Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier xxxx

Multi-User Hybrid Beamforming Relying on Learning-Aided Link-Adaptation for mmWave Systems

K. SATYANARAYANA[1], (Student Member, IEEE), MOHAMMED EL-HAJJAR[1], (SENIOR MEMBER, IEEE), ALAIN MOURAD[2], AND LAJOS HANZO[1], (FELLOW, IEEE)

[1]University of Southampton, Southampton SO17 1BJ, U.K
[2]InterDigital Inc., London EC2A 3QR, U.K.

The financial support of the InterDigital and the financial support of the EPSRC projects EP/Noo4558/1, EP/PO34284/1, of the Royal Society's GRFC Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom is gratefully acknowledged.

ABSTRACT

Hybrid beamforming (HBF) relying on a large antenna array is conceived for millimeter wave (mmWave) systems, where the beamforming (BF) gain compensates for the propagation loss experienced. The BF gain required for a successful transmission depends on the user's distance from the base station (BS). For the geographically separated users of a multi-user mmWave system, the BF gain requirements of different users tend to be different. On the other hand, the BF gain is directly related to the number of antenna elements (AEs) of the array. Therefore, in this paper, we propose a HBF design for the downlink of multi-user mmWave systems, where the number of AEs employed at the BS for attaining BF gains per user is dependent on the user's distance. We then propose grouping of the RF chains at the BS, where each group of RF chains serves a specific group of users depending on the nature of the channel. Furthermore, to support the escalating data rate demands, the exploitation of link-adaptation techniques constitutes a promising solution, since the rate can be maximized for *each* link while maintaining a specific target bit error rate (BER). However, given the time-varying nature of the wireless channel and the non-linearities of the amplifiers, especially at mmWave frequencies, the performance of conventional link adaptation relying on pre-defined threshold values degrades significantly. Therefore, we additionally propose a two-stage link adaptation scheme. Specifically, in the first stage we switch on or off both the digital precoder and the combiner depending on the nature of the channel, while in the second stage a machine-learning assisted link-adaptation is proposed, where the receiver predicts whether to request spatial multiplexing- or diversity-aided transmission from the BS for every new channel realization. We demonstrate by simulation that having both a digital precoder and a combiner in a single dominant path scenario is redundant. Furthermore, our simulations show that the learning assisted adaptation provides significantly higher data rates than that of the conventional link-adaptation, where the reconfiguration decision is simply based on pre-defined threshold values.

INDEX TERMS Millimeter Wave, MIMO, Beamforming, Machine Learning.

NOMENCLATURE

ADC	Analog-to-Digital Converter
AE	Antenna Element
AoA	Angle-of-Arrival
AoD	Angle-of-Departure
AMC	Adaptive Modulation and Coding
BER	Bit Error Rate
BF	Beamforming
BPSK	Binary Phase Shift Keying
BS	Base Station
DAC	Digital-to-Analog Converter
ECR	Energy Consumption Ratio
HBF	Hybrid Beamforming
KNN	K-Nearest Neighborhood
LOS	Line-of-Sight
MIMO	Multiple-Input Multiple-Output
mmWave	Millimeter Wave

MU	Multi-User
NLOS	Non Line-of-Sight
NOMA	Non-Orthogonal Multiple Access
OFDM	Orthogonal Frequency-Division Multiplexing
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TPC	Transmit Precoder
ZF	Zero-Forcing

I. INTRODUCTION

GIVEN the dearth of spectral resources in the face of increasing data rate demands of mobile users in the sub-6 GHz band, harnessing millimeter wave (mmWave) frequencies has the benefit of large bandwidths to support high data rates [1]. However, an important challenge in harnessing mmWave frequencies is that they suffer from high propagation losses because of the attenuation imposed by atmospheric absorption, foliage density and rain-induced fading [1]. To mitigate the propagation losses, typically directional transmission is employed, where large antenna arrays are used to derive BF gain [2]. Conventionally, directional transmission is achieved by invoking digital signal processing elements relying on analog-to-digital/digital-to-analog converters (ADCs/DACs) for each of the RF chains. However, since large antenna arrays have to be employed at mmWave frequencies for attaining high BF gain, dedicating an ADC/DAC to each of the RF chains would impose high cost, complexity and power consumption. Hence, to circumvent the need for a large number of power hungry ADCs/DACs and to reduce the hardware complexity, a HBF design is conceived, where the signals generated by digital signal processing in the baseband relying on a few RF chains are fed to analog phase shifters in the RF stage before transmission from the antennas [3]–[5].

A vast body of literature is focused on transceiver designs for multi-user (MU) mmWave systems. Sayeed and Brady [6] proposed a transceiver architecture that exploits the concept of beamspace MIMO, where the data is multiplexed on orthogonal spatial beams. Liang *et al.* [7] conceived a low-complexity hybrid precoder for multiuser MIMO systems that is reminiscent of zero-forcing (ZF) precoding design. More particularly, Bogale *et al.* [8] determined the number of RF chains required at the BS for matching the performance of the potentially excessive-complexity full-RF digital BF solution for downlink multi-user mmWave systems. A beam domain reference signal was proposed by Han *et al.* [9] for downlink communications in order to maximize the BF gain in the desired direction.

On the other hand, to support the escalating data rate demands, typically link-adaptation is used to maximize the data rate while simultaneously meeting the bit error rate (BER) targets [10], where a threshold is used as a criterion to adapt the link based on the pre-defined modes of transmission. In other words, a BER versus rate look-up table is constructed for each of the legitimate transmission modes. Then the SNR after post-processing is compared against the average SNR threshold values, and the specific mode having the highest throughput as well as satisfying the BER target is activated. Prior work on link-adaptation, such as in [11]–[14], employs adaptation relying on hard threshold values of the average SNR. By contrast the authors of [15], [16] employ adaptation based on the rapidly-fluctuating time-variant channels. However, the ever-changing wireless channel and the non-linearities in the amplifiers erode the performance of conventional adaptation [17]. This is because the decision activated on the transmission scheme is based on the distorted threshold values due to the time-varying nature of the channel and owing to the non-linearities in the amplifiers. Hence, to circumvent the limitations of conventional link-adaptation, machine learning algorithms may be invoked based on the training data used for observation, regardless of the nature of imperfections imposed at the various processing stages [18]. In other words, no threshold values are used for making a decision on the transmission scheme activated. Instead, a more confident decision is made based on the model learned during the training stage.

In the literature, machine learning assisted algorithms have been studied in the context of adaptive modulation and coding (AMC). To increase the accuracy of link-adaptation, Daniels *et al.* [17] conceived a framework for overcoming the limitations of AMC aided MIMO-OFDM relying on supervised learning algorithms [19], such as the K-Nearest Neighborhood (KNN). In addition to KNN, Daniels *et al.* [20] also presented an online AMC, where support vector machines were employed. In [21], link adaptation has been proposed for single carrier frequency domain equalization, again using the KNN algorithm. More recently, a broader class of machine learning algorithms, namely deep learning methods have been applied in both the context of indoor localization [22] as well as in detection [23].

Against this backdrop, in this paper, we invoke a supervised learning based algorithm, where the decision/prediction is made based on the observation samples collected during the training phase. Both the BER and the instantaneous post processing SNR are taken as feature spaces to capture the channel conditions as well as the implementation losses imposed by the imperfections of the amplifiers. In our paper, the dimension of the feature space is 2. It is also instructive to note that if the dimension of the feature space is high, machine learning algorithms would result in erroneous solutions, unless they are provided with an exceedingly large number of training samples for decision-making [19].

On the other hand, in the mmWave HBF literature, analog BF is always combined with digital BF, regardless of the nature of the channel. Contrary to popular belief, we show in this paper that activating both a digital precoder and combiner when the channel has only a single dominant path is redundant. Therefore, switching off both the digital precoder

and the digital combiner and activating analog only BF is energy-efficient.

Against this background, our contributions are summarized as follows.

1) *We propose a transmitter design for MU-MIMO mmWave systems, where the users may be geographically separated both in terms of their distance and beam directions. More explicitly, we propose a user-specific allocation of the number of transmit AEs, where the number of active AEs for a user is decided based upon the user-specific BF gain required. We aim for creating this design by activating a subset of the phase shifters from the total number of available phase shifters. Hence, this philosophy makes our design energy-efficient, since not all the phase shifters are active at any given time, which is in contrast to the conventional multi-user mmWave systems, where all the phase shifters are active.*
2) *We then propose to appropriately group the RF chains at the BS in order to serve each user depending on the user's specific channel conditions. More explicitly, the number of RF chains that form a group at the BS to serve a single user depends on the number of possible beam directions for that user. However, the maximum number of RF chains that can be assigned for serving a user is limited in order to set aside a fair number of RF chains for the other groups serving different users.*
3) *Having grouped the RF chains, we then focus our attention on the per-user link design. More explicitly, we demonstrate that it suffices to have analog only BF in a scenario where there is a single dominant propagation path, which makes having any digital precoder/combiner redundant for that specific link.*
4) *We propose a learning assisted adaptive transceiver design for each user link based on the near-instantaneous post-processed SNR, where the adaptation switches between multiplexing versus diversity oriented transmission modes as well as by appropriately configuring the modulation employed so as to facilitate both high-reliability and high-rate operation. The receiver relies on the instantaneous post-processed SNR to decide on the transmitter's multiplexing versus diversity aided transmission mode and on the choice of the specific modulation scheme to be employed with the aid of supervised learning relying on the feed-back information forwarded to the BS. In this paper, we invoke the KNN classification technique at the receiver for decision making, as a design example.*
5) *We show through simulations that at a target BER of 10^{-3}, the learning-assisted adaptation achieves a significantly higher rate with a SNR gain of about 5 dB, while maintaining the required target BER compared to that of conventional link-adaptation carried out based on hard threshold values. Furthermore, we demonstrate by simulations that the performance of the system relying on HBF and on analog only BF is similar in an environment, where there is only a single dominant path for communications.*
6) *A qualitative discussion on the complexity of the KNN algorithm is presented.*

Notations: We use upper case boldface, $\mathbf{A}$, for matrices and lower case boldface, $\mathbf{a}$, for vectors. We use $(.)^T$, $(.)^H$, $\|.\|_{\mathrm{F}}$, $\mathrm{Tr}(.)$ $\mathbb{E}(.)$ for the transpose, Hermitian transpose, Frobenius norm, trace and expectation operator, respectively. We adopt $\mathbf{A}(m,n)$ to denote the m^{th} row and n^{th} column of $\mathbf{A}$, $\mathbf{I}_N$ is the identity matrix of size $N \times N$, and $\mathbf{A} \succ 0$ indicates that $\mathbf{A}$ is a positive definite matrix. Finally, we use $\mathcal{CN}, \mathcal{U}$, and i.i.d. to represent complex-valued normal distribution, uniform distribution, and independent and identical distribution, respectively.

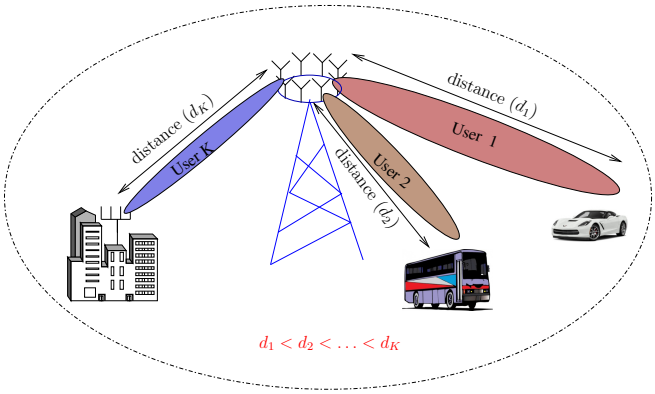


FIGURE 1: *System Model.*

The rest of the paper is organized as follows. Sec. II details the system model and HBF conceived for MU mmWave systems, while Sec. III discusses both the conventional and our learning-aided link-adaptation designs. Our simulation results and conclusions are presented in Sec. V and Sec. VI, respectively.

II. SYSTEM MODEL

In this section, we present the system model considered and describe the HBF employed at both the transmitter and receiver, followed by further discussions on the concept of link-adaptation in the context of our system model.

Let us consider the BS communicating with K users, each equipped with N_u antennas and N_u^{RF} RF chains, where the users may be geographically separated from each other, as shown in Fig. 1. In this design, the BS is equipped with N_t antennas and N_t^{RF} chains, where the BS processes the signal digitally using N_t^{RF} chains in the baseband and then the processed signal is phase shifted using $N_t^{\mathrm{RF}} N_t$ phase shifters in the radio-frequency (RF) stage before its transmission from the N_t antennas, as shown in Fig. 2. This design is referred to as fully-connected hybrid beamforming, where every RF chain of the design in Fig. 2 is connected to N_t AEs using N_t phase shifters. It is also important to emphasize that the attainable BF gain is dictated by the number of

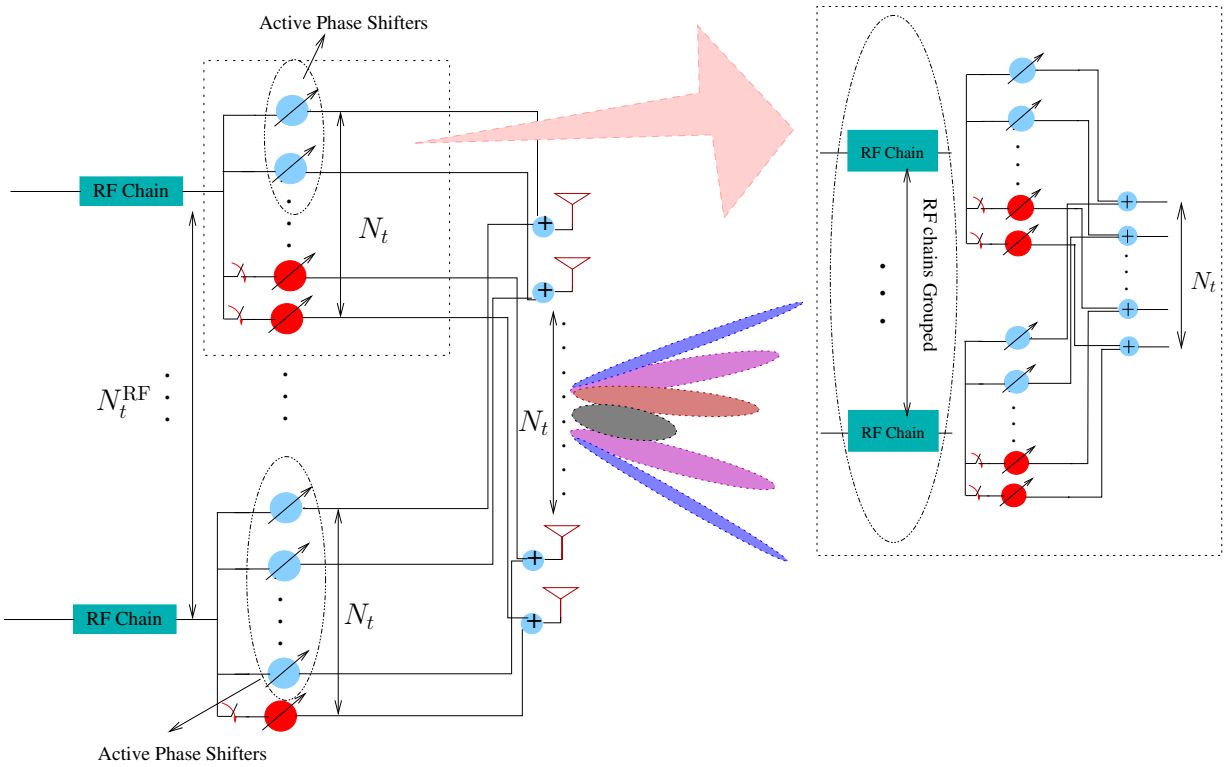


FIGURE 2: *Pictorial illustration of the number of active phase shifters at any time. Note that the design on the left side of the figure shows a single RF chain serving a user by employing a specific number of active phase shifters. However, there may be a group of RF chains that may serve multiple users that share similar channel conditions. The design on the right side of the figure shows the grouping of RF chains to serve a specific group of users. It can also be interpreted as a set of fully-connected RF chains with the same number of active phase shifters to serve a user .*

active AEs, which is equivalent to the number of active phase shifters, since the output of the phase shifters is fed to the AEs as shown in Fig. 2.

It is instructive to note that the aim of the BF in mmWave systems is to compensate for the propagation loss involved. Therefore, in our design of Fig. 2 the number of phase shifters active at any given time is distance-dependent. This is because the BF gain required to compensate for the propagation loss for each user may be different, since it is dependent on the user's distance from the BS. Let us consider Fig. 1 again as a 'toy' example. In this figure, user 1 and user 2 are located at distances of d_1 and d_2, respectively, from the BS, where $d_1 < d_2$. Since user 1 is closer to the BS, the propagation loss[1] experienced by the user 1 is lower than that of user 2, which is farther from the BS. As a result, the BF gain required to compensate for the loss is higher for user 2 than that of user 1. Therefore, the number of active phase shifters required at the BS for user 1 in order to compensate for the path loss is lower than the number of active phase shifters needed for user 2. This philosophy makes our design more energy-efficient than the conventional design where all the phase shifters remain active at a given time.

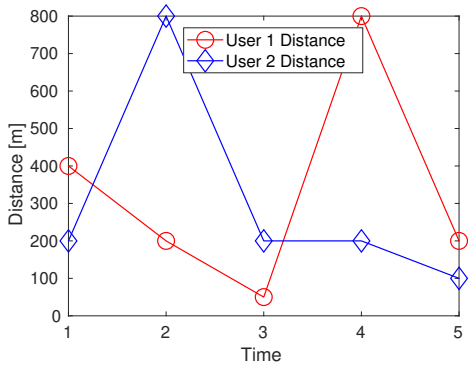
We also note that the beam of user 2 is narrower than that of user 1 as a benefit of having a higher-gain radiation pattern due to having more active phase shifters for user 2 so as to compensate for his/her higher propagation loss [1]. In other words, a high BF gain is achieved by activating a large number of phase shifters.

To expound further, let us consider Fig. 2, which shows the active phase shifters at the BS at a given point of time. Note that the BS seen in the figure shows a fully-connected design, where all the phase shifters are connected to all the transmit antennas and each RF chain is connected to all the phase shifters. It can be seen in the figure that the number of active phase shifters shown in blue color of the first and the last RF chains are different, since the BF gain required for the respective users is different.

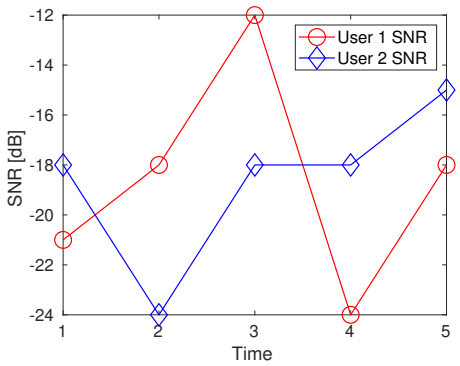
Having discussed the number of active phase shifters, we now focus our attention on the specific allocation of the RF chains, where more than one RF chain may be connected to the same number of active phase shifters[2] in a fully-connected fashion as shown at the right side of Fig. 2 [4], [5]. In other words, a plurality of RF chains grouped together may

[1]The propagation loss considered in this paper is free-space loss; however, in practice, other large-scale fading factors such as foliage density, attenuation due to rain-induced fading, and shadowing should also be considered. Our design will still work when considering all these factors, but we opted to focus on the free space loss for the sake of simplifying the discussion.

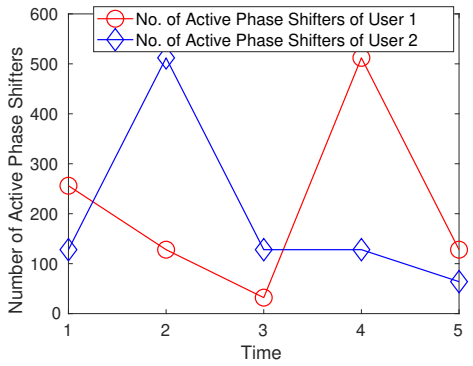
[2]Note that the rest of the phase shifters are switched off.



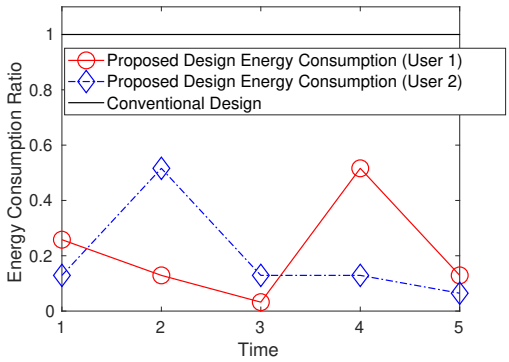
(a) *Distance of users' from the transmitter.*



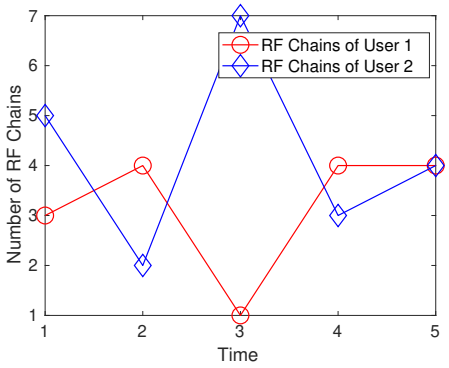
(b) *SNR of users' versus time.*



(c) *Number of phase shifters activated per user versus time.*



(d) *Energy consumption ratio of the proposed design with respect to the conventional design, where all the phase shifters remain active all the time, in terms of the usage of phase shifters.*



(e) Number of RF chains grouped to serve each user versus time.

FIGURE 3: *Graphical illustration of the variation in the number of active phase shifters depending on the user's distance and SNR, and of the the number of RF chains grouped to serve each user as a function of time.*

either serve a single user or a group of users. It is important to emphasize that first we assign a single RF chain per user at the BS and only the remaining RF chains will be distributed accordingly. More explicitly, the assignment of multiple RF chains to each user at the BS depends on the availability of beams. This is because the number of RF chains grouped together to serve a user is equal to the number of beams available for transmission to that user. More explicitly, if the user's channel is capable of supporting a single beam, then the number of RF chains in the group cannot be more than one. Hence, for the sake of fairness, the total number of RF chains, which is N_t^{RF}, at the BS is grouped in such a way that more RF chains are grouped for serving users having more possible beam directions in order to grant them additional degrees of freedom in the angular domain. However, the maximum number of RF chains in a group is limited, which will be discussed later in the paper. Note that this design can also be readily extended to sub-array-connected design [4],

[24].

To elaborate a little further, Fig. 3(a) illustrates the distance of users 1 and 2 from the BS terminal versus time. In this illustration, we assume that the BS is equipped with 1024 AEs and 32 RF chains, where each RF chain is connected to 1024 phase shifters[3]. It can be seen from Fig. 3(b) that as the users' distance of Fig. 3(a) increases from the BS, the SNR of each user decreases because of the propagation loss. This loss can be compensated by BF gain, where the number of activated phase shifters required to perform BF increases with the propagation loss. This becomes evident from Fig. 3(c), which illustrates the number of active phase shifters for the users of Fig. 3(a). As an example, let us consider user 1 of Fig. 3 (a) at time instant 4, where the user is 800 meters (m) away from the BS, while the SNR observed is -24 dB, as shown in Fig. 3(b). Accordingly, the number of active phase shifters needed to compensate for the path loss and to achieve an SNR of 3 dB is 512[4], as presented in Fig. 3(c). In other words, 512 of the 1024 phase shifters connected to a RF chain are required to achieve a BF gain of $10\log(512) \approx 27$ dB [1]. Similarly, observe for user 2, who is at a distance of 200 m, only 128 of the 1024 phase shifters have to be activated for achieving same SNR of 3 dB. In contrast to the conventional design, where all 1024 phase shifters are activated regardless of the users' distance, our design becomes more energy-efficient by appropriately adapting the number of active phase shifters, which will become explicit from Fig. 3(d).

It can be seen from Fig. 3(d) that the energy consumed by the users of our design is markedly lower than that of the conventional design, where all the phase shifters remain active all the time. More explicitly, in the conventional design, all phase shifters are used without considering the BF gain requirements, hence wasting energy owing to its higher-than-necessary BF gain. By contrast, our design activates exactly the required number of phase shifters, while attaining exactly the required BF gain. For example, we have seen that user 1 at time instant 4 would utilize 512 out of 1024, while user 2 utilizes only 128 out of 1024 phase shifters, which corresponds to 50% and 19%, respectively, of the energy consumed by the conventional design. The energy consumption ratio (ECR) is calculated as the ratio of the number of active phase shifters to the total number of phase shifters, given by

$$\text{ECR} = \frac{\text{Number of active phase shifters}}{\text{Total number of phase shifters}}. \quad (1)$$

Observe in Fig. 3(d) that energy consumption ratio of the conventional design is 1, while it varies for our design depending on the number of activated phase shifters.

Having judiciously activated the required number of phase shifters, the RF chains are arranged as groups to serve users depending on the channel conditions of the user. Let us again consider Fig. 3(e), where at time instant 4, the number of RF chains serving user 1 is 4 out of the available 32, while 3 out of the available 32 RF chains invoked for serving user 2. The choice in the number of RF chains that form a group is decided by the number of potential beam directions of the users. However, the maximum number of RF chains serving a user depends on the number of RF chains allocated to the preceding group. For example, let us assume there are 8 users and each user has 10 potential beam directions. Let us furthermore assume that there are 32 RF chains. In this scenario, for the sake of fairness, every user is served by 4 RF chains. On the other hand, for example, user 2 has 3 beams, which means that it can be served by a maximum of 3 RF chains, because having additional RF chains would be redundant[5]. So in this setting, user 4 may be served by 5 RF chains since the user has 6 more additional beams available for data transmission.

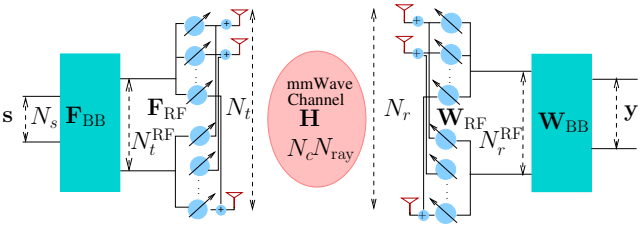


FIGURE 4: *Hybrid architecture.*

After allocating the phase shifters and RF chains to the users, each user link can be modeled relying on the system model of Fig. 4, where user[6] k receives its signal transmitted from N_{t_k} active antennas, which is then digitally processed using N_k^{RF} RF chains. Then the vector of received signal for user k is given by

$$\mathbf{y}_k = \mathbf{W}_{\text{BB}}^{k^H} \mathbf{W}_{\text{RF}}^{k^H} \mathbf{H}_k \mathbf{F}_{\text{RF}}^k \mathbf{F}_{\text{BB}}^k \mathbf{s} + \mathbf{W}_{\text{BB}}^{k^H} \mathbf{W}_{\text{RF}}^{k^H} \mathbf{n}, \quad (2)$$

where $\mathbf{n}_k$ is the Gaussian noise distributed as $\mathcal{CN} \sim (0, \sigma^2)$, $\mathbf{W}_{\text{BB}}^k$ is the baseband combiner of size $N_u^{\text{RF}} \times N_s^u$, $\mathbf{W}_{\text{RF}}^k$ is the analog RF combining matrix of size $N_u \times N_u^{\text{RF}}$, $\mathbf{F}_{\text{RF}}^k$ is the analog RF beamforming matrix of size $N_{t_k} \times N_k^{\text{RF}}$, and $\mathbf{F}_{\text{BB}}^k$ is the baseband precoder matrix of size $N_k^{\text{RF}} \times N_{s_k}$. We also note that N_k^{RF} is the number of RF chains grouped for serving user k, while $N_{t_k} (N_{t_k} \leq N_t)$ is the number of active phase shifter of the group k. Furthermore, $KN_u = N_r, KN_u^{\text{RF}} = N_r^{\text{RF}}$ and $KN_s^u = N_s$.

Additionally, $\mathbf{H}_k$ is the statistical spatial channel model given by

$$\mathbf{H}_k = \sqrt{\frac{N_u N_{t_k}}{N_c N_{\text{ray}}}} \sum_{n_c=1}^{N_c} \sum_{n_{\text{ray}}=1}^{N_{\text{ray}}} \alpha_{n_c}^{n_{\text{ray}}} \mathbf{a}_{r_u}(\phi_{n_c}^{n_{\text{ray}}}) \mathbf{a}_{t_k}^T(\theta_{n_c}^{n_{\text{ray}}}), \quad (3)$$

[3]The total number of phase shifters is equal to 1024*32= 32768.

[4]The mathematical relationship between the number of AEs required and the BF gain value is given by Eq. (13) of [1].

[5]The number of RF chains is equal to the number of beams for transmission.

[6]As a design example, we assumed K groups and each group has one user. However, that one user may be served by multiple RF chains provided that the user's channel supports multiple beams.

where $\mathbf{a}_{r_u}(\phi)$ and $\mathbf{a}_{t_k}(\theta)$ are the antenna response vectors at the angle of arrival ϕ and the angle of departure θ, respectively. To elaborate further, N_c is the number of clusters while N_{ray} is the number of rays in the cluster, and $\alpha_{n_c}^{n_{\text{ray}}}$ is the Rayleigh fading coefficient whose phase is uniform and amplitude is complex Gaussian distributed as $\mathcal{CN} \sim (0, 1)$ with mean 0 and variance 1.

It is important to emphasize that the architecture of Fig. 2 may be deemed to be equivalent to that of Fig. 4, where the BS of Fig. 2 processes the signals digitally in the baseband using a digital transmit precoder (TPC) matrix $\mathbf{F}_{\text{BB}}$ of size $N_t^{\text{RF}} \times N_s$ and then the digitally precoded signal is phase shifted using the RF beamformer matrix $\mathbf{F}_{\text{RF}}$ of size $N_t \times N_t^{\text{RF}}$ before transmission. Then the collective downlink received signal vector $\mathbf{y}$ after both RF and baseband processing using the constituent matrices $\mathbf{W}_{\text{RF}}$ and $\mathbf{W}_{\text{BB}}$ of sizes $N_r \times N_r^{\text{RF}}$ and $N_r^{\text{RF}} \times N_s$, respectively, is given by

$$\mathbf{y} = \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{n}, \tag{4}$$

where $\mathbf{n}$ is the Gaussian noise, $\mathbf{W}_{\text{RF}} = \text{diag}\left[\mathbf{W}_{\text{RF}}^1, \ldots, \mathbf{W}_{\text{RF}}^K\right]$, $\mathbf{W}_{\text{BB}} = \text{diag}\left[\mathbf{W}_{\text{BB}}^1, \ldots, \mathbf{W}_{\text{BB}}^K\right]$. Furthermore, $\mathbf{H} = \text{diag}\left[\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K\right]$ and $\mathbf{s}$ is the transmit symbol vector of size N_s, while $\mathbf{y} = [y_1 \ldots y_k \ldots y_K]^T$. Similarly, $\mathbf{F}_{\text{RF}} = \text{diag}\left[\mathbf{F}_{\text{RF}}^1, \mathbf{F}_{\text{RF}}^2, \ldots, \mathbf{F}_{\text{RF}}^K\right]$, while $\mathbf{F}_{\text{BB}} = \text{diag}\left[\mathbf{F}_{\text{BB}}^1, \mathbf{F}_{\text{BB}}^2, \ldots, \mathbf{F}_{\text{BB}}^K\right]$.

The flowchart of the proposed design is presented in Fig. 5. First the BF gain required for each of the K users is calculated, as detailed in [1]. Then we compute the number of AEs necessary to compensate for the propagation loss. Having obtained the number of active AEs required, the RF chains are appropriately distributed at the BS to serve individual users. However, first we assign a single RF chain per user at the BS and only then will the remaining RF chains be distributed, where the number of beam-pairs available for communication is decided based on the criterion discussed in Sec. II. Having allocated the RF chains, then link-adaptation is carried out for every user. The link-adaptation is a two-stage process for each user: switching on/off the digital precoder,followed by learning assisted AMC transmission.

Having discussed the allocation of phase shifters and RF chains, we now focus our attention to a single-link of (2), where the BS design its $\mathbf{F}_{\text{RF}}^k$ and $\mathbf{F}_{\text{BB}}^k$ as well as the modulation and transmission scheme depending on the nature of the channel.

A. HBF

The matrices $\mathbf{F}_{\text{RF}}^k$ and $\mathbf{F}_{\text{BB}}^k$ are designed by maximizing the capacity of the hybrid precoder. More explicitly, the objective function formulation results in minimizing the Frobenius norm between the optimal matrix and the matrix product $\mathbf{F}_{\text{RF}}^k \mathbf{F}_{\text{BB}}^k$ [25], which is formulated as:

$$\min_{\mathbf{F}_{\text{RF}}^k, \mathbf{F}_{\text{BB}}^k} \|\mathbf{F}_{\text{opt}}^k - \mathbf{F}_{\text{RF}}^k \mathbf{F}_{\text{BB}}^k\|_{\text{F}}, \tag{5}$$

$$s.t. \quad \|\mathbf{F}_{\text{RF}}^k \mathbf{F}_{\text{BB}}^k\|_{\text{F}}^2 = N_s, \tag{6}$$

$$|\mathbf{F}_{\text{RF}}^k(m, n)| = 1. \tag{7}$$

The optimal precoder matrix $\mathbf{F}_{\text{opt}}^k$ of the system model in (4) is obtained by the singular value decomposition (SVD) of the channel matrix $\mathbf{H}_k = \mathbf{U}_k \Sigma_k \mathbf{V}_k^H$, where the first N_s columns of the right singular matrix $\mathbf{V}_k$ are chosen to construct the matrix $\mathbf{F}_{\text{opt}}^k$.

The near-optimal solution of (5) when the columns of $\mathbf{F}_{\text{opt}}^k$ are the right singular vectors of the channel matrix $\mathbf{H}_k$ is given by Ghauch *et al.* [25]. However, since the focus of the paper is not on the decomposition of the optimal digital precoder matrix $\mathbf{F}_{\text{opt}}^k$ into its hybrid product, we have adopted the approach of Ghauch *et al.* [25] for the hybrid precoder decomposition. Similarly, the solutions for $\mathbf{W}_{\text{RF}}^k$ and $\mathbf{W}_{\text{BB}}^k$ are obtained by decomposing the left-singular vectors of the channel matrix $\mathbf{H}_k$, as in (5).

Traditionally, the link-adaptation is carried out based on the average threshold values. Upon receiving the signal, the receiver calculates the instantaneous post-processing SNR, based on which the receiver makes the decision concerning the most appropriate transmission mode by comparing it against the pre-defined average SNR threshold values [14].

The total post-processing SNR, which is the SNR calculated after combining using the matrix $\mathbf{W}_k$ at the receiver of user k, for a given channel realization $\mathbf{H}_k$ and noise variance σ_n^2 is given by

$$\text{SNR} = \frac{\text{tr}\left(\left(\mathbf{W}^{k^H} \mathbf{H}_k \mathbf{F}^k\right)^H \mathbf{W}^{k^H} \mathbf{H}_k \mathbf{F}^k\right)}{\text{tr}\left(\mathbf{W}^k \mathbf{W}^{k^H} \sigma_n^2\right)}, \tag{8}$$

where $\mathbf{W}^k = \mathbf{W}_{\text{RF}}^k \mathbf{W}_{\text{BB}}^k$ and $\mathbf{F}^k = \mathbf{F}_{\text{RF}}^k \mathbf{F}_{\text{BB}}^k$.

Remark 1: In a scenario where the users' beams are close to each other, interference leaked from one beam into another should be accounted for (2). This philosophy is akin to that of non-orthogonal multiple access (NOMA) systems. In this setting, the matrices $\mathbf{W}_{\text{BB}}^{k^H} \mathbf{W}_{\text{RF}}^{k^H}$ may be designed to minimize the interference, as detailed in [26], while the design of the matrices $\mathbf{F}_{\text{RF}}^k$ and $\mathbf{F}_{\text{BB}}^k$ is discussed in the next subsection.

B. IMPROVED ENERGY-EFFICIENT HBF

The HBF presented in the previous section does not consider the nature of the channel. However, it is important to emphasize that when the channel has only a single dominant path, employing both a digital precoder matrix $\mathbf{F}_{\text{BB}}$ and a combiner matrix $\mathbf{W}_{\text{BB}}$ is redundant, as we will show later in this paper. This is because when the channel has only a single dominant path, analog BF using phase shifters efficiently captures the signal. Mathematically, this is equivalent to setting the columns of $\mathbf{F}_{\text{BB}}$ and $\mathbf{W}_{\text{BB}}$ matrices from the Identity $\mathbf{I}$.

Therefore, in this section, we propose switching off the digital precoder and combiner when the channel has only a single dominant path for communication, since in this scenario analog only BF efficiently captures the signal [24] and having any digital processing in the baseband is redundant. Fig. 6 (a) shows the BER of the system when the channel has only a single dominant path. It can be seen in the figure

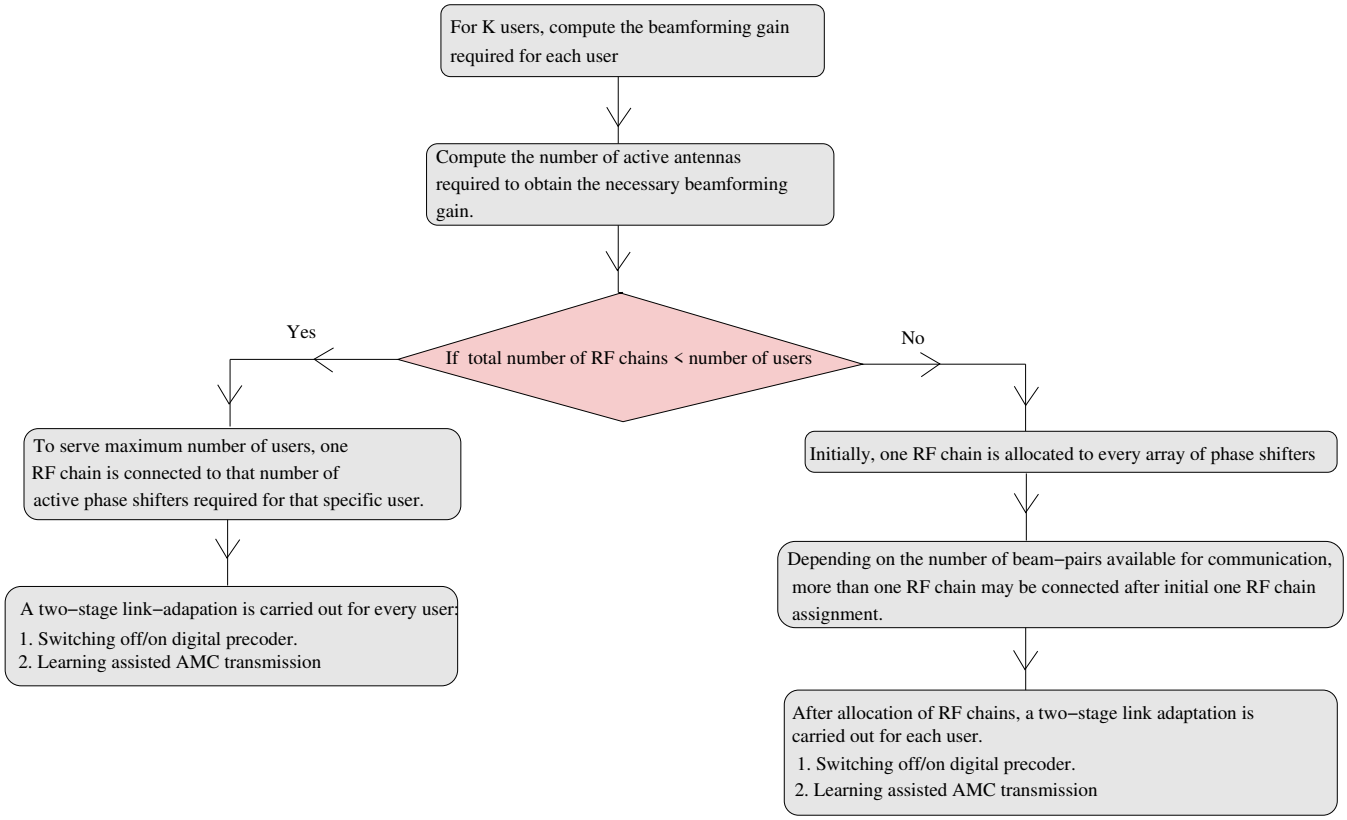


FIGURE 5: *Flowchart of the proposed design.*

that the performances of the system with both analog and digital BF and that of analog only BF is identical. Similarly, Fig. 6 (b) shows the achievable rate of system for both designs. It is evident from the figure that the rate of the system without digital processing is identical to that of the system combined with digital processing. In other words, the analog phase shifters in the RF steer and combine efficiently in the direction of the channel response vectors at the BS and receiver, respectively.

In the next section, we detail both the conventional and the proposed link-adaptation.

Remark 2: It is instructive to note that the system considered is not very different from the scenario of having unknown interferences. The proposed design can be readily extended to the situation when there is interference from the unintended transmitters by using the SINR instead of the SNR. In this setting the denominator in (8) of the revised manuscript would contain one more additional term, which is the interference. Moreover, according to the central limit theorem, when there is a large number of interferes, the interference caused by them can be modeled as additional Gaussian noise. In this case (8) would remain the same, except for an increased noise variance.

III. ADAPTIVE TRANSCEIVER DESIGN

In this section, we discuss the adaptive transceiver design proposed while contrasting it with the conventional link-adaptation. Note that this adaptation is followed after having made the decision on whether to switch off the digital precoder and combiner depending on the channel conditions.

A. CONVENTIONAL ADAPTATION

In this section, we first describe the adaptive system when the HBF is used. Then later we discuss the adaptation when there is only a single dominant path. In conventional adaptation, the receiver makes a decision concerning the mode of transmission using the post-processed SNR based on the pre-defined threshold SNR values, which are set to meet a required BER. Fig. 7(a) shows the plot of average BER against the average SNR for different transmission schemes using 64×32 element MIMO scheme relying on two RF chains at both the BS as well as at the receiver and communicating over the mmWave channel model of (3). In this plot, two spatial streams are used for spatial multiplexing, while only a single spatial stream is used when aiming for diversity. Observe from the plot that the diversity-oriented QPSK scheme performs better than spatial multiplexing using two BPSK streams. Furthermore, spatial multiplexing with two QPSK streams achieves better BER than diversity-aided 16 QAM. Therefore, Fig. 7(b) excludes both the inferior spatial multiplexing aided BPSK streams and the diversity aided 16 QAM, hence only considering the schemes that provide a better BER for a given rate.

Following the conventional link-adaptation, the specific

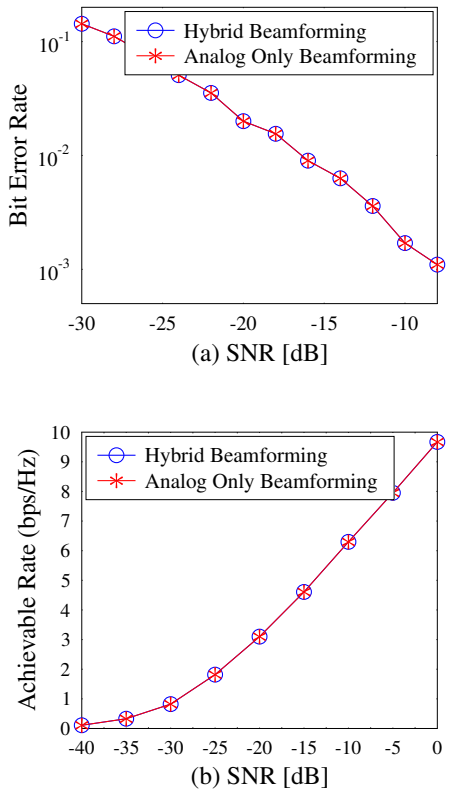


FIGURE 6: *BER and achievable rate with HBF (analog and digital) and using only analog BF, when the channel has only one dominant path.*

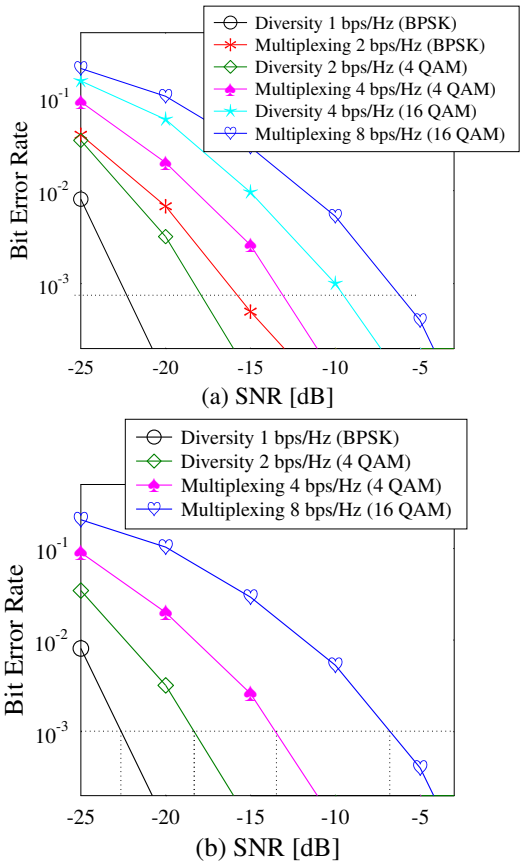


FIGURE 7: *Average BER versus average SNR of spatial diversity and spatial multiplexing for different transmission rates. In this configuration, two spatial streams are transmitted using a* 64×32 *element MIMO with* N_t^{RF} *and* $N_r^{RF} = 2$*, while channel is NLOS in nature.*

threshold values designed for each scheme characterized in Fig. 7(b) that attain the target BER of 10^{-3} are shown by vertical lines. After the calculation of the instantaneous post-processed SNR the receiver decides on both the type of transmission scheme as well as on the modulation mode by comparing it against the pre-defined threshold values. Explicitly, the receiver compares the post-processed SNR against the vertical lines in Fig. 7(b) and relays the requested mode information to the BS. We note that the post-processed SNR values are calculated offline and stored in a memory so that the receiver does not have to do any calculations.

On the other hand, in the case of only one dominant path, the digital BF is switched off and adaptation is performed amongst the modulation schemes only, since the channel does not support multiplexing/diversity.

B. PROPOSED LEARNING ASSISTED ADAPTATION

In contrast to conventional adaptation, in the proposed design, the receiver relies on learning, hence dispensing with any pre-defined threshold values. Here we conceive a classification algorithm. However, most classification algorithms require an explicit functional mapping between the feature set and the classifiers. The feature set in our work includes the SNR and the BER, while the classifiers are spatial multiplexing associated with different modulations modes and spatial diversity also using different modulations. Unfortunately, there is a paucity of information about the functional mapping between the two. Hence, we conceive a KNN classification algorithm, which is non-parametric, since it does not require any information about the functional mapping [18]. Furthermore, the theoretical assumptions made related to its mathematical tractability may become invalid in practical environments, where the data may not obey a specific distribution. In such scenarios, the KNN algorithm would be a promising choice, because again, it does not depend on any assumptions or knowledge about the data distribution [18]. The operating principle of the KNN algorithm is illustrated in Fig. 8, where the training data of two classifiers are distinguished with squares and circles. When a testing data point is given, a circle with the testing data point as its center is drawn so that it encircles K points from the classifiers, as shown in Fig. 8. Then the class associated with more points in the circle will be chosen by invoking majority voting. If

the number of points in each class is the same, then the class having a better throughput is selected. The feature set used

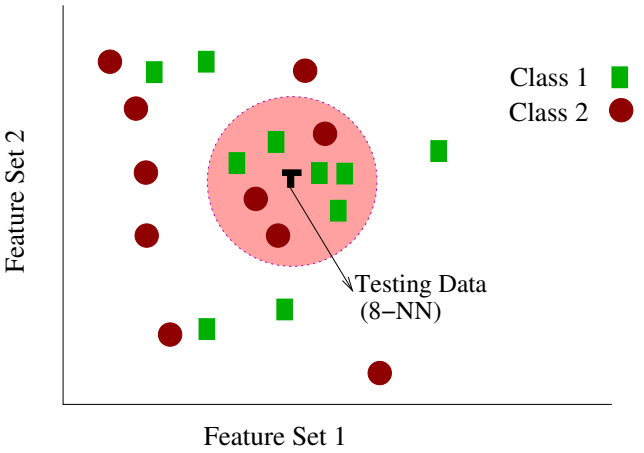


FIGURE 8: *Illustration of K-Nearest Neighborhood Algorithm.*

for link-adaptation is $\mathcal{F} = \{\text{SNR}, \text{BER}\}$, and the class set is defined as $C = \{\text{Div}_{\text{BPSK}}, \text{Div}_{\text{QPSK}}, \text{Mux}_{\text{QPSK}}, \text{Mux}_{16\text{QAM}}\}$[7] for NLOS channel, while $C = \{\text{QPSK}, 16\text{QAM}\}$ when the channel has only a single dominant path since it has no multiplexing/diversity gain. In other words, when the channel has only a single dominant path, adaptation during the second stage only reconfigures the modulation scheme.

The proposed learning adaptation is comprised of two phases: the training phase and the testing phase. In the training phase, both the BER and the instantaneous post-processing SNR are calculated for each channel realization and stored in memory. Having accumulated the training data, the testing phase ensues. When a new data point is received, the post-processed SNR is calculated. Then, equipped with the post-processed SNR and the required BER as the parameters, the K nearest neighbors are chosen from the set C. Finally, the specific class which has more points in the neighborhood is selected. To elaborate further, Fig. 9 shows the training data for 100 channel realizations at 5 different noise levels. This figure was obtained by plotting the *instantaneous BER versus instantaneous SNR using (8), as opposed to the average BER versus average SNR portrayed in* Fig. 7.

It is worth observing from Fig. 9 that the selection of the class from set C based on the average threshold values would result in low rates because of the wide-ranging scattering of the instantaneous post-processing SNR values. It can clearly be seen in Fig. 9 that the boundaries are not hard, which makes it difficult to decide the choice of class based on the threshold values of the conventional link-adaptation. During the testing phase, upon estimating the channel state information, the receiver calculates the post-processing SNR assuming that spatial multiplexing is used relying on (8) and then finds the K-nearest neighbors with the aid of the post-processed SNR calculated and the required BER, followed by selecting the class through majority voting.

[7]Div denotes diversity, while Mux denotes multiplexing.

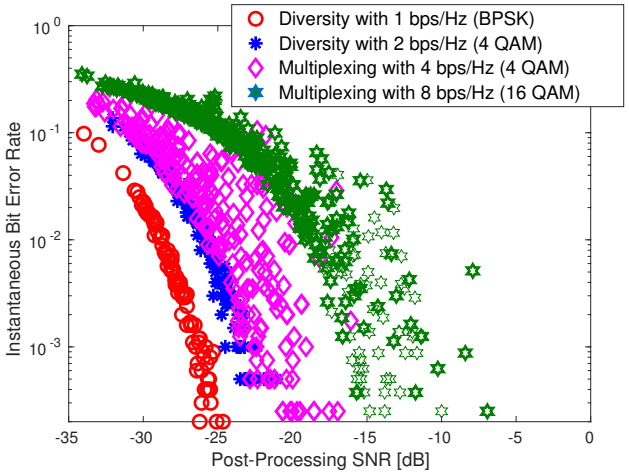


FIGURE 9: *Transmission scheme profiles for NLOS.*

Remark 3: When two different color points of Fig. 9 overlap, which means that the instantaneous BERs of the two schemes are similar, the specific transmission scheme having the higher rate is selected.

Remark 4: The learning algorithm invoked in this paper is a non-parametric classification method [18], which does not depend on the bandwidth, or sampling rate. To expound a little further, it only considers the post-processing SNR, as well as the target BER as the feature sets and then decides upon the specific class having the higher rate while satisfying the target BER for the post-processing SNR observed.

IV. COMPLEXITY

In this section, we present a qualitative discussion on the KNN algorithm's complexity in practical implementations. There are two significant components of the KNN, which dominate the complexity, namely the search complexity and the memory required for saving data points. Given the advances in the storage capabilities of the devices, the latter may be of less concern than the former. Hence, we focus our discussions on the search complexity. Let us assume that we have n training samples in d dimensions. Then a brute-force KNN search would have the search complexity of $O(kdn)$ for K nearest-neighbors. Although in our paper we only deal with the BER, rate and post-processing SNR, the brute-force search across the design-space of these three parameters results in a high complexity for higher dimensions. Therefore, this method is not effective as it does not exploit the structure of the training data. However, there is a vast body of literature on fast-KNN techniques search [27]–[29], which restructure the training data into clusters. By restructuring the data into clusters each having $O(\sqrt{n})$ objects in each cluster, the search complexity is reduced to $O(kd\sqrt{n}\log nd + kn)$, which is lower than the brute-force KNN search complexity [28], [29].

On the other hand, in conventional link-adaptation, a BER and rate look-up table is constructed for different average SNR values. Then the SNR after post-processing is compared against the average SNR values and the scheme whose SNR value is higher than the threshold is selected. Although the computational complexity in conventional link-adaptation is lower, the quality-of-the-service of the system is significantly affected.

V. SIMULATION RESULTS

In this section, we present our simulation results for characterizing the rate achieved by the proposed learning assisted link adaptation for a user k. We performed Monte Carlo simulations for studying the performance gap between the proposed design and the conventional design, where the average is calculated using 100 channel realizations and a total of 1000 symbols are used for each channel realization. Furthermore, in these results, the desired/required BER is set to 10^{-3}. The simulation parameters are listed in Table 1.

TABLE 1: Simulation parameters.

Parameters	*Values*
N_t	64
N_r	32
N_t^{RF}	2
N_r^{RF}	2
N_s	1, 2
K	20
$\alpha_{n_c}^{n_{\text{ray}}}$	$\mathcal{CN}(0,1)$
$\phi_{n_c}^{n_{\text{ray}}}$	Laplacian distributed
$\theta_{n_c}^{n_{\text{ray}}}$	Laplacian distributed
BER	10^{-3}

Fig. 10(a) shows the histogram-based probability density function (PDF) versus the average SNR for all the classes in the set C using learning. It can be seen from Fig. 10(a) that at -20 dB of post-processed SNR there is a similar probability of transmission for the classes of $\text{Div}_{\text{BPSK}}, \text{Div}_{\text{QPSK}}$. Furthermore, the probability of the class Div_{BPSK} falls gradually as the SNR increases, while the Mux_{QPSK} mode starts to share the probability of being activated with the Div_{QPSK} mode, where the probabilities are obtained by evaluating the relative frequency of each class using the proposed learning assisted link-adaptation based on both the post-processing SNR and on the BER target. Similarly, the probability of choosing the Mux_{QPSK} mode increases at high SNRs. By contrast, the conventional link-adaptation shown in Fig. 10(b) has a very low probability of selecting the class Div_{QPSK} in the SNR region $-20 < \text{SNR [dB]} < -15$. Similarly, Mux_{QPSK} has a low probability, when we have $-10 < \text{SNR [dB]} < -3$ as opposed to Fig. 10(a), where the decision is made based on the pre-defined threshold values.

Fig. 11(a) shows the throughput of the system for both the proposed as well as for the conventional link-adaptations. The curves in Fig. 11(a) are obtained by calculating the average number of bits transmitted per channel realization using learning and conventional link-adaptations, respectively.

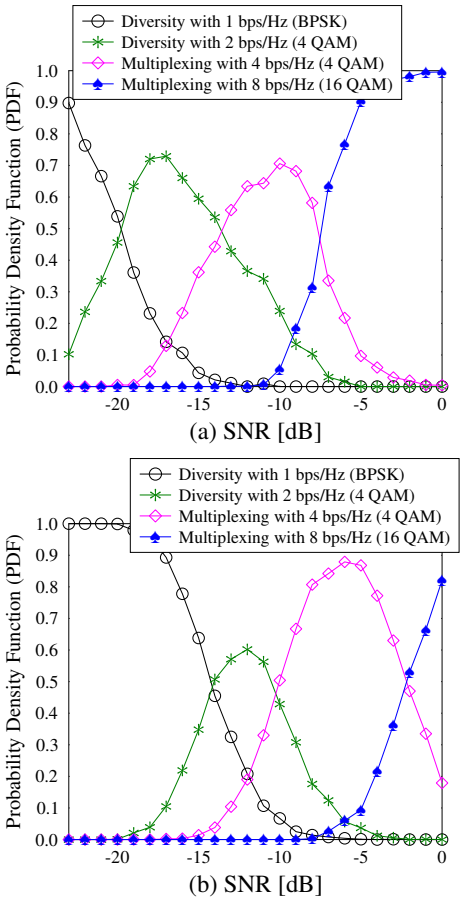


FIGURE 10: *Probability density functions for different classes as a function of the average SNR. Simulation parameters used are listed in Table I. (a) Learning assisted link-adaptation. (b) Conventional link-adaptation. In this setting, the channel is NLOS in nature.*

It is readily seen that the proposed learning assisted link-adaptation achieves a superior throughput and the throughput difference is substantial, especially at the border lines, during the transition between adjacent classes. In Fig. 11(a), for example, it can be clearly seen that in the high-SNR region the proposed design outperforms the conventional adaptation by about 5 dB.

On the other hand, as an example we have also shown in Fig. 11(b) the throughput of the system when the channel has only a single dominant path, where the adaptation takes place between the QPSK and 16QAM schemes. It can be seen from Fig. 11(b) that the learning-aided design achieves superior performance over the conventional design. Again, the digital precoder is switched off in this design.

Fig. 12(a) shows the BER performance of both the proposed design and of the conventional design as a function of the average SNR. It can be observed from Fig. 12(a) that the proposed design meets the target BER of 10^{-3} whilst providing a higher data rate, as shown in Fig. 11 (a). Although the conventional adaptation typically provides a lower BER

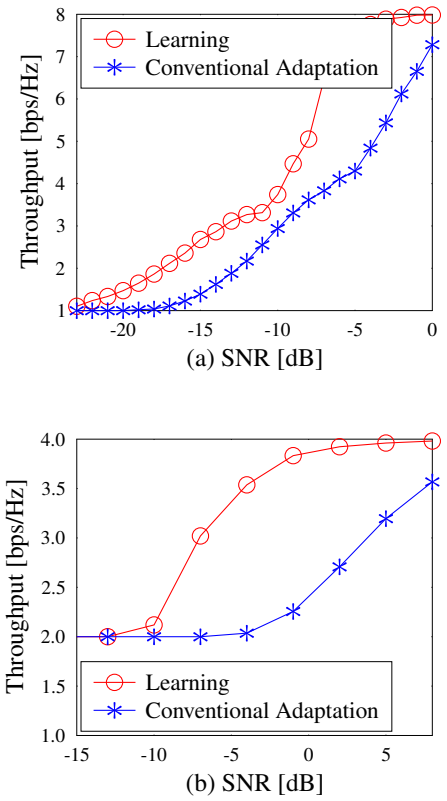


FIGURE 11: *Capacity of the proposed design and of the conventional adaptation as a function of the average SNR — (a) NLOS (b) Only one dominant path. Simulation parameters used are listed in Table 1.*

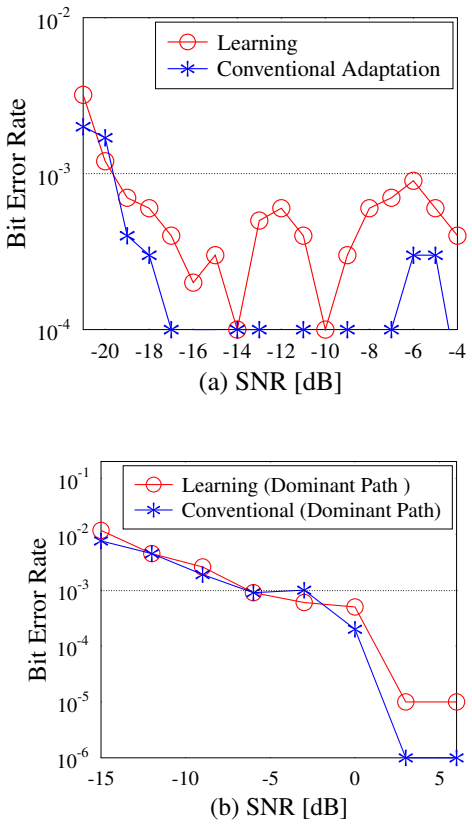


FIGURE 12: *BER performance of the proposed design and of the conventional design as a function of the average SNR —(a) NLOS (b) Only one dominant path. Simulation parameters used are listed in Table 1.*

than the target, it fails to reach the highest data rate possible. Similarly, Fig. 12(b) shows the BER performance when the channel has only a single dominant path, where the learning assisted design meets the target BER of 10^{-3}, whilst also providing a higher data rate, as shown in Fig. 11(b).

Remark 5: The KNN algorithm's superior performance can be attributed to the learning strategy, which is invoked offline. Furthermore, the KNN algorithm also records the post-processing SNR, the instantaneous BER and the rate as a data point during its operation transmission and updates its data points. These data points are further used for decision-making during the next transmission time slot. These attributes make the KNN algorithm aided system superior to the conventional technique.

VI. CONCLUSIONS

We proposed a transmitter design for MU mmWave systems, where the number of phase shifters is dependent on the BF gain required to compensate for the propagation loss of each user. In this design we activate exactly the required number of phase shifters from the available set of phase shifters. We then proposed grouping of the RF chains at the BS to serve each user depending on the user's channel, followed by a machine-learning assisted link-adaptation scheme conceived for mmWave systems, where the receiver predicts the most appropriate type of spatial multiplexing versus diversity aided transmission as well as the most suitable modulation mode to be employed by the BS for every new channel realization. Furthermore, we proposed switching off the digital precoder and combiner when the channel has only a single dominant path, where the communication is established using the dominant path. We demonstrated by simulation that the proposed learning assisted adaptation readily meets the target BER, while providing significantly higher data rate than the conventional link-adaptation relying on SNR threshold values.

REFERENCES

[1] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," IEEE Commun. Surveys Tuts., vol. 20, no. 2, pp. 870–913, Secondquarter 2018.

[2] G. Maccartney, M. Samimi, and T. Rappaport, "Exploiting directionality for millimeter-wave wireless system improvement," in Proc. ICC, June 2015, pp. 2416–2422.

[3] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," IEEE Trans. Signal

Process., vol. 53, no. 11, pp. 4091–4103, Nov 2005.
[4] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Dual-function hybrid beamforming and transmit diversity aided millimeter wave architecture," IEEE Trans. Veh. Technol., vol. 67, no. 3, pp. 2798–2803, March 2018.
[5] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Millimeter wave hybrid beamforming with DFT-MUB aided precoder codebook design," in Proc. VTC (Fall), Sep. 2017, pp. 1–5.
[6] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in Proc. Globecom, Dec 2013, pp. 3679–3684.
[7] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," IEEE Wireless Communications Letters, vol. 3, no. 6, pp. 653–656, Dec 2014.
[8] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," IEEE Trans. Wireless Commun., vol. 15, no. 5, pp. 3311–3326, May 2016.
[9] S. Han, C. I, Z. Xu, and S. Wang, "Reference signals design for hybrid analog and digital beamforming," IEEE Commun. Lett., vol. 18, no. 7, pp. 1191–1193, July 2014.
[10] L. Hanzo, C. H. Wong, and M. S. Yee, Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA, MC-CDMA and OFDM Systems. New York, NY, USA: John Wiley & Sons, Inc., 2002.
[11] L. Chen et al., "Energy-efficient link adaptation on Rayleigh fading channel for OSTBC MIMO system with imperfect CSIT," IEEE Trans. Veh. Technol., vol. 62, no. 4, pp. 1577–1585, May 2013.
[12] E. Eraslan and B. Daneshrad, "Low-complexity link adaptation for energy efficiency maximization in MIMO-OFDM systems," IEEE Trans. Wireless Commun., vol. 16, no. 8, pp. 5102–5114, Aug 2017.
[13] T. L. Jensen et al., "Fast link adaptation for MIMO-OFDM," IEEE Trans. Veh. Technol., vol. 59, no. 8, pp. 3766–3778, Oct 2010.
[14] F. Peng, J. Zhang, and W. E. Ryan, "Adaptive modulation and coding for IEEE 802.11n," in Proc. WCNC, March 2007, pp. 656–661.
[15] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. M. Mourad, and L. Hanzo, "Adaptive transceiver design for C-RAN in mmwave communications," IEEE Access, vol. 6, pp. 16 770–16 782, 2018.
[16] M. El-Hajjar, S. Zummo, and L. Hanzo, "Near-instantaneously adaptive cooperative uplink schemes based on space-time block codes and V-BLAST," in Proc. VTC, April 2007, pp. 2200–2204.
[17] R. C. Daniels, C. M. Caramanis, and R. W. Heath, "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering," IEEE Trans. Veh. Technol., vol. 59, no. 1, pp. 114–126, Jan 2010.
[18] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification (2nd Edition). Wiley-Interscience, 2000.
[19] M. Mohri, A. Rostamizadeh, and A. Talwalkar, Foundations of Machine Learning (1st Edition). The MIT Press, 2012.
[20] R. Daniels and R. W. Heath, "Online adaptive modulation and coding with support vector machines," in Proc. European Wireless Conf., April 2010, pp. 718–724.
[21] Z. Puljiz, M. Park, and R. W. Heath, "A machine learning approach to link adaptation for SC-FDE system," in Proc. Globecom, Dec 2011, pp. 1–5.
[22] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," IEEE Trans. Veh. Technol., vol. 66, no. 1, pp. 763–776, Jan 2017.
[23] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in Proc. Int. Workshop Signal Process. Adv. Wireless Commun., July 2017, pp. 1–5.
[24] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Adaptive transceiver design for C-RAN in mmwave communications," IEEE Access, vol. 6, pp. 16 770–16 782, 2018.
[25] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," IEEE J. Sel. Topics Sig. Proc., vol. 10, no. 3, pp. 528–542, April 2016.
[26] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," IEEE Trans. Veh. Technol., vol. 68, no. 2, pp. 1394–1404, Feb 2019.
[27] S. Ougiaroglou and G. Evangelidis, "RHC: a non-parametric cluster-based data reduction for efficient k-NN classification," Pattern Analysis and Applications, vol. 19, no. 1, pp. 93–109, 2016.
[28] X. Wang, "A fast exact k-nearest neighbors algorithm for high dimensional search using k-means clustering and triangle inequality," in Int. Jr. Conf. Neural Netw., April 2012, pp. 2351—2358.
[29] Rokach and Lior, A Survey of Clustering Algorithms: Data Mining and Knowledge Discovery Handbook (2nd Edition). Springer US, 2010.

K. SATYANARAYANA (www.satyanarayana.xyz) received his B. Tech. degree in Electrical Engineering from Indian Institute of Technology Madras, India, in 2014. During Jul'14-Aug'15, he worked as a research assistant at Indian Institute of Science, Bangalore. Currently, Satya is a research scholar in Wireless Communications at the University of Southampton in liaison with InterDigital Europe, London, UK. His research interests include millimeter wave communications, HBF, with an emphasis on transceiver algorithms for wireless communication systems and multi-functional MIMO.

MOHAMMED EL-HAJJAR is an Associate Professor in the department of Electronics and Computer Science in the University of Southampton. He received his PhD in Wireless Communications from the University of Southampton, UK in 2008. Following the PhD, he joined Imagination Technologies as a design engineer, where he worked on designing and developing ImaginationâĂŹs multi-standard communications platform, which resulted in three patents. He is the recipient of several academic awards and has published a Wiley-IEEE book and in excess of 80 journal and conference papers. Mohammed's research interests include the design of intelligent and energy-efficient transceivers, MIMO, millimeter wave communications, cross-layer optimization for large-scale networks and Radio over fiber network design.

ALAIN A. M. MOURAD holds a PhD degree in Telecommunications from ENST Bretagne in France. He has over 15 yearsâĂŹ experience in the wireless networks industry. He is currently leading the research and development of Next Generation Radio Access Networks at InterDigital International Labs (London, Berlin, Seoul). Prior to joining InterDigital, Dr. Mourad was a Principal Engineer at Samsung Electronics R&D (UK) and previously a Senior Engineer at Mitsubishi Electric R&D Centre Europe (France). Throughout his career, Dr. Mourad has been active in the research and standardization of recent communication networks (5G/4G/3G) and broadcasting systems (ATSC 3.0/DVB-NGH/DVB-T2). He has held various leadership roles in the industry, invented over 35 granted patents and several other patent applications, and authored over 50 peer-reviewed publications. He received the Inventor of the Year Award from Samsung Electronics R&D (UK) twice in 2012 and 2013, and in 2016 InterDigital Innovation Award for the âĂIJidea, creation, and execution of InterDigital EuropeâĂİ.

LAJOS HANZO (http://www-mobile.ecs.soton.ac.uk) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 100+ PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1600+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research CouncilâĂŹs Advanced Fellow Grant and the Royal SocietyâĂŹs Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://wwwmobile.ecs.soton.ac.uk Lajos has 30 000+ citations and an H-index of 68.