User Association and the Alignment-throughput Tradeoff in Millimeter Wave Networks

Hossein Shokri-Ghadikolaei†, Yuzhe Xu†, Lazaros Gkatzikis‡, and Carlo Fischione†
†Electrical Engineering, KTH Royal Institute of Technology, 100 44, Stockholm, Sweden
‡Mathematical and Algorithmic Sciences Lab, France Research Center, Huawei Technologies Co. Ltd., Paris, France
emails: {hshokri, yuzhe, carlofi}@kth.se, lazaros.gkatzikis@huawei.com

***Abstract*—Millimeter wave (mmWave) communication is a promising candidate for future extremely high data rate, wireless networks. The main challenges of mmWave communications are deafness (misalignment between the beams of the transmitter and receiver) and blockage (severe attenuation due to obstacles). Due to deafness, prior to link establishment between a client and its access point, a time consuming alignment/beam training procedure is necessary, whose complexity depends on the operating beamwidth. Addressing blockage may require a reassociation to non-blocked access points, which in turn imposes additional alignment overhead. This paper introduces a unifying framework to maximize network throughput considering both deafness and blockage. A distributed auction-based solution is proposed, where the clients and access points act asynchronously to achieve optimal association along with the optimal operating beamwidth. It is shown that the proposed algorithm provably converges to a solution that maximizes the aggregate network utility within a desired bound. Convergence time and performance bounds are derived in closed-forms. Numerical results confirm superior throughput performance of the proposed solution compared to existing approaches, and highlight the existence of a tradeoff between alignment overhead and achievable throughput that affects the optimal association.**

I. Introduction

Millimeter wave (mmWave) communication technology is a promising candidate to enable new applications with extremely high data rate requirements such as virtual reality, wireless backup connections in data centers, and mobile backhauling. The main characteristics of mmWave systems are small wavelength, severe path-loss, vulnerability to obstacles, and large number of antenna elements both at the transmitter and at the receiver [1]–[4].

Beamforming using a large number of antenna elements is the key technique to completely compensate the huge path-loss of mmWave communications and to establish a link with sufficiently large signal-to-noise ratio (SNR). In general, there are three main beamforming architectures: digital, analog, and hybrid [5]. Digital beamforming considers one RF chain per antenna element, which imposes unaffordable complexity and cost to mmWave networks that generally accommodated large number of antenna elements in each transceiver. Analog beamforming uses an array of phase shifters, connected to only one RF chain to achieve enough directivity gain and compensate severe path-loss. Due to its simplicity, current mmWave standards, namely IEEE 802.15.3c and IEEE 802.11ad, adopt this beamforming architecture [6], [7]. The main disadvantage of analog beamforming is the lack of multiplexing gain, which substantially reduces the achievable throughput in mmWave networks. Hybrid analog-digital beamforming is a promising architecture for mmWave networks, allowing the use of a very large number of antenna elements with a limited number of RF chains. The analog beamformer provides the required directivity gain, and the digital beamformer provides multiplexing gain [5]. The key point is that the digital beamformer is applied for the effective channel, consisting of the analog beamforming weights and the actual channel matrix, which has much smaller dimension than the actual channel, significantly reducing complexity and cost of the digital beamformer. From the available proposals in the study group of IEEE 802.11ay, which deals with the development of next generation 60 GHz mmWave networks, analog beamforming of the legacy mmWave standards will be replaced by hybrid beamforming.[1]

At the medium access control (MAC) layer, directional communication entails the *deafness* problem [3], which results from the misalignment between the beams of the transmitter and the intended receiver. IEEE 802.15.3c and IEEE 802.11ad address deafness through a two-stage beam training procedure [2]. Initially, a coarse grained sector-level sweep is performed, followed by a beam-level alignment phase. An exhaustive search over all possible transmission and reception directions is applied in each level. The beam training alleviates the complexity of traditional digital beamforming [3], but introduces a new alignment overhead, i.e., the time required to identify the best beams. This overhead depends on the number of directions that have to be searched, which in turn depend on the selected transmission and reception beamwidths (operating beamwidths). On the other hand, pencil beams promise significant spatial gains [8], [9], and therefore affect the optimal resource allocation decision.

Besides beam training, efficient association is particularly important in mmWave networks due to the smaller size of cells, denser access points (AP) deployments, and frequent handovers [3]. Given that association directly affects the long-term resource allocation [10], it has attracted the research interest in general [10]–[12] and in millimeter wave networks [13]–[15] in particular. The current mmWave standards use minimum-distance association, translated to association based on the maximum received signal strength indicator (RSSI) [6], [7]. Poor performance of RSSI-based association in the presence of non-uniform spatial distribution of clients, and heterogenous APs with different number of antenna elements and different

[1] Detailed information can be found in http://www.ieee802.org/11/Reports/ng60_update.htm. IEEE 802.11ay was approved in May 2015, and the study group has not released any stable document so far.

transmission powers is shown in [10]–[15]. Unfortunately, alternative association procedures proposed in the aforementioned literature do not consider the special characteristics of mmWave networks, namely blockage and directionality. Random blockage increases the frequency of reassociation, and hence introduces the need for fast and low-complexity association approaches. The optimal directionality level (beamwidth) is another important missing aspect of the existing association proposals. Notice that the operating beamwidth determines the complexity of mmWave link establishment and also the achievable transmission rate for each mmWave link. Moreover, current association solutions of mmWave networks are developed only for analog beamforming, further limiting their applicability to future mmWave networks based on hybrid beamforming.

In this paper, the problem of jointly optimizing operating beamwidth and association of the clients to APs in a hybrid beamforming architecture, is investigated. We first capture the interplay between beam training and achievable throughput of a mmWave link between a client and an AP. Then, we formulate a network throughput maximization problem by adopting the optimal association and operating beamwidth for all clients and APs. The resulting optimization problem is combinatorial and non-convex. Thus, we transform it to a multi-assignment problem, for which we develop a low-complexity distributed auction algorithm. In the proposed algorithm, the clients and APs act asynchronously to reach the optimal association and to find the optimal operating beamwidth. We show the convergence of the proposed algorithm to the optimal solution that maximizes network throughput within a desired bound. We also analyze the convergence time of the proposed auction algorithm. We conclude that mmWave networks can achieve significant throughput enhancements compared to the standard approaches that do not consider blockage, directionality, or optimization over the operating beamwidth (and its consequences such as alignment-throughput tradeoff).

The rest of the paper is organized as follows: In Section II, we introduce the system model and formulate the joint beamwidth selection and association optimization problem. In Section III, we present the proposed solution approach. Numerical results are reported in Section IV, followed by concluding remarks in Section V.

II. System Model and Problem Formulation

Consider a mmWave network consisting of a set of APs and clients, all operating at the mmWave frequency bands. We consider a hybrid beamforming architecture for each AP i with n_i RF chain (analog beams) to support both directivity and multiplexing gains [3]. To mathematically model the association problem with hybrid beamforming, we replace each AP i with n_i virtual APs, located at the same position, each having one RF chain. For sake of simplicity, we use “AP” to denote a virtual AP (i.e., an RF chain of an AP) for the rest of the paper. We denote by $\mathcal{C}$ the set of clients, by $\mathcal{A}$ the set of all APs, by p the transmission power of an AP, by σ the power of white Gaussian noise, and by g_{ij}^c the channel gain between AP i and client j, capturing long-term components of channel attenuation such as distance-dependent path loss. Here, similarly to [11], we neglect the impact of fast fading on the received signal and consequently on the signal-to-interference-plus-noise ratio (SINR), as the association will be executed on a large time scale compared to the instantaneous channel fluctuations, and therefore fast fading components can be averaged out. This channel model is common in the design of long-term resource allocation procedures [10], [11]. We also assume that multiuser interference is negligible, which is a valid assumption for small- to modest-sized mmWave networks [8], [9].

Fig. 1. Illustration of virtual time slot.

We define a virtual time slot, shown in Fig. 1, as the time difference between two consecutive executions of the association optimization problem. In each virtual time slot, a new association is computed and announced to the clients. Then, each client will establish a link towards its intended AP, which imposes an alignment (beam training) overhead [16]. In this paper, we assume that the APs (clients) are aware at which sector their associated clients (APs) are located, prior to alignment phase. This assumption is valid for use cases with low-mobility clients (e.g., in pedestrian scenarios). During the alignment phase, AP i and client j have to decide on the optimal refined beams within their sectors, by searching over all possible combinations, through a sequence of pilot transmissions. Let T_p denote the time required for a pilot transmission, which has to be performed for every combination of beam vectors of the client and AP, and ψ_i^a, φ_i^a, ψ_j^u, and φ_j^u be sector-level and beam-level beamwidths at the AP i and client/user j and receiver sides of link i, respectively. For clarity, we will use superscript a to denote the parameters related to an AP and u for users/clients. Under the exhaustive search approach adopted by existing mmWave standards [6], [7], the total duration of the alignment phase between AP i and client j, denoted by τ_{ij}, can be approximated by [16]

$$\tau_{ij}\left(\varphi_i^a, \varphi_j^u\right) = T_p \frac{\psi_i^a \psi_j^u}{\varphi_i^a \varphi_j^u} \, . \tag{1}$$

Upon completion of the beam training phase, the communication link is established, and the data transmission phase starts. Here, we have assumed that there is no outage after the alignment phase. Alignment time τ_{ij} should not exceed total virtual time slot duration T, so

$$\varphi_i^a \varphi_j^u \geq \frac{T_p}{T} \psi_i^a \psi_j^u \, . \tag{2}$$

Besides, since alignment takes place within the sector-level beamwidths, we have $\varphi_{i,\min}^a \leq \varphi_i^a \leq \psi_i^a$ for all $i \in \mathcal{A}$ and $\varphi_{j,\min}^u \leq \varphi_j^u \leq \psi_j^u$ for all $j \in \mathcal{C}$, where $\varphi_{i,\min}^a$ and $\varphi_{j,\min}^u$ are the minimum possible operating beamwidth of AP i and client j, respectively. The minimum possible operating beamwidth depends on the number of antenna elements and antenna configurations [17].

Let g_{ij}^a be the directivity gain that AP i contributes to the link between AP i and client j (transmission gain), and g_{ij}^u be the directivity gain that client j contributes to the link between AP i and client j (reception gain). At the MAC layer, we can approximate the actual antenna pattern by a sector antenna model with negligible sidelobe emission, where the directivity gain is a constant for all angles in the main lobe and zero otherwise, thus $g^a_{ij} = 2\pi/\varphi^a_i$ and $g^u_{ij} = 2\pi/\varphi^u_j$. Then, the power received by client j from AP i is $pg^a_{ij}g^c_{ij}g^u_{ij}$. Assuming negligible multiuser interference, the SNR at client j due to the transmission of AP i is

$$\mathrm{SNR}_{ij} = \frac{pg^a_{ij}g^c_{ij}g^u_{ij}}{\sigma}\,, \tag{3}$$

which depends on the transmission power p, operating beamwidths φ^a_i and φ^u_j, and network topology. Link between AP i and client j can support a maximum transmission rate of $\log_2(1 + \mathrm{SNR}_{ij})$ for the remaining $T - \tau_{ij}$ seconds. Therefore, achievable throughput of client j from AP i per virtual time slot, denoted by c_{ij}, is

$$c_{ij} = \left(1 - \frac{\tau_{ij}}{T}\right)\log_2(1 + \mathrm{SNR}_{ij})\,, \tag{4}$$

where τ_{ij} and SNR_{ij} are given in (1) and (3), respectively. Narrower transmission and reception beamwidths lead to higher SNR and hence higher data rate. This gain is obtained at the cost of higher alignment time that leaves less time for the data transmission phase. This reveals a tradeoff between the time devoted to the alignment phase and the effective data rate, captured in (4).

Let x_{ij} be a binary association variable, equal to 1 if and only if client j is associated to AP i. Given that the network topology is known a priori (that is, g^c_{ij} are known for every AP i and client j), the optimal association attempts to find the optimal values for φ^a_i, φ^u_j, x_{ij} to maximize some network utility. In this paper, we consider network throughput maximization. Although fairness and other network utilities can also be considered [11], we use network throughput maximization to provide a benchmark for the network throughput performance under user association and beamwidth selection in mmWave networks. We collect all control variables x_{ij} in matrix $\boldsymbol{X}$, and collect all variables φ^a_i and φ^u_j in vectors $\boldsymbol{\phi}^a$, $\boldsymbol{\phi}^u$. The optimal association problem can be formally stated as

$$\begin{aligned}
\max_{\boldsymbol{\phi}^a,\boldsymbol{\phi}^u,\mathbf{X}} \quad & \sum_{j\in\mathcal{C}}\sum_{i\in\mathcal{A}} c_{ij}x_{ij} && && \text{(5a)}\\
\text{s.t.} \quad & \sum_{j\in\mathcal{C}} x_{ij} \le 1, && \forall i \in \mathcal{A}\,, && \text{(5b)}\\
& \sum_{i\in\mathcal{A}} x_{ij} \le 1, && \forall j \in \mathcal{C}\,, && \text{(5c)}\\
& \psi^a_i\psi^u_j T_P/T \le \varphi^a_i\varphi^u_j, && \forall i \in \mathcal{A}, j \in \mathcal{C}\,, && \text{(5d)}\\
& \varphi^a_{i,\min} \le \varphi^a_i \le \psi^a_i\,, && \forall i \in \mathcal{A}\,, && \text{(5e)}\\
& \varphi^u_{j,\min} \le \varphi^u_j \le \psi^u_j\,, && \forall j \in \mathcal{C}\,, && \text{(5f)}\\
& x_{ij} \in \{0,1\}\,, && \forall i \in \mathcal{A}, j \in \mathcal{C}\,. && \text{(5g)}
\end{aligned}$$

Observe that, for notational simplicity, function arguments have been discarded. Constraint (5b) guarantees that AP serves at most one user, since each AP has only one RF chain, while constraint (5b) guarantees association to only one AP, mitigating joint scheduling requirements among APs. The solution of optimization problem (5) provides a long-term association policy and the optimal operating beamwidth for mmWave networks. This solution is valid as long as the inputs of the optimization problem, i.e., network topology, remain unchanged. Upon observing any change in the input (e.g., due to appearance of an obstacle), optimization problem (5) needs to be re-executed, and new associations will be announced to the clients. In the following section, we develop efficient solution method based on auction algorithms to solve this optimization problem.

III. Distributed Solution Approach

To solve optimization problem (5), we decompose it into two subproblems: (1) optimizing over operating beamwidth $\boldsymbol{\phi}^b$ and $\boldsymbol{\phi}^u$ for all pairs of clients/APs to find the optimal c^*_{ij}, and (2) substituting c^*_{ij} into a simplified optimization problem, formulated in the following, to find the optimal association. Since we consider negligible multiuser interference, joint optimization over operating beamwidths of all links is reduced to individual optimization over single pair of client-AP, which is solvable by a simple gradient descent algorithm [16]. In the following, we develop a solution method for the second subproblem. Given c^*_{ij} for all $i \in \mathcal{A}$ and $j \in \mathcal{C}$, optimization problem (5) is reduced to

$$\begin{aligned}
\max_{\mathbf{X}} \quad & \sum_{j\in\mathcal{C}}\sum_{i\in\mathcal{A}} c^*_{ij}x_{ij} && \text{(6a)}\\
\text{s.t.} \quad & \sum_{j\in\mathcal{C}} x_{ij} \le 1, \quad \forall i \in \mathcal{A}\,, && \text{(6b)}\\
& \sum_{i\in\mathcal{A}} x_{ij} \le 1, \quad \forall j \in \mathcal{C}\,, && \text{(6c)}\\
& 0 \le x_{ij}, \qquad \forall i \in \mathcal{A}, j \in \mathcal{C}\,. && \text{(6d)}
\end{aligned}$$

Note that optimization problem (6) is an asymmetric assignment problem in which relaxing constraint (5g) to constraint (6d) will not affect the optimality of the final solution [18].

To develop a distributed solution approach for asymmetric assignment problem (6), we convert it into a typical minimum cost flow problem following the methodology proposed in [18]. We replace *max* with *min*, reversing in parallel the sign of c^*_{ij}, and we introduce a virtual *supersource client* s that is connected to each AP i through a zero cost artificial arc (s, i) with feasible flow range $[0, \infty)$ for each AP i. The supersource client s generates traffic equal to $M - N$ units while the supply for each remaining client is of one unit, where M and N are the numbers of APs and clients, respectively. As a consequence, one unit of traffic is the output of each AP i. We refer the reader to [18, Section 7.2.2] for more details. The resulting problem is

$$\begin{aligned}
\min_{\mathbf{X}} \quad & \sum_{i\in\mathcal{A}}\sum_{j\in\mathcal{C}} -c^*_{ij}x_{ij} && && \text{(7a)}\\
\text{s.t.} \quad & \sum_{i\in\mathcal{A}} x_{ij} = 1\,, && \forall j \in \mathcal{A}\,, && \text{(7b)}\\
& \sum_{j\in\mathcal{C}} x_{ij} + x_{is} = 1\,, && \forall i \in \mathcal{A}\,, && \text{(7c)}\\
& \sum_{i\in\mathcal{A}} x_{is} = M - N\,, && && \text{(7d)}\\
& 0 \le x_{ij}, x_{is}\,, && \forall i \in \mathcal{A}, j \in \mathcal{C}\,, && \text{(7e)}
\end{aligned}$$

where the decision variables x_{ij} are extended to include also s. By using the terminology of network optimization, x_{ij} has the meaning of amount of flow between i and j. The first two constraints ensure that the flow *supply* of each client j is one unit, and a flow of one unit will reach every AP i respectively. The third constraint declares that s is the supersource client and the flow that generates is of $M-N$ units. The last two constraints declare that the flow of each arc may be infinite, where an arc between i and j denotes the connection (i,j). A solution to the minimum cost flow problem (7) is the same to the initial problem (6) [18, Section 7].

By using the duality theory for the minimum cost flow problems [18, Section 4.2], we formulate the dual problem

$$\min_{p_i,\pi_j,\lambda} \quad \sum_{j\in\mathcal{C}} \pi_j + \sum_{i\in\mathcal{A}} p_i - (M-N)\lambda \tag{8a}$$

$$\text{s.t.} \quad \pi_j + p_i \geq c^*_{ij}\,, \qquad \forall i \in \mathcal{A}, j \in \mathcal{C}\,, \tag{8b}$$

$$\lambda \leq p_i\,, \qquad \forall i \in \mathcal{A}\,. \tag{8c}$$

where $-\pi_j$ is the Lagrangian multiplier introduced to represent the price (or benefit due to the negative sign) of each client j, p_i represents the price for AP i and λ the price for s. The optimal solution to problem (8) allows us to derive the optimal solution to (6) [18, Sections 4.2, 5]. In order to solve (8) we need some technical intermediate results. We start by giving the definition of ϵ-*Complementary Slackness* (ϵ-CS)

Definition ϵ-CS: Let ϵ be a positive scalar. An assignment $\mathcal{S}$ and a pair (π, p) are said to satisfy ϵ-CS if

$$\pi_j + p_i \geq c^*_{ij} - \epsilon\,,\ \forall i \in \mathcal{A}, j \in \mathcal{C}\,,$$
$$\pi_j + p_i = c^*_{ij}\,,\ \forall (i,j) \in \mathcal{S}\,,$$
$$p_k \leq \min_{i\in\mathcal{A}(\mathcal{S})} p_i\,,\ \forall k \notin \mathcal{A}(\mathcal{S})\,,$$

where the set $\mathcal{A}(\mathcal{S}) = \{i | (i,j) \in \mathcal{S}, \forall i\}$ contains all the APs assigned under assignment $\mathcal{S}$.

Proposition 3.1: Consider problems (6) and (8). Let $\mathcal{S}$ be a feasible solution for problem (6) and consider a dual variable pair (π, p). If ϵ-CS conditions are satisfied by $\mathcal{S}$ and (π, p), then $\mathcal{S}$ is within $N\epsilon$ of the optimal solution of optimization problem (6).

Proof: The proof is similar to the proof [18, Proposition 7.7]. ■

Based on Proposition 3.1, we are now in the position to present the solution method to problem (8) in the form of a fully distributed auction mechanism. The distributed solution method is based on the application of Algorithm 1a by the clients, and Algorithm 1b by the APs. In the solution algorithms, the vector $P_j \in \mathbb{R}^M$ denotes the prices vector for the APs (stored in client j), p_i denotes the price of AP i (stored in AP i). In what follows we present the basic steps to establish the distributed algorithms.

Initially the prices of all APs are set to zero in both algorithms. On the clients side (Algorithm 1a), every client j fulfilling condition in Line 7 finds the best AP i_j using the local knowledge of the prices. In Lines 8∼11, client j calculates the largest bid for AP i_j. Then, it sends the request to AP i_j. On the APs side (Algorithm 1b), when AP i_j receives the request from clients with different bids, it chooses the best

Algorithm 1a Distributed Auction Algorithm for Client j

1: Initialize $i_j = \varnothing, P_j = \mathbf{0}$
2: **while true do**
3: **if** receive **no** and new price p_{i_j} from i_j **then**
4: Disconnect from AP i_j ▷ Released Connection
5: $[P_j]_{i_j} \leftarrow p_{i_j}$, and $i_j \leftarrow \varnothing$
6: **end if**

7: **if** $i_j = \varnothing$ **then** ▷ Distributed Auction for client
8: $i'_j \leftarrow \arg\max_{i\in\mathcal{A}} \left\{c^*_{ij} - [P_j]_i\right\}$,
9: $u_j \leftarrow \max_{i\in\mathcal{A}} \left\{c^*_{ij} - [P_j]_i\right\}$,
10: $\omega_j \leftarrow \max_{i\in\mathcal{A}\setminus\{i'_j\}} \left\{c^*_{ij} - [P_j]_i\right\}$,
11: $b_{i_j j} \leftarrow p_{i_j} + u_j - \omega_j + \epsilon$
12: Send request with $b_{i_j j}$ to AP i'_j, and wait response

13: Receive respond, (**yes** or **no**) and $p_{i'_j}$ ▷ After receiving response
14: **if** respond contains **yes then**
15: Connect to AP i'_j, and $i_j \leftarrow i'_j$
16: **end if**
17: $[P_j]_{i_j} \leftarrow p_{i'_j}$
18: **end if**
19: **end while**

Algorithm 1b Distributed Auction Algorithm for AP i

1: Initialize the client $j_i = \varnothing$, and price $p_i = 0$
2: **if** receive request from client j and b_j **then**
3: **if** $b_j - p_i \geq \epsilon$ **then** ▷ Respond to Connection Request
4: Send **yes** and p_i, to client j
5: Send **no** and p_i, to client j_i
6: Connect to client i, and $i_j \leftarrow i$, $p_i \rightarrow b_j$
7: **else**
8: Send **no** and p_i, to client j
9: **end if**
10: **end if**

client j_i that provides highest bid and higher price compared to the old price p_{i_j}. AP i_j updates its price and feedbacks the latest price to the clients as described in Lines 4∼6 and Line 8. The auction algorithm terminates when there are no client requests.

Proposition 3.2: Consider M APs, N clients. The distributed auction algorithms given in Algorithm 1 terminate within a finite number of iterations bounded by $NM^2\lceil\Delta/\epsilon\rceil$, where $\Delta = \max c^*_{ij} - \min c^*_{ij}$.

Proof: A proof is given in the extended version [19] and omitted from this paper to save space. ■

Proposition 3.3: Let ϵ be a desired positive constant. The final assignment obtained by Algorithm 1 is within $N\epsilon$ of the optimal assignment benefit of problem (6). The final assignment is optimal if all c^*_{ij} is integer, and $\epsilon < 1/N$.

Proof: A proof is given in the extended version [19] and omitted from this paper to save space. ■

IV. Numerical Results

In this section, we evaluate the performance of the proposed association solution. To this end, we simulate a mmWave network, with a set of APs and clients uniformly distributed at random in a 100x100 m^2 area. All APs and clients operate at 60 GHz and the operating bandwidth is 1.2 GHz. We assume 90° sector-level beams both at APs and clients. Transmission power is 0 dBm, ambient noise power is -83 dBm, and path loss exponent is 3. To measure the average performance of the algorithms, we simulate 1000 different topologies, each for

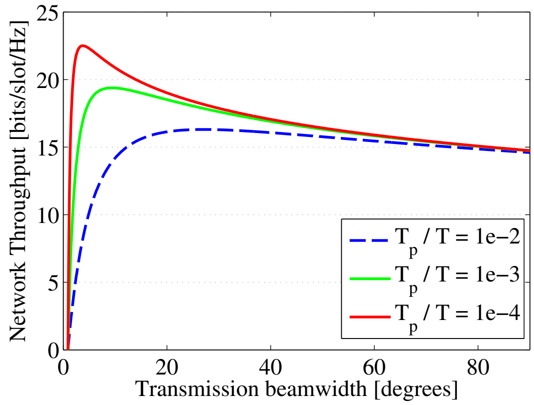

Fig. 2. Alignment-throughput tradeoff in mmWave networks.

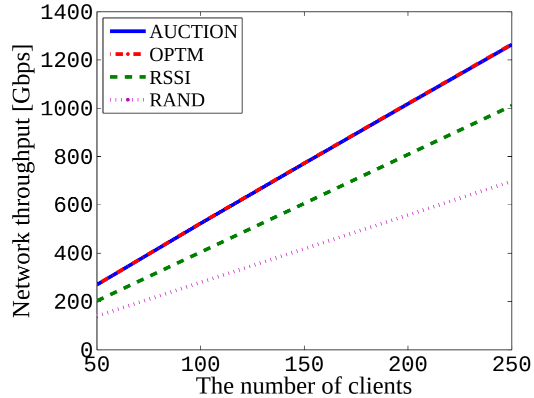

Fig. 3. Average throughput of AUCTION, OPTM, RAND, and RSSI, where $M - N = 50$.

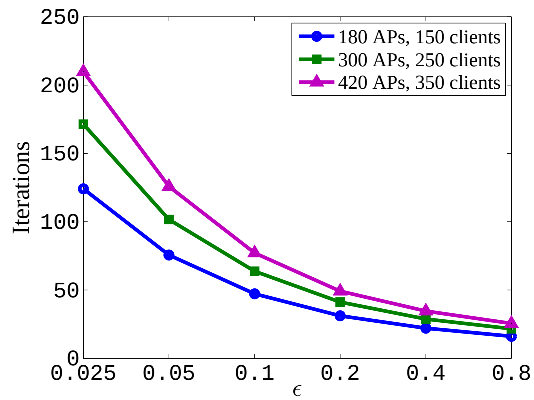

Fig. 4. Convergence performances of Algorithm 1 for problem (5), u, when ϵ varies.

500 consecutive virtual time slots, and evaluate the average performance. We use typical values for ϵ ($\epsilon = 0.1$) wherever it is not stated in the description. Obstacles may block the LOS link between clients and APs with probability 0.1 and for a duration of 1 second. We compare the performance of the developed association solution to the following approaches:

- RAND: Random association, where each client is randomly with uniform distribution associated to an available AP;
- (RSSI): RSSI-based association, where each client is associated to AP that can provide the highest RSSI; and
- (OPTIM): Optimal association, where the associations are the solution of optimization problem (5) using exhaustive search.

Fig. 2 demonstrates the alignment-throughput tradeoff for a single AP-client pair. For narrow beamwidths, beam training overhead is dominating, whereas as the operating beamwidths increase, directivity gain becomes more important. Moreover, reduced pilot transmission overhead T_p/T enables the AP and client to execute more beam training iterations using the same time budget. As a result, performance is improved, and narrower beams are more beneficial. The optimal operating beamwidth can be derived using a simple gradient descent algorithm [16], and will be used to transform optimization problem (5) to (6), so as to apply the proposed auction-based association approach.

Fig. 3 shows the average objective value (logarithmic scale of the network throughput averaged over all experiments) of optimization problem (5), obtained by the Algorithms 1 (AUCTION) after termination, in comparison to RAND, RSSI, and OPTM. Notice that the clients are uniformly distributed at random in the area covered by the existing APs. Results indicate that the distributed auction algorithm performs very close to the optimal policy, which is consistent with Proposition 3.3, and outperforms both RSSI and RAND approaches. In particular, the proposed method leads to 12% and 32% throughput enhancement compared to RSSI and RAND approaches, respectively.

Fig. 4 shows the average convergence time of the proposed distributed auction algorithms for different values of ϵ, used in the iterations of the proposed distributed algorithm. From this figure, it is evident that there is a noticeable effect of ϵ on the number of iterations till convergence. As stated by Proposition 3.2, the auction algorithms are faster for larger ϵ values. To elaborate, we depict deviation from the optimal solution (maximum network throughput) in Fig. 5. Increasing ϵ leads to faster convergence to a less accurate suboptimal solution of optimization problem (5), as studied in Proposition 3.3. The distance between the network throughput corresponds to the suboptimal association solution and that of the optimal association, denoted by $\Delta_{\max}$, increases with ϵ. Furthermore, $\Delta_{\max}$ is bounded by $N\epsilon$, which is consistent with Proposition 3.3.

V. Conclusion

In this paper, the problem of jointly optimizing operating beamwidth and user association to access points was investigated. The objective was to maximize the network throughput. The resulting optimization problem is combinatorial and nonconvex. By transforming it to a multi-assignment problem, a distributed algorithm auction algorithm was developed to solve the problem. The convergence time and optimality of the proposed algorithm were derived. It was shown that the proposed association algorithm results in significant throughput enhancements compared to the standard approaches that do not consider optimization over operating beamwidth or its consequences such as alignment-throughput tradeoff.

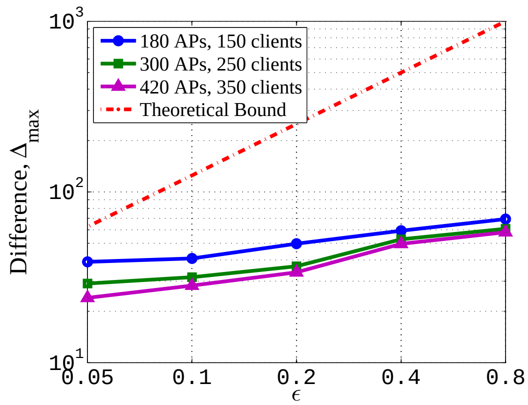

Fig. 5. Maximum distance from the optimal objective value of problem (5) to the resulting solutions obtained by Algorithm 1, when ϵ varies.

References

[1] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[2] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson Education, 2014.

[3] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Trans. Commun.*, 2015, to be published.

[4] H. Shokri-Ghadikolaei, C. Fischione, P. Popovski, and M. Zorzi, "Design aspects of short range millimeter wave wireless networks: A MAC layer perspective," *IEEE Netw.*, 2015, to be published.

[5] S. Sun, T. Rappapport, R. Heath, A. Nix, and S. Rangan, "MIMO for millimeter wave wireless communications: Beamforming, spatial multiplexing, or both?" *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 110–121, Dec. 2014.

[6] "IEEE 802.15.3c Part 15.3: Wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs) amendment 2: Millimeter-wave-based alternative physical layer extension," Oct. 2009.

[7] "IEEE 802.11ad. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications - amendment 3: Enhancements for very high throughput in the 60 GHz band," Dec. 2012.

[8] S. Singh, R. Mudumbai, and U. Madhow, "Interference analysis for highly directional 60-GHz mesh networks: The case for rethinking medium access control," *IEEE/ACM Trans. Netw.*, vol. 19, no. 5, pp. 1513–1527, Oct. 2011.

[9] H. Shokri-Ghadikolaei and C. Fischione, "The transitional behavior of interference in millimeter wave networks," *KTH Royal Institute of Technology, Tech. Rep.*, May 2015.

[10] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in HetNets: Old myths and open problems," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 18–25, Apr. 2014.

[11] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.

[12] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[13] A. Mesodiakaki, F. Adelantado, L. Alonso, and C. Verikoukis, "Energy-efficient user association in cognitive heterogeneous networks," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 22–29, Jul. 2014.

[14] G. Athanasiou, C. Weeraddana, and C. Fischione, "Auction-based resource allocation in millimeter-wave wireless access networks," *IEEE Commun. Lett.*, vol. 17, no. 11, pp. 2108–2111, Nov. 2013.

[15] G. Athanasiou, C. Weeraddana, C. Fischione, and L. Tassiulas, "Optimizing client association in 60 GHz wireless access networks," *IEEE/ACM Trans. Netw.*, vol. 23, no. 3, pp. 836–850, Jun. 2015.

[16] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione, "Beam-searching and transmission scheduling in millimeter wave communications," in *Proc. IEEE International Conference on Communications (ICC)*, 2015.

[17] C. A. Balanis, *Antenna theory: Analysis and design*. John Wiley & Sons, 2012.

[18] D. P. Bertsekas, *Network Optimization: Continuous and Discrete Models*. Belmont, Massachusetts: Athena Scientific, 1998.

[19] H. Shokri-Ghadikolaei, Y. Xu, and C. Fischione, "The interplay between alignment overhead and association in mmWave networks," *KTH Royal Institute of Technology, Tech. Rep.*, Aug. 2015, available upon request.