

1: $\mathbf{R}_m = 0_{N_m \times L_m}$

2: for $m \leftarrow 1$ to $M$ do

3: \hspace{1em} for $k \leftarrow 1$ to $K$ do

4: \hspace{2em} $r_{k,m} = \frac{1}{\sqrt{L_m N_m}} e^{(\xi + \angle h_{k,m}^{(i)})}$, for FHP

5: \hspace{2em} $\hat{r}_{k,m}^{(i)} = \frac{1}{\sqrt{L_m}} e^{j(\xi + \angle h_{k,m}^{(i)})}$, for PHP

6: \hspace{1em} end for

7: end for

8: for $i \leftarrow 1$ to $2^M - 1$ do

9: \hspace{1em} Solve $\mathcal{P}_3$ for fixed $\mathbf{R}_m$ and $z_i$, obtain $\mathbf{D}_k^i(z_i)$. 

10: end for

11: $o = \arg\min_i \sum_{k=1}^{K} \text{Tr} \left( \hat{\mathbf{R}} \mathbf{D}_k^i(z_i) \right) + \kappa(z_i)$

12: Optimal sleeping mode indicator $z_o$.

13: Minimum sum RF transmit power $P_{tx}^* = \sum_{m=1}^{M} \sum_{k=1}^{K} \text{Tr} \left( \mathbf{Q}_m \mathbf{D}_k^i(z_o) \right)$,

14: Minimum sum power $P^* = \sum_{k=1}^{K} \text{Tr} \left( \hat{\mathbf{R}} \mathbf{D}_k^i(z_o) \right) + \kappa(z_o)$.

C. Lagrangian Analysis

In order to gain insights of the cooperative transmissions and the user associations, in this subsection, we analyze the Lagrangian dual problem of the optimization problem and provide conditions for the user association strategy.

Conditioned on analog precoders, the Lagrangian of the optimization problem $\mathcal{P}_3$ is given by

\[
\mathcal{L}(\mathbf{D}_k, \lambda_k, \mu_k) = \sum_{k=1}^{K} \text{Tr} \left( \hat{\mathbf{R}} \mathbf{D}_k \right) + \kappa(z) + \sum_{k=1}^{K} \lambda_k \times \left( (2^{\tau_k} - 1) \left( \sum_{k' \neq k, k' = 1}^{K} \text{Tr} \left( \mathbf{D}_k \hat{\mathbf{H}}_{k'} \right) + \sigma_k^2 \right) - \text{Tr} \left( \mathbf{D}_k \hat{\mathbf{H}}_k \right) \right) + \sum_{m=1}^{M} \mu_m \left( \sum_{k=1}^{K} \text{Tr} \left( \mathbf{Q}_m \mathbf{D}_k \right) - z_m P_{\max,m} \right),
\]

(33)
where $\lambda_k$ and $\mu_k$ are non-negative Lagrange multipliers. The dual function is given by

$$g(\lambda_k, \mu_m) = \min_{\{D_k\}} \mathcal{L}(D_k, \lambda_k, \mu_k)$$

$$= \kappa(z) + \sum_{k=1}^{K} \lambda_k(2^{\tau_k} - 1)\sigma_k^2 - \sum_{m=1}^{M} \mu_m z_m P_{\text{max},m} + \min_{\{D_k\}} \sum_{k=1}^{K} \text{Tr} \left( Y_k D_k \right), \quad (34)$$

where

$$Y_k = \hat{R} + \sum_{k' \neq k, k'=1}^{K} \lambda_{k'}(2^{\tau_{k'}} - 1)\hat{H}_{k'} + \sum_{m=1}^{M} \mu_m Q_m - \lambda_k \hat{H}_k. \quad (35)$$

The minimum of (34) is $-\infty$ except for $Y_k \geq 0, \forall k$. Thus, the Lagrange dual problem is

$$\max_{\{\lambda_k, \mu_m\}} \kappa(z) + \sum_{k=1}^{K} \lambda_k(2^{\tau_k} - 1)\sigma_k^2 - \sum_{m=1}^{M} \mu_m z_m P_{\text{max},m}$$

$$\text{s.t.} \quad Y_k \geq 0, \quad \forall k. \quad (36)$$

Let $\lambda^*_k$ and $\mu^*_k, \forall k$ denote the optimal solutions for (36) and $D^*_k$ be the optimal solution for (34). Because $D^*_k = d^*_k(d^*_k)^H$ and strong duality holds, the optimal digital precoders $d^*_k$ can be obtained by

$$\frac{\partial \mathcal{L}(D_k, \lambda^*_k, \mu^*_m)}{\partial d_k} = Y_k d^*_k = 0. \quad (37)$$

Because $Y_k$ is a block diagonal matrix, for each diagonal element, we have

$$\left( b_m R_m^H R_m + \sum_{k' \neq k, k'=1}^{K} \lambda_{k'}(2^{\tau_{k'}} - 1)R_{m}^H h_{k',m} h_{k',m}^H R_m + \sum_{m'=1}^{M} \mu_{m'} Q_{m,m'} - \lambda_k R_m^H h_{k,m} h_{k,m}^H R_m \right) d_{k,m} = 0, \quad (38)$$

hence, for a user $k$ to be served by BS $m$, the optimal digital precoder should satisfy

$$d^*_k = c_{k,m}(B_{k,m})^{-1}R_m^H h_{k,m}, \quad (39)$$

where

$$c_{k,m} = \lambda^*_k h_{k,m}^H R_m d_{k,m}^* \quad (40)$$

$$B_{k,m} = b_m R_m^H R_m + \sum_{k' \neq k, k'=1}^{K} \lambda_{k'}(2^{\tau_{k'}} - 1)R_{m}^H h_{k',m} h_{k',m}^H R_m + \sum_{m'=1}^{M} \mu_{m'} Q_{m,m'}. \quad (41)$$

Multiplying (38) with $(d^*_{k,m})^H$ from the left and plugging in (39), we have

$$c_{k,m}^2 \left( h_{k,m}^H R_m (B_{k,m})^{-1}R_m^H h_{k,m} - \lambda_k^* \left( h_{k,m}^H R_m (B_{k,m})^{-1}R_m^H h_{k,m} \right)^2 \right) = 0. \quad (42)$$
Hence, if a user $k$ should be served by BS $m$, the optimal multiplier should satisfy

$$\lambda_k^* = \frac{1}{h_{k,m}^H R_m^{-1} R_m^H h_{k,m}}$$  \hspace{0.5cm} \text{(43)}$$

otherwise, the RF transmit power is set to 0 in order to satisfy (42). Therefore, (43) is a necessary condition for user $k$ to be associated to BS $m$. Furthermore, according to the feasibility constraint $Y_k \geq 0$, the Lagrange multiplier is found to satisfy

$$\lambda_k \leq \frac{1}{h_{k,m}^H R_m^{-1} R_m^H h_{k,m}}, \forall m,$$  \hspace{0.5cm} \text{(44)}$$

where the equality holds for BSs with maximum $h_{k,m}^H R_m^{-1} R_m^H h_{k,m}$. Hence, the set of serving BSs of user $k$ is given by

$$\left\{ m \mid \arg \max_m h_{k,m}^H R_m^{-1} R_m^H h_{k,m} \right\}. \hspace{0.5cm} \text{(45)}$$

The above expression gives the optimal user association principle in terms of the total transmit power. A user is served by multiple BSs if (45) has multiple elements, otherwise, the user is served by a single BS. According to (45), the user association strategy is not simple as it is based on the RF precoders, the target spectral efficiency, the transmit power limit and the interference power. The serving BS of a user depends on the norm of the effective channel after analog precoding $R_m^H h_{k,m}$ which is weighted by $B_{k,m}$ such that BSs with less interference power or higher transmit power limit are chosen.

**IV. Simulation Results**

In this section, we present the simulation results for the FHP and the PHP, where the precoders are obtained by our proposed algorithm, and compare the results with those given by the FDP. We first introduce the baseline simulation environment including the channel model and the hardware parameters. Then, we analyze the simulation results in terms of three aspects:

- **Beam patterns.** In Section IV-B (Fig. 2), we plot the beam patterns of the FHP, the PHP and the FDP, which provides direct visualization of the directional transmissions and interference management achieved by the cooperative precoding.

- **Power consumption.** Section IV-C (Figs. 3-4) shows that, for the same RF transmit power compared to the FDP, the hybrid precoding under the case of the FHP achieves close spectral efficiency as in the FDP. Moreover, Fig. 4 studies the effect of hardware power consumption on the sum power consumption for different precoding architectures.

- **Value of cooperation.** In Section IV-D (Figs. 5-8), we show how the BS cooperation reduces the sum RF transmit power, the sum power of the BSs, as well as the infeasible solutions.
(a) BS1, $N_1 = 64$. (b) BS2, $N_2 = 64$. (c) BS1, $N_1 = 128$. (d) BS2, $N_2 = 128$.

Fig. 2: Beam patterns in a network with $M = 2$ BSs and $K = 4$ users. All BSs are set to be active, the number of RF chains is $L_m = 4$ per BS and the target spectral efficiency is $\tau_k = 4$ bit/s/Hz per user. The AOD is assumed to be at $-60^\circ, -30^\circ, 30^\circ, 60^\circ$. The beam gains are shown in dB scale and are normalized with respect to the largest beam gain in the FDP. (a) and (b): Beam patterns at BS1 and BS2 with 64 antennas, respectively. (c) and (d): Beam patterns at BS1 and BS2 with 128 antennas, respectively. Here, the solid black line is for the PHP, the red dash-dot curve is for the FHP and the dash blue curve represents the results for the FDP.

A. Simulation Environment

Since large antenna arrays and directive transmissions make the mmWave multi-path channel sparser than the lower frequency channel, we use a clustered channel model to characterize the channel sparsity. The channel model considers a few dominant signal paths and clusters the multi-path components according to different reflections. The channel vector between BS $m$ and user $k$ is given by [13]–[15]

$$h_{k,m} = \sqrt{\rho_{k,m} N_m} \frac{N_{\text{cl}}}{N_{\text{cl}} N_{\text{ray}}} \sum_{i=1}^{N_{\text{cl}}} \sum_{l=1}^{N_{\text{ray}}} \alpha_{i,l} a_m^r(\theta_{i,l}) a_m^c(\phi_{i,l})^H,$$

where $N_{\text{cl}}$ is the number of clusters, $N_{\text{ray}}$ is the number of paths within a cluster, $\alpha_{i,l} \sim \mathcal{CN}(0,1)$ is the gain of the $l$-th path in the $i$-th cluster and $a_m(\theta_{i,l})$ is the normalized transmit antenna array response vector. For an $N_m$-element uniform linear array, it is given by

$$a_m(\theta) = \frac{1}{\sqrt{N_m}} \left[ 1, e^{jkd \sin(\theta)}, \ldots, e^{jkd(N_m-1)\sin(\theta)} \right]^T,$$

where $\theta_{i,l}$ is the angle of departure (AOD), $k = 2\pi/\Lambda$, $\Lambda$ is the wavelength and $d = \Lambda/2$ is the antenna spacing. Also, $\theta_{i,l}$ is assumed to follow a truncated Laplace distribution with mean cluster angle $\bar{\theta}_i \sim \mathcal{U}(\theta_{\text{min}}, \theta_{\text{max}})$ and angular spread $\sigma_\theta$, which assumes the transmitter uses sector transmissions. For a uniform planar array of $U \times V$ elements, the array response vector is given by

$$a_m(\theta, \phi) = \frac{1}{\sqrt{N_m}} \left[ 1, e^{jkd(\sin(\theta) \sin(\phi) + \cos(\phi))}, \ldots, e^{jkd((U-1)\sin(\theta) \sin(\phi) + (V-1)\cos(\phi))} \right]^T,$$
TABLE I: Simulation parameters.

| Parameter | Value       |
|-----------|-------------|
| $P_{PS}$  | 40 [mW]    |
| $P_{DAC}$ | 200 [mW]   |
| $P_{RF}$  | 40 [mW]    |
| $P_{\text{max},m}$ | 55 [dBm] |
| $\eta_m$  | 0.3        |
| $\Delta_m$| 0.15       |

where $\theta$ and $\phi$ denote the azimuth and elevation angles, respectively.

The path loss in mmWave channels differs greatly depending on the LOS and NLOS environment [9], [31], [39]. In order to incorporate both the LOS and NLOS models together, the path loss of LOS or NLOS transmissions is determined by a LOS probability as a function of the transmission distance. In this way, the path loss between BS $m$ and user $k$ is given by

$$\rho_{k,m} = \mathbb{I}(p_L(\hat{d})) P_{\text{LOS}}^{-1} + \left(1 - \mathbb{I}(p_L(\hat{d}))\right) P_{\text{NLOS}}^{-1},$$

(49)

where $\hat{d}$ is the distance and $\mathbb{I}(p_L(\hat{d}))$ is a Bernoulli random variable. The LOS probability is given by $p_L(\hat{d}) = e^{-\beta \hat{d}}$ [40], where the exponentially decaying probability models the fact that the probability of LOS decreases with distance and $\beta$ models the average blockage density that causes the NLOS condition.

The path loss follows the close-in free space reference distance model and is of the form [31], [39]

$$P_{\text{LOS/NLOS}}[\text{dB}] = 20 \log_{10} \left(\frac{4\pi}{\lambda}\right) + 10\bar{n}_{\text{LOS/NLOS}} \log_{10}(\hat{d}) + X_{\text{LOS/NLOS}}, \quad \hat{d} \geq 1 \text{ m},$$

(50)

where $\bar{n}_{\text{LOS/NLOS}}$ and $X_{\text{LOS/NLOS}}$ are the LOS and NLOS dependent path loss exponent and log-normal distributed shadowing with standard deviation $\sigma_{\text{LOS/NLOS}}$, respectively.

We choose hardware power consumption values given in Table I for reference in our simulations. Such parameter settings are in harmony with, e.g., [21], [32], and have been selected based on our discussion with Ericsson, so that we provide fair comparisons for different architectures. The simulation results are averaged over $10^5$ channel realizations, and in each realization users are randomly dropped in an area of $200 \text{ m} \times 200 \text{ m}$.

B. On Beam Pattern

In Fig. 2, we compare the beam patterns of the FDP, the FHP, and the PHP for $M = 2$ BSs, $K = 4$ users and $L_m = 4$ RF chains per BS. All BSs are set to be active and the target spectral efficiency is $\tau_k = 4 \text{ bit/s/Hz}$ for all users. The shown beam patterns of different array sizes are
Fig. 3: Sum RF transmit power $P_{tx}^*$ with the optimal sleep mode pattern. The network parameters are $M = 2$ BSs, $K = 4$ users, $N_m = 64$ antennas and $L_m = 4$ RF chains per BS.

Based on different channel realizations and consider only the azimuth domain, although we have observed the same qualitative behavior in all tested channel realizations and in the elevation domain.

In Figs. 2(a) and 2(b), we can see the two BSs jointly serve four users with main lobes pointed at AODs. By comparing the beam patterns among different architectures, it confirms that our proposed FHP gives close performance to the FDP in terms of the main lobe angles and maximum gains. Also, we notice that the PHP generates wider beams and thus may cause more interference to other directions. Figures 2(c) and 2(d) show that, by increasing the number of antennas, the beam main lobes become narrower and more energy focusing, especially for the PHP. Hence, if a large antenna array is available, it is possible to achieve energy-focusing narrow beams with PHP, while keeping the number of PSs and complexity low.

C. On Power Consumption

In this subsection, we assess the power consumption of different architectures and show the effect of the number of antennas and the sleep mode scalar $a$ on the power consumption.

Setting $M = 2$ BSs, $K = 4$ users, $N_m = 64$ antennas, $L_m = 4$ RF chains per BS, Fig. 3 shows the sum RF transmit power consumption $P_{tx}^*$ versus target spectral efficiency. It confirms that, for a given RF transmit power, the FHP scheme can achieve close spectral efficiency compared to the FDP and it is also true for other BS settings not shown here. However, the PHP needs to
transmit with more power than the FDP or the FHP to achieve a given target spectral efficiency. Intuitively, this is because the reduced number of RF chains and PSs leads to less control over the precoders. In order to satisfy a high target spectral efficiency, it becomes difficult for the PHP to form as energy-focusing beams as the FDP or the FHP, thus the PHP needs to increase the RF transmit power.

To examine the effect of antennas and PSs on the power consumption, we compare the sum power consumption for two PS power values $P_{PS} = 10$ mW and $P_{PS} = 40$ mW in Fig. 4. Here, the results are presented for the cases with the number of BSs $M = 2$, the number of users $K = 4$ and the target spectral efficiency $\tau_k = 4$ bit/s/Hz per user. Figure 4 shows that, for small array sizes, the sum power consumption of the FHP decreases with the number of antennas as a result of increased beamforming gains and, thus, reduced RF transmit power. However, the sum power consumption starts to increase for larger number of antennas and the increase rate is larger for a larger $P_{PS}$. This is intuitively so because the power consumption is dominated by the increased hardware power when the number of antenna is large. Also, Figure 4 shows that the sum power consumption of the PHP keeps decreasing with the number of antennas since the number of PSs needed in the PHP increases less quickly than that of the FHP and the power consumption is dominated by the decreasing RF transmit power. Therefore, depending on how fast the hardware power increases with the number of antennas, there exists a cross-over point.
where the PHP outperforms the FHP, in terms of the sum power consumption for a given target spectral efficiency. The result indicates that, in order for the PHP to consume less power than the FHP, the RF transmit power gap between the PHP and the FHP needs to be reduced, which can be achieved by increasing the number of antennas or improving the PHP architecture such that a larger beamforming gain is achieved.

D. On the Value of Cooperation

In this subsection, we show simulation results for different number of BSs and analyze the value of cooperation, in terms of the power consumption, infeasibility probability, as well as the BS cooperation probability.

Setting the number of antennas $N_m = 64$, the number of users $K = 4$, the number of RF chains $L_m = 4$ per BS and the target spectral efficiency $\tau_k = 4$ bit/s/Hz per user, Fig. 5(a) and 5(b) show the sum RF transmit power and the sum power consumption versus the number of cooperative BSs, respectively. As seen in Fig. 5(a) and (b), there is small difference between the case with all BSs being active and the case with the sleep mode for all architectures. This is, intuitively so, because when $M < 3$ the average BS activation probability of the FHP scheme is close to 100% and it decreases to around 30% at $M = 5$. Even though more power is saved in
Fig. 6: Infeasibility probability versus the number of BSs. A large $\beta$ defines a denser blockage environment with less LOS transmissions. The network parameters are the number of antennas $N_m = 64$, the number of users $K = 4$, the number of RF chains $L_m = 4$ per BS and the target spectral efficiency $\tau_k = 4 \text{ bit/s/Hz}$ per user.

Fig. 7: BS joint transmission probability versus the number of BSs. A larger $\beta$ defines a denser blockage environment with less LOS transmissions. The network parameters are the number of antennas $N_m = 64$, the number of users $K = 4$, the number of RF chains $L_m = 4$ per BS and the target spectral efficiency $\tau_k = 4 \text{ bit/s/Hz}$ per user.

the sleep mode, for a sleep-mode scalar $a = 0.5$, the saved hardware power is small compared to the RF transmit power. Especially, for the PHP, as the hardware power consumption of the PHP is the least among the considered architectures, negligible difference between the all active case and the sleep mode case is observed, hence, the sum power consumption of the PHP tends to be RF-power-dominated.
Fig. 8: CDF of the average RF transmit power for individual BSs. The network parameters are set to $N_m = 64$, $K = 4$, $L_m = 4$ per BS and $\tau_k = 4$ bit/s/Hz per user.

In Fig. 5(a), the sum RF transmit power is shown to decrease with the number of BSs for all architectures. Comparing a cooperative network having 2 cooperative BSs with the cases having 1 BS, the sum RF transmit power consumption is reduced by 65%, 71% and 56% for the FDP, the FHP and the PHP (see Fig. 5(a)), respectively. The result shows that the network densification and cooperative transmissions lead to a better chance for a user to be served by BSs with good channel conditions, thus, requiring less RF transmit power to achieve a target spectral efficiency$^1$. In Fig. 5(b), the sum power consumption is reduced by 54%, 64% and 55% for the FDP, the FHP and the PHP, respectively, when $M = 1 \to 2$. For $M > 4$, the PHP starts to consume less sum power than the FHP, which is due to the fact that the hardware increase has less effect on the sum power consumption than that of the FHP, and the sum power consumption of the PHP is dominated by the decreasing RF transmit power. The results in Fig. 5 indicate that, due to the network densification, cooperative transmissions have the potential to reduce both the RF transmit power and the sum power consumption of the network.

To examine the feasibility of the proposed precoding algorithm under different blocking conditions, the infeasibility probability is illustrated in Fig. 6 for the number of antennas $N_m = 64$, the number of users $K = 4$, the number of RF chains $L_m = 4$ per BS and the target spectral

---

$^1$Indeed, cooperation between the BSs is at the cost of backhauling [41], which is not considered in this work.
efficiency $\tau_k = 4$ bit/s/Hz per user. Here, the infeasibility probability is defined as the probability that a feasible precoder solution cannot be found to support the target spectral efficiency. For a fixed $\beta$ which models the blockage density, the results show that the infeasibility probability drops substantially from single BS transmission to cooperative transmissions and converges to 0 as $M$ increases. For a given $M$, as expected, the infeasibility probability increases with $\beta$ as a result of the increased NLOS transmissions. The results indicate that cooperative networks is effective in reducing the infeasibility probability in areas with dense blockages and many NLOS transmissions.

In Fig 7, the joint transmission probability, defined as the probability that a user is served by more than one BS, is shown for the cases with $N_m = 64$, $K = 4$, $L_m = 4$ per BS and $\tau_k = 4$ bit/s/Hz per user. For $M > 2$, the joint transmission probability of the FHP and the PHP increases with $\beta$, as it becomes difficult for a single BS to satisfy the target spectral efficiency when the probability of NLOS transmission is high. Also, we notice that the joint transmission probability decreases when the number of BSs approaches the number of users. Intuitively, this is because joint transmissions from multiple BSs impose additional interference to other users and the algorithm tries to assign one BS per user. Even though the joint transmission probability is small, BS joint transmission still provides performance gain in the following two aspects. First, by jointly performing the user association, a user is optimally associated with BSs having good channel conditions and least transmit power that satisfies the quality-of-service constraints. Second, the BS sleep strategy reduces the average activation probability for larger number of cooperating BSs, thus optimizing the sum power consumption. Additionally, Fig 7 shows that, for the same $\beta$, the joint transmission probability of the PHP is higher than that of the FHP, since the infeasibility probability of the PHP is larger, thus requiring more help from other BSs.

To check the effect of number of BSs on the RF transmit power range, Fig. 8 shows the RF transmit power cumulative distribution function (CDF) of individual BSs for $M = \{1, 2, 4\}$, $N_m = 64$, $K = L_m = 4$, $\forall m$, and $\tau_k = 4$ bit/s/Hz, $\forall k$. In general, cooperative transmissions reduce the transmit power variations by optimizing the user associations, which might be very useful in Effective Isotropic Radiated Power (EIRP) limited deployment scenarios. In Fig. 8, the RF transmit power for the case with 4 BSs has less variance than the cases with smaller number of BSs. The 95-th percentile of the transmit power varies from approximately 2.4W in the 4-BS cooperation case to 235W in the 1-BS case for both the FHP and the PHP. The power variation may affect the performance of hardware components such as power amplifiers considering the
output power dynamic range.

The simulations in this subsection verify that a cooperative network with more BSs can achieve lower RF/sum power consumption, and is effective in reducing the infeasibility probability and transmit power variations. However, we also notice that the BS cooperation increases the backhaul and coordination overhead [41], studying their effects on the network performance is out of the scope of this paper and is left for future work.

V. CONCLUSION

In this paper, we proposed a hybrid beamforming algorithm that enables joint transmissions in a cooperative multi-cell multi-user mmWave network. By allowing joint transmissions from multiple BSs to each user, we first showed that minimization of the sum power consumption with minimum user rate constraints and maximum RF transmit power constraints can be decoupled into independent analog precoding and digital precoding problems. Next, we analyzed the Lagrangian dual problem of the convex digital precoder optimization problem which gave the conditions for the optimal user association strategy that minimizes the sum power consumption of the network. Simulation results have been presented for various architectures including the FDP, the FHP and the PHP. Simulations on the power consumption verifies that the FHP achieves close RF transmit power compared to the FDP and the hardware power consumption has a large impact on the sum power consumption of the hybrid precoding architectures. Simulations on different number of BSs show that the FDP, FHP and PHP schemes are each best under different number of cooperative BSs. The results demonstrate that cooperative transmissions have the advantage of reducing the sum power consumption of the network, the infeasibility probability and the RF transmit power variations.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What will 5G be?” IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[2] X. Lin, J. Li, R. Baldemair, T. Cheng, S. Parkvall, D. Larsson, H. Koorapaty, M. Frenne, S. Falahati, A. Grövlen, and K. Werner, “5G new radio: Unveiling the essentials of the next generation wireless access technology,” Jun. 2018. [Online]. Available: https://arxiv.org/abs/1806.06898
[3] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, present, and future,” IEEE J. Sel. Areas Commun., vol. 30, no. 3, pp. 497–508, Apr. 2012.
[4] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[5] J. G. Andrews, F. Baccelli, and R. K. Ganti, “A tractable approach to coverage and rate in cellular networks,” *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Dec. 2011.

[6] C. Fang, B. Makki, X. Xu, and T. Svensson, “Equal gain combining in poisson networks with spatially correlated interference signals,” *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 628–631, Dec. 2016.

[7] mmMAGIC, “Measurement results and final mmMAGIC channel models,” Deliverable D2.2, May 2017. [Online]. Available: https://bit.ly/2NCbKYs

[8] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, “Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design,” *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

[9] T. Bai and R. W. Heath, “Coverage and rate analysis for millimeter-wave cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[10] “3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on channel model for frequency spectrum above 6 GHz (Release 14),” 3GPP TR 38.900, Jul. 2017.

[11] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini, “Channel feasibility for outdoor non-line-of-sight mmwave mobile communication,” in *Proc. IEEE VTC Fall*, Québec City, Canada, Sep. 2012, pp. 1–6.

[12] C. Fang, B. Makki, and T. Svensson, “Coverage analysis for millimeter wave uplink cellular networks with partial zero-forcing receivers,” in *Proc. IEEE WiOpt*, Paris, France, May 2017, pp. 1–6.

[13] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[14] C. Rusu, R. Méndez-Rial, N. González-Prelcic, and R. W. Heath, “Low complexity hybrid precoding strategies for millimeter wave communication systems,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8380–8393, Dec. 2016.

[15] W. Ni and X. Dong, “Hybrid block diagonalization for massive multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.

[16] F. Sohrabi and W. Yu, “Hybrid digital and analog beamforming design for large-scale antenna arrays,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[17] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, “Energy-efficient hybrid analog and digital precoding for MmWave MIMO systems with large antenna arrays,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[18] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, “Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[19] L. Zhou and Y. Ohashi, “Hybrid precoding with data stream adaptation for high throughput mmwave MIMO systems,” in *Proc. IEEE WCNC*, Doha, Qatar, Apr. 2016, pp. 1–6.

[20] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, “Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2274–2287, Apr. 2017.

[21] R. Méndez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, “Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?” *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.

[22] C. Lin and G. Y. Li, “Energy-efficient design of indoor mmwave and sub-THz systems with antenna arrays,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4660–4672, Jul. 2016.

[23] S. He, C. Qi, Y. Wu, and Y. Huang, “Energy-efficient transceiver design for hybrid sub-array architecture MIMO systems,” *IEEE Access*, vol. 4, pp. 9895–9905, Jan. 2016.

[24] X. Ge, Y. Sun, H. Gharavi, and J. Thompson, “Joint optimization of computation and communication power in multi-user massive MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4051–4063, Jun. 2018.
[25] 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 J. Sel. Areas Commun., vol. 31, no. 2, pp. 226–240, Feb. 2013.

[26] J. Li, E. Björnson, T. Svensson, T. Eriksson, and M. Debbah, “Joint precoding and load balancing optimization for energy-efficient heterogeneous networks,” IEEE Trans. Wireless Commun., vol. 14, no. 10, pp. 5810–5822, Oct. 2015.

[27] Y. Cheng, M. Pesavento, and A. Philipp, “Joint network optimization and downlink beamforming for CoMP transmissions using mixed integer conic programming,” IEEE Trans. Signal Process., vol. 61, no. 16, pp. 3972–3987, Aug. 2013.

[28] H. J. Yang, W. Shin, B. C. Jung, C. Suh, and A. Paulraj, “Opportunistic downlink interference alignment for multi-cell mimo networks,” IEEE Trans. Wireless Commun., vol. 16, no. 3, pp. 1533–1548, Mar. 2017.

[29] A. Michaloliakos, W. C. Ao, and K. Psounis, “Joint user-beam selection for hybrid beamforming in asynchronously coordinated multi-cell networks,” in Proc. IEEE IFA, San Diego, USA, Jan. 2016, pp. 1–10.

[30] S. Sun, T. S. Rappaport, M. Shafi, and H. Tataria, “Analytical framework of hybrid beamforming in multi-cell millimeter-wave systems,” IEEE Trans. Wireless Commun., vol. 17, no. 11, pp. 7528–7543, Nov. 2018.

[31] G. R. MacCartney and T. S. Rappaport, “Millimeter-wave base station diversity for 5G coordinated multipoint (CoMP) applications,” IEEE Trans. Wireless Commun., vol. 18, no. 7, pp. 3395–3410, Jul. 2019.

[32] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi, “Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs,” IEEE Trans. Wireless Commun., vol. PP, no. 99, pp. 1–1, Dec. 2017.

[33] C. Fang, B. Makki, J. Li, and T. Svensson, “Coordinated hybrid precoding for energy-efficient millimeter wave systems,” in Proc. IEEE SPAWC, Kalamata, Greece, Jun. 2018, pp. 1–5.

[34] T. V. Chien, E. Björnson, and E. G. Larsson, “Joint power allocation and user association optimization for massive MIMO systems,” IEEE Trans. Wireless Commun., vol. 15, no. 9, pp. 6384–6399, Sep. 2016.

[35] B. Makki and T. Eriksson, “On the performance of MIMO-ARQ systems with channel state information at the receiver,” IEEE Trans. Commun., vol. 62, no. 5, pp. 1588–1603, May 2014.

[36] “Energy efficiency analysis of the reference systems, areas of improvements and target breakdown,” Energy Aware Radio Netw. Technol., Tokyo, Japan, EU FP7 INFO-ICT-247733 EARTH, D2.3, Jan. 2012.

[37] B. Debaillie, C. Desset, and F. Louagie, “A flexible and future-proof power model for cellular base stations,” in Proc. IEEE VTC spring, Glasgow, Scotland, May 2015, pp. 1–7.

[38] D. J. Love and R. W. Heath, “Equal gain transmission in multiple-input multiple-output wireless systems,” IEEE Trans. Commun., vol. 51, no. 7, pp. 1102–1110, Jul. 2003.

[39] M. K. Samimi, T. S. Rappaport, and G. R. MacCartney, “Probabilistic omnidirectional path loss models for millimeter-wave outdoor communications,” IEEE Wireless Commun. Lett., vol. 4, no. 4, pp. 357–360, Aug. 2015.

[40] T. Bai, R. Vaze, and R. W. Heath, “Analysis of blockage effects on urban cellular networks,” IEEE Trans. Wireless Commun., vol. 13, no. 9, pp. 5070–5083, Sep. 2014.

[41] O. M. Teyeb, A. Muhammad, G. Mildh, E. Dahlman, F. Barac, and B. Makki, “Integrated access backhauling networks,” 2019. [Online]. Available: http://arxiv.org/abs/1906.09298