Millimeter-Wave Beam Training Acceleration through Low-Complexity Hybrid Transceivers

Danilo De Donno, Joan Palacios, and Joerg Widmer

***Abstract*—Millimeter-wave (mm-wave) communication systems can provide much higher data rates than systems operating at lower frequencies, but achieving such rates over sufficiently large distances requires highly directional beamforming at both the transmitter and receiver. These antenna beams have to be aligned very precisely in order to obtain sufficient link margin. In this paper, we first propose a parallel-adaptive beam training protocol which significantly accelerates the link establishment between mm-wave devices by exploiting the ability of hybrid analog-digital beamforming antennas to scan multiple spatial sectors simultaneously. Second, we deal with practical constraints of mm-wave transceivers and design a novel greedy geometric algorithm to synthesize sector beam patterns featuring configurable beamwidth and multi-beam radiation as required by the proposed beam training protocol. These multi-beam patterns are then also used for concurrent data communication over multiple paths, in case several suitable directions are found during the beam training. Simulation results show that our algorithm is able to shape antenna patterns very close to those attained by a fully-digital beamforming architecture, yet requires lower complexity hardware compared with state-of-the-art solutions. Exploiting such multi-beam antenna patterns, our parallel beam training protocol can provide up to 82% effective rate increase and 70% search time decrease compared to existing sequential protocols. The acceleration of the beam training phase shifts the optimum balance between the search overhead and the achieved directivity gain so that the best performance is reached with a training load 30% to 60% lower than that of sequential beam training.**

***Index Terms*—Millimeter wave systems, hybrid beamforming, beam training, channel estimation, antenna patterns.**

I. Introduction

DUE to the spectrum fragmentation and the limited bandwidth below 6 GHz, current wireless communication systems cannot support user data rates of several gigabits per second and above in a commercially viable manner. Such bandwidth shortage has motivated the exploration of vacant spectrum at higher frequencies, and in particular in the millimeter-wave (mm-wave) frequency bands, which the future fifth-generation (5G) wireless standard is expected to exploit [1], [2]. The higher propagation loss and unfavorable atmospheric absorption make data transmission over relatively long distances a serious challenge at mm-wave frequencies. Fortunately, the short wavelength allows more antenna elements to be integrated into mm-wave devices, thus enabling the implementation of massive multiple-input multiple-output (MIMO) and large-sized phased arrays to overcome range limitations. However, directional communications complicate the link establishment between mm-wave devices, which must perform a time-consuming beam training procedure to determine suitable directions of transmission and reception. This incurs significant overhead and waste of network resources.

Channel estimation as performed in sub-6 GHz MIMO systems is, at present, impractical because the high cost, power consumption, and complexity of mixed-signal hardware at mm-wave frequencies prevent the use of a dedicated RF chain per antenna element. For this reason, most of the literature proposes analog beamforming (BF) architectures which rely on networks of RF phase shifters to control the phase of the signal at each antenna element [3]–[7]. However, compared to fully digital BF, where both phase and amplitude of the transmitted/received signals can be precisely controlled, analog BF is sub-optimal because of the constant amplitude as well as the low phase resolution of RF phase shifters currently available for mm-wave frequencies [8], [9]. Such constraints have a significant impact on the beamforming performance, resulting in antenna radiation patterns with (i) high sidelobe level (SLL), (ii) poor flatness over the covered sector, and (iii) difficult-to-control beamwidth. This latter point is crucial for the design of low-overhead, multi-level beam training protocols where the angular sector covered by the beam is wider in the earlier training stages and is progressively narrowed down to a converging antenna pattern for subsequent data transmissions [3]–[5]. Furthermore, since the transceiver can only use one beam direction at a time, analog BF lacks parallel, multi-stream processing thus resulting in reduced throughput and more time required for beam training. This leads to the well-known tradeoff between the time devoted to beam training overhead and the achieved directivity gain [10]. That is, increasing the alignment overhead leaves less time for data transmission, but, at the same time, leads to higher transmission rates.

Hybrid analog-digital architectures aim at approaching the performance of fully digital BF by dividing the precoding/combining operations between the analog and digital domains while having fewer RF chains compared to pure digital architectures. The availability of multiple RF chains allows hybrid analog-digital architectures to shape beam patterns very close to those attained by a fully digital architecture, i.e., with limited overlap between adjacent beams and excellent flatness over the covered sectors. Moreover, since multiple RF chains provide parallel, multi-stream capabilities, hybrid analog-digital architectures are able to speed up the beam training phase by scanning multiple spatial directions simul-

Manuscript received June xx, 2016.

D. De Donno and J. Widmer are with the IMDEA Networks Institute, Madrid, Spain (e-mail: danilo.dedonno@imdea.org; joerg.widmer@imdea.org).

J. Palacios is with IMDEA Networks Institute, Madrid, Spain, and also with Universidad Carlos III de Madrid, Madrid, Spain (e-mail: joan.palacios@imdea.org).

taneously, hence increasing the effective data rate.

Despite the substantial research interest in hybrid BF for mm-wave systems, only very few works in the literature have investigated the opportunity of leveraging the multi-directional scanning capabilities of hybrid transceivers to accelerate the beam training phase [11]–[13]. Most of the literature on mm-wave hybrid BF has, until now, considered sequential, single-directional scanning during the beam training phase, focusing the attention, instead, on the design of optimal precoders to achieve larger precoding gains during the data trasmission phase. However, as will be discussed in the next section, those solutions exhibit high computational complexity and rely on unrealistic hardware assumptions [14]–[19].

In this paper, taking our prior work in [12], [20], [21] as a starting point, we consider the problem of link establishment between two mm-wave devices and generalize the design of adaptive beam training protocols (i.e., protocols where the antenna sectors are adaptively narrowed based on the bisection concept) by accommodating mm-wave transceivers able to scan multiple sectors simultaneously. We propose an adaptive, low-overhead beam training strategy which is significantly faster and more effective than sequential scanning approaches in the literature. To do that, we take into account practical hardware constraints of mm-wave hybrid transceivers and design a lightweight, yet effective algorithm to synthesize sector beam patterns with configurable beamwidth and multi-beam radiation characteristics as required by the proposed protocol. The main contributions of the paper are as follows:

- We propose to speed-up the beam training process in mm-wave systems by exploiting the simultaneous, multi-directional transmission and reception capabilities of hybrid analog-digital transceivers. Based on this concept, we design an adaptive beam training protocol that provides a combinatorial reduction in the number of beam search steps compared to existing approaches based on sequential, single-directional scanning.
- We develop, to the authors knowledge for the first time in the mm-wave context, the mathematical tools for an hybrid receiver to reliably separate multiple interacting streams impinging on its antennas and select the sector pair (i.e., the transmit sector at the transmitting device and the receive sector at the receiving device) which provides the maximum beamformed gain.
- We formulate an optimization problem to approximate fully-digital, multi-beam antenna patterns by means of a hybrid analog-digital transceiver architecture with much fewer RF chains than antenna elements. Compared with existing approaches in the literature which require high-resolution RF phase shifters, our approach is able to achieve near optimal performance with only 2-bit phase quantization.
- We propose and implement a lightweight algorithm, based on geometric intuitions, to efficiently solve the above optimization problem with much lower computational complexity and higher accuracy than state-of-the-art strategies. Our design enables simultaneous, multi-directional scanning and, therefore, can be leveraged in the framework of beam training to accelerate the mm-wave link establishment. Once the beam training procedure is accomplished, our algorithm allows also to synthesize the hybrid precoders/combiners to be used for the communication itself through multiple, parallel data streams transmitted/received simultaneously over different paths.

Simulation results show that our solution is able to synthesize beam patterns almost indistinguishable from those shaped by fully-digital BF, while requiring lower complexity hardware compared with the literature. The speed-up achieved by the proposed beam training protocol with simultaneous, multi-directional transmission/reception translates into a considerable rate increase compared to state-of-the-art strategies either adopting sequential, single-directional scanning or parallel, multi-directional scanning at the receiver only. Furthermore, while the tradeoff between beam training overhead and achievable rate has been already investigated for a simple line-of-sight (LOS) channel [10], we consider, on the contrary, a realistic channel model and propose low-complexity hybrid beamforming techniques that significantly reduce the latency associated with the point where the optimum tradeoff between training overhead and data rate is reached.

The paper is organized as follows. The system model is presented in §III. In §IV, we generalize the design of adaptive beam training protocols by distinguishing between two main strategies, namely sequential adaptive and parallel adaptive beam training. In §V, we design a lightweight, greedy-geometric algorithm to approximate fully-digital, multi-beam antenna patterns by means of a low-complexity hybrid architecture. Simulation results assessing the performance of the proposed algorithms are given in §VI, before drawing the main conclusions of the paper in §VII.

We use the following notation along the paper. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex normal distribution with mean μ and variance σ^2, $\mathbf{A}$ is a matrix, $\mathbf{a}$ is a vector, and $\mathcal{A}$ denotes a set. $\|\mathbf{a}\|_2$ is the Euclidean norm of $\mathbf{a}$, $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$, $|\mathbf{A}|$ is its determinant, while $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^{-1}$ denote its transpose, Hermitian, and inverse respectively. $[\mathbf{A}]_{\mathcal{B},:}$ ($[\mathbf{A}]_{:,\mathcal{B}}$) are the rows (columns) of the matrix $\mathbf{A}$ indexed by the set $\mathcal{B}$, and $\mathbf{I}$ denotes the identity matrix. $\mathbb{E}[\cdot]$ is the expectation operator and $\lceil \cdot \rceil$ denotes the ceiling function.

II. Related work

As outlined in the introduction, the implementation of fully digital BF with a large number of antenna elements is currently infeasible at mm-wave frequencies because of cost and implementation issues. At the same time, despite being commonly adopted in mm-wave devices[1], analog BF represents a sub-optimal solution due to its inability to process multiple parallel streams and to synthesize beam patterns highly configurable in terms of required beamwidth, steering direction, flatness, and SLL. Different strategies are proposed in the literature

[1]For example, the Talon AD7200 Multi-Band Wi-Fi Router is equipped with the Intel Wireless Gigabit W13100 802.11ad radio based on analog beamforming (http://www.intel.com/content/www/us/en/wireless-products/wigig-sink-w13100.html).

to provide analog BF architectures with variable beamwidth capabilities. For example, quadratic phase excitation beamforming is proposed in [22] to synthesize wide beam array patterns exhibiting, however, poor performance in terms of gain flatness over the covered sector. Commercially available mm-wave devices based on analog BF, instead, adopt antenna subset selection strategies to synthesize variable beamwidth radiation patterns [23]. Selection, however, provides limited equivalent isotropic radiated power (EIRP), especially when one power amplifier per antenna element is used [24]. In this case, in fact, assuming a M-element antenna array with only m_{on} active elements, the EIRP is reduced by a factor of $(m_{\text{on}}/M)^2$ — one m_{on}/M factor accounts for the array gain reduction while the other m_{on}/M factor accounts for the reduced number of active amplifiers. Several beam training protocols using analog BF and that are based on the idea of adaptively narrowing down the antenna pattern beamwidth are proposed in the literature [3]–[5]. Further improvements of this idea are discussed in [6], where a numerical Rosenbrock search is considered to reduce the number of probed sector combinations. However, in addition to the aforementioned limitations in terms of beam shape quality, all these solutions based on analog BF [3]–[6] share the disadvantage of converging towards only one communication beam, thus making de facto impossible the exploitation of spatial multiplexing gain.

Hybrid analog-digital BF represents a good compromise between cost/complexity and performance. In fact, the availability of multiple RF chains allows hybrid analog-digital architectures to overcome the limitations of analog BF in terms of beam pattern synthesis and multi-stream processing. Among the several applications, this latter ability can be exploited, for example, to accelerate the link establishment between mm-wave devices. Most of the literature on mm-wave hybrid BF proposes high complexity solutions built upon unrealistic hardware assumptions based, for example, on the availability of a considerable number of RF chains and RF phase shifters with large (or even infinite) number of quantization bits. Moreover, the vast majority of work ignores the potential of parallel, multi-stream processing to speed-up the beam training phase. Instead, the design of hybrid precoders and combiners to maximize the rate during the data transmission phase is the most investigated topic.

In [14], a codebook to be used for sequential, single-directional scanning is developed under the assumption that hybrid analog-digital BF is available only at the transmitter while the receiver is equipped with a single antenna. The hybrid codebook is designed by minimizing the mean square error (MSE) between the code vector's beam pattern and its corresponding ideal beam pattern. The MSE minimization is accomplished by the orthogonal matching pursuit (OMP) algorithm with unconstrained digitally-controlled RF phase shifters (i.e., with infinite number of quantization bits). A similar OMP-based approach with unconstrained RF phase shifters is discussed also in [15]. In [16], the OMP algorithmic concept is used to design hybrid precoders/combiners assuming perfect channel knowledge, but without investigating how to shape optimal beam patterns and leverage them for beam training. The design of hybrid analog-digital codebooks with variable beamwidths was pioneered by Heath et al. in [17]. The ability to generate antenna patterns with various beamwidths makes such a codebook particularly attractive for the design of adaptive, low-overhead beam training protocols. The approach in [17] focuses on sequential, single-directional scanning during beam training and assumes that RF phase shifters with a large number of quantization bits are available at mm-wave frequencies. However, the current state of silicon technologies makes the design of mm-wave RF phase shifters with high phase shift resolution impractical [9]. Unlike [17] which only uses RF phase shifters, [18] proposes an even more complex hybrid BF architecture in which mm-wave devices are equipped with both antenna switches and RF phase shifters with 6-bit phase quantization. An antenna selection strategy is exploited in [18] to optimize the beam pattern design over all quantized phase values and different antenna subsets. However, antenna subset selection for hybrid architectures shares the aforementioned EIRP reduction disadvantage encountered in analog BF architectures. In [19], RF phase shifters with fewer quantization bits are exploited for the design of hybrid analog-digital codebooks consisting, however, only of narrow beams with fixed beamwidth.

To the best of the authors' knowledge, the only works that investigate multi-directional scanning via hybrid BF in the context of mm-wave initial access are [11], [13]. In [11], the cell discovery problem is considered and detectors of the synchronization signals transmitted by the mm-wave base station are derived only for the cases of analog and digital BF. The gain provided by the ability of mobile stations to scan multiple directions simultaneously via hybrid BF is estimated (i.e., without any actual implementation or derivation) based on the performance achieved by analog and digital BF. In [13], a channel estimation scheme with reduced signaling is proposed using a genetic algorithm to synthesize the hybrid analog-digital precoders and combiners. The multi-stream capabilities of mm-wave hybrid transceivers are exploited only at the receiver to detect the training sequences simultaneously from multiple directions, while the transmitter does not employ multi-directional transmission as with the scheme we propose in this paper. We compare against this algorithm in § VI-C.

III. System model

We consider the mm-wave system shown in Fig. 1 and focus on the beam training phase during which a base station (BS) equipped with a uniform linear array (ULA) of M_{BS} isotropic radiators and N_{BS} RF transceiver chains uses $D_{\text{BS}} \leq N_{\text{BS}}$ streams of data to simultaneously transmit training packets in D_{BS} different directions at each time slot. On the other side, a mobile station (MS) equipped with a ULA of M_{MS} isotropic radiators and N_{MS} RF transceiver chains performs received power measurements simultaneously over D_{MS} directions.

More concretely, we assume that the BS adopts Golay training sequences which, for each sector, are encoded using orthogonal Walsh spreading codewords. Since the training sequences are simultaneously transmitted through different sectors, the use of orthogonal codes helps to reduce the inter-sector interference. At the same time, Golay sequences possess

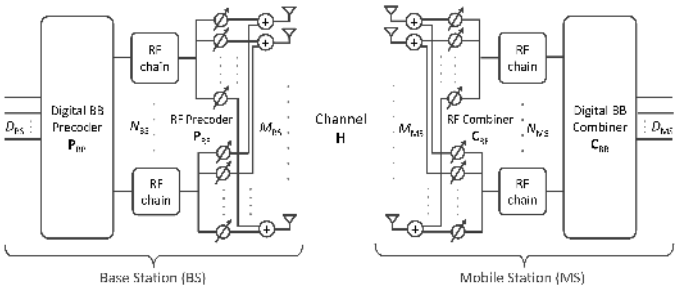


Fig. 1. Overview of BS-MS mm-wave transceiver architecture for hybrid analog-digital BF.

very good auto-correlation, which protects the Walsh codes from reduced orthogonality due to multipath [7]. Ultimately, for each transmit sector $i = 1, 2, ..., D_{\text{BS}}$, the BS emits a training signal $s_i[t]$, for $t = 1, 2, ..., N_s$. The overall set of transmitted symbols can be arranged into a matrix $\mathbf{S}$, with dimensions $D_{\text{BS}} \times N_s$, where the i-th row contains the time-domain sequence transmitted over the i-th direction.

The BS applies an $N_{\text{BS}} \times D_{\text{BS}}$ digital baseband precoder $\mathbf{P}_{\text{BB}}$ followed by an $M_{\text{BS}} \times N_{\text{BS}}$ RF precoder, $\mathbf{P}_{\text{RF}}$, to the training symbol matrix $\mathbf{S}$. To adhere to the transmit power constraint, we impose the normalization $\|[\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}]_{:,i}\|_2^2 = 1$, for $i = 1, 2, ..., D_{\text{BS}}$. If $\mathbf{P} = \mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}$ is the $M_{\text{BS}} \times D_{\text{BS}}$ combined BS precoding matrix and the radiated power is equally distributed among the D_{BS} directions, the discrete-time transmitted signal is then

$$\mathbf{X} = \rho\mathbf{P}\mathbf{S}\,, \tag{1}$$

where ρ is a normalization constant to enforce the total radiated power constraint. In case of orthogonal encoded training sequences, $\rho = \sqrt{P_t/D_{\text{BS}}}$, where P_t is the average transmit power.

We adopt a block-fading channel model in which the MS observes the received signal

$$\mathbf{R} = \mathbf{H}\mathbf{X} + \mathbf{N}\,, \tag{2}$$

where $\mathbf{H}$ is the $M_{\text{MS}} \times M_{\text{BS}}$ matrix representing the mm-wave channel between the BS and MS, and $\mathbf{N}$ is a $M_{\text{MS}} \times N_s$ noise matrix with independent complex entries distributed as Gaussian random variables with mean zero and variance σ^2.

The MS configures its front-end to concurrently receive over $D_{\text{MS}} \leq N_{\text{MS}}$ different directions. To do that, it applies an $M_{\text{MS}} \times N_{\text{MS}}$ RF combiner $\mathbf{C}_{\text{RF}}$ followed by a $N_{\text{MS}} \times D_{\text{MS}}$ digital baseband combiner $\mathbf{C}_{\text{BB}}$. The discrete-time signal received by the MS is the $D_{\text{MS}} \times N_s$ matrix $\mathbf{Y}$ given by

$$\mathbf{Y} = \mathbf{C}^H\mathbf{R} = \sqrt{\frac{P_t}{D_{\text{BS}}}}\mathbf{C}^H\mathbf{H}\mathbf{P}\mathbf{S} + \mathbf{C}^H\mathbf{N}\,, \tag{3}$$

where $\mathbf{C} = \mathbf{C}_{\text{RF}}\mathbf{C}_{\text{BB}}$ is the $M_{\text{MS}} \times D_{\text{MS}}$ MS combiner matrix.

We adopt the 3GPP-style mm-wave channel model presented in [25], which is based on real-world measurements at 28 GHz in New York City. According to the measurements, the mm-wave channel can be modeled as a combination of clusters, each composed of several sub-paths. Specifically, in case of a static scenario without mobility, the channel matrix is described by:

$$\mathbf{H} = \sqrt{\frac{M_{\text{BS}}M_{\text{MS}}}{L}}\sum_{k=1}^{K}\sum_{\ell=1}^{L}\alpha_{k\ell}\mathbf{u}_{\text{MS}}(\psi_{k\ell}^{\text{MS}})\mathbf{u}_{\text{BS}}^H(\psi_{k\ell}^{\text{BS}})\,, \tag{4}$$

where K is the number of clusters, L is the number of sub-paths per cluster, $\alpha_{k\ell}$ is the complex gain on the ℓ-th sub-path of the k-th cluster, and $\mathbf{u}_{\text{MS(BS)}}(\cdot)$ is the antenna array response vector at the MS (BS). While the study presented in this paper can be applied to arbitrary antenna arrays, we consider ULAs whose $\mathbf{u}_{\text{MS}}(\cdot)$ can be written as:

$$\mathbf{u}_{\text{MS}}(\psi_{k\ell}^{\text{MS}}) = \sqrt{1/M_{\text{MS}}}\left(1,\, \mathrm{e}^{j\Psi_{k\ell}},\, ...,\, \mathrm{e}^{j(M_{\text{MS}}-1)\Psi_{k\ell}}\right)^T\,, \tag{5}$$

where $\Psi_{k\ell} = (2\pi/\lambda)\Delta\sin(\psi_{k\ell}^{\text{MS}})$, λ is the wavelength, and Δ is the distance between antenna elements. The ULA response at the BS, $\mathbf{u}_{\text{BS}}(\cdot)$, can be written in a similar way. Note that, in order to simplify the notation, we consider BS and MS implementing horizontal (2-D) beamforming only, which implies that all scattering happens in the azimuthal domain. Extension to planar antenna arrays and, therefore, to 3-D beamforming is straightforward. The channel gain coefficients are assumed to be Rayleigh distributed, i.e., $\alpha_{k\ell} \sim \mathcal{CN}(0, \gamma_k 10^{-0.1PL})$, $k = 1, 2, ..., K$ with γ_k the fraction of power in the k-th cluster and PL the omnidirectional path loss which can be calculated according to the equations in [25, Table I]. As in [25], we assume that, depending on the separation distance d, the BS-MS link can be in one of three conditions: LOS, non-LOS (NLOS) or outage. Statistically, the probability functions for these three states can be modeled as:

$$p_{\text{out}}(d) = \max\left(0, 1 - e^{-a_{\text{out}}d + b_{\text{out}}}\right) \tag{6a}$$

$$p_{\text{LOS}}(d) = (1 - p_{\text{out}}(d))e^{-a_{\text{los}}d} \tag{6b}$$

$$p_{\text{NLOS}}(d) = 1 - p_{\text{out}}(d) - p_{\text{LOS}}(d)\,, \tag{6c}$$

where the parameters a_{los}=0.0149 m^{-1}, a_{out}=0.0334 m^{-1}, and b_{out}=5.2 are fitted from measurement data. Note that, depending on the BS-MS link condition, the model adopts different parameters for PL calculation (see [25, Table I]).

In the next section, we illustrate how the simultaneous multi-directional scanning can be leveraged to design low-overhead beam training protocols for fast and adaptive channel estimation.

IV. Adaptive beam training

In this section, taking [17] and our prior work in [12], [20], [21] as a starting point, we generalize the design of adaptive beam training protocols (i.e., protocols where the antenna sectors are adaptively narrowed based on the bisection concept) by accommodating BSs and MSs able to simultaneously scan multiple directions. Based on the system model in the previous section, we can distinguish two main types of adaptive beam training: (1) *sequential adaptive beam training* (SA-BT), where D_{BS}=D_{MS}=1, and (2) *parallel adaptive beam training* (PA-BT), where $D_{\text{BS}} > 1$ or $D_{\text{MS}} > 1$. For both types, we assume the availability of a feedback channel through which the MS feeds the search results back to the BS. However, such requirement can be easily relaxed by using a ping-pong approach similar to the one described in [26].

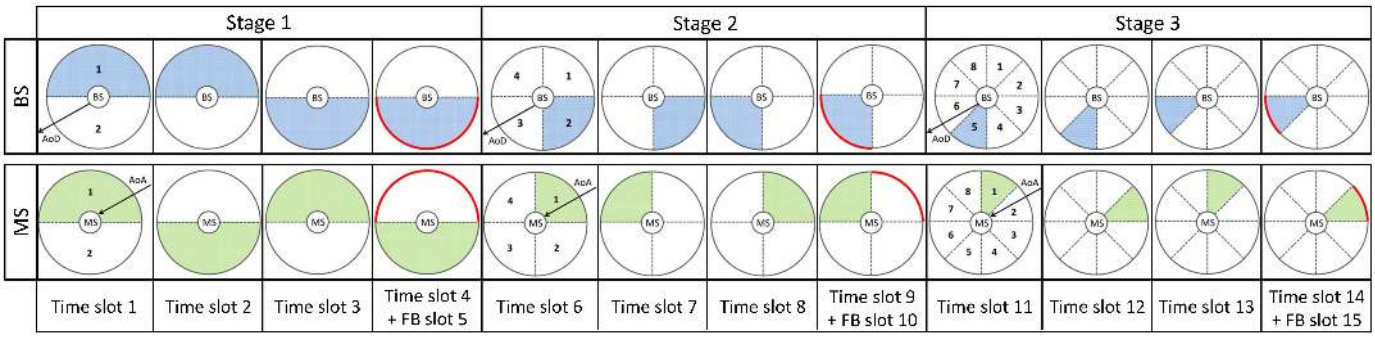


Fig. 2. An example of the adaptive azimuthal domain partitioning performed by BS (blue sectors) and MS (green sectors) during the SA-BT protocol with K_{BS}=K_{MS}=2. The sectors with red border are the ones selected at the end of each stage. Note that here BS and MS can respectively transmit and receive training sequences through only one antenna sector at each time slot.

A. Sequential adaptive beam training (SA-BT)

In SA-BT, BS and MS can respectively transmit and receive training sequences through only one antenna sector at each time slot. Concretely, in the first stage, the BS (MS) divides its $[0, 2\pi]$ azimuthal domain into K_{BS} (K_{MS}) sectors and designs the K_{BS} (K_{MS}) hybrid analog-digital precoding (combininig) vectors giving rise to transmit (receive) antenna patterns covering such sectors. Assuming that a different transmit/receive sector pair is used at each time slot, a total of $K_{\text{BS}} \cdot K_{\text{MS}}$ time slots are necessary to train all the possible transmit/receive sector combinations. At the end of the stage, the MS compares the received training sequences to determine the BS/MS sector pair providing maximum received power. This translates into selecting the sector pair which is highly likely to contain the most dominant Angle of Departure (AoD) and Angle of Arrival (AoA) of the mm-wave channel. The MS communicates the search results to the BS using an additional feedback slot. Then, BS and MS prepare for the later stages where the selected sectors are further divided into smaller subsets until the AoD/AoA pair is estimated with the desired resolution. Implementations of the SA-BT protocol are presented in [17] and in our prior work [20].

An example of partitioning of the azimuthal domain performed by the BS and MS during the first three stages of the SA-BT protocol with K_{BS}=K_{MS}=2 is shown in Fig. 2. In the first stage, the BS transmits sequentially through Sector #1 and Sector #2. At the same time, the MS receives sequentially from its own sectors and performs signal power measurements. Sector #1 and Sector #2 are selected respectively as the BS transmit sector and MS receive sector providing maximum beamformed channel gain. The results of the estimation performed at the MS are communicated to the BS via an uplink feedback slot. In the second stage, the sectors selected in the previous stage are further divided into two 90° sectors which are trained sequentially as in the first stage. The process is repeated until the required angular resolution of the AoD/AoA estimation is achieved. The SA-BT protocol can be easily extended in order to estimate more than one dominant path in the channel. This can be done by cycling the procedure just described over the number of dominant paths L_{est} to be estimated and removing, at each iteration, the contribution of the already estimated paths.

B. Parallel adaptive beam training (PA-BT)

Starting from the aforementioned SA-BT algorithm, we design in this subsection a PA-BT protocol which takes advantage of the ability of hybrid analog-digital transceivers to scan multiple directions simultaneously as modeled in §III.

In the first stage of the protocol, the BS divides its $[0, 2\pi]$ azimuthal domain into K_{BS} sectors and designs the hybrid analog-digital precoder $\mathbf{P}$ to transmit the training sequences over $D_{\text{BS}} \leq \min(K_{\text{BS}}, N_{\text{BS}})$ sectors simultaneously as described in the previous section. The number of steps required to perform a full 2π scan is therefore $\lceil K_{\text{BS}}/D_{\text{BS}} \rceil$. At the same time, the MS divides its azimuthal domain into K_{MS} sectors and applies the hybrid combiner $\mathbf{C}$ at $\lceil K_{\text{MS}}/D_{\text{MS}} \rceil$ successive time slots. In each time slot, the BS signals are received by the MS in parallel (i.e., simultaneously) from $D_{\text{MS}} \leq \min(K_{\text{MS}}, N_{\text{MS}})$ spatial directions. As shown in Fig. 3 for the case with K_{BS}=K_{MS}=D_{BS}=D_{MS}=2, only one time slot is required in Stage 1 since the whole azimuthal domain is covered all at once using simultaneous transmission at the BS and simultaneous reception at the MS.

The objective for the MS is to select the sector pair (i.e., the combination of transmit BS sector and receive MS sector) that maximizes the received signal power. This translates into the estimation of the $D_{\text{MS}} \times D_{\text{BS}}$ joint channel matrix $\overline{\mathbf{Y}} = \sqrt{P_t}\mathbf{C}^H\mathbf{H}\mathbf{P}$ from the measured signal $\mathbf{Y}$ in Eq. 3. In fact, each element $[\overline{\mathbf{Y}}]_{i,j}$ contains the beamformed channel gain achieved when the MS is receiving over $[\mathbf{C}]_{:,i}$ and the BS is transmitting over $[\mathbf{P}]_{:,j}$. The power gain over such precoder/combiner configuration is $\left|[\overline{\mathbf{Y}}]_{i,j}\right|^2$.

Given the orthogonality among the rows of $\mathbf{S}$, the matrix $\mathbf{F} = \frac{1}{N_s}\mathbf{S}^H$ such that $\mathbf{SF} = \mathbf{I}$ can be defined. Then, since the independent and identically distributed (IID) noise components remain independent under orthonormal transformation, we can calculate

$$\mathbf{Y}' = \sqrt{D_{\text{BS}}}\mathbf{Y}\mathbf{F} = \sqrt{P_t}\mathbf{C}^H\mathbf{H}\mathbf{P} + \sqrt{D_{\text{BS}}}\mathbf{C}^H\mathbf{N}', \qquad (7)$$

where $\mathbf{N}'$ is a $M_{\text{MS}} \times D_{\text{BS}}$ white noise matrix with independent complex elements distributed as Gaussian random variables with mean zero and variance σ^2/N_s. As is evident from Eq. 7,

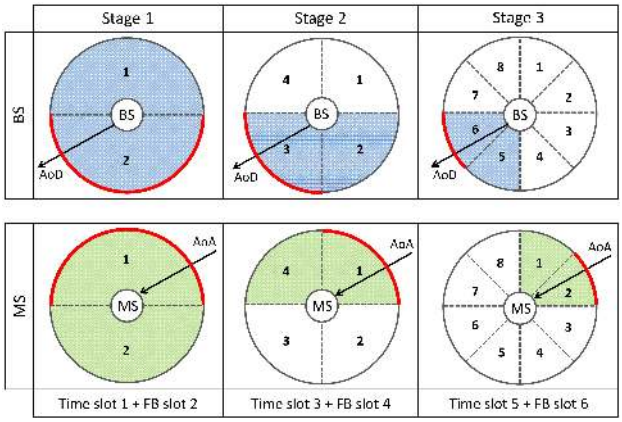


Fig. 3. An example of the adaptive azimuthal domain partitioning performed by BS (blue sectors) and MS (green sectors) during the PA-BT protocol with K_{BS}=K_{MS}=D_{BS}=D_{MS}=2. At each stage, the sectors filled with the same color are covered simultaneously by the respective device while the sectors with red border are the ones estimated to provide maximum beamformed channel gain.

since $\mathbf{Y}'$ is Gaussian distributed with mean $\overline{\mathbf{Y}}$, the expected value of $\overline{\mathbf{Y}}$ can be computed as $\hat{\overline{\mathbf{Y}}} = \sqrt{D_{\text{BS}}}\mathbf{YF}$. Note that D_{BS}, N_s, and the training sequences adopted by the BS can be considered as cell configuration parameters which have been acquired by the MS during the cell discovery phase.

The MS calculates the received signal power for each BS-MS sector pair as $|[\hat{\overline{\mathbf{Y}}}]_{i,j}|^2$, with $i = 1, 2, ..., D_{\text{MS}}$ and $j = 1, 2, ..., D_{\text{BS}}$, to determine the one with the maximum beamformed channel gain. This translates into selecting the sector pair which is highly likely to contain the most dominant AoD and AoA of the mm-wave channel. The MS uses an additional feedback slot to communicate the search results to the BS. In the subsequent stages, the procedure is repeated with the selected sectors further divided into smaller subsets until the final AoD/AoA pair is estimated with the desired resolution. As is evident from Fig. 3, the ability of hybrid transceivers to scan multiple sectors simultaneously provides a combinatorial reduction in the number of time slots required by PA-BT against SA-BT. It should be noted that increasing D_{BS} and D_{MS} likewise reduces the number of RF chains per scanned direction, which affects the quality of the beam patterns and, consequently, the channel estimation. Such a qualitative tradeoff will be quantitatively evaluated in §VI-A.

Also the proposed PA-BT protocol can be easily extended to estimate more than one dominant AoD/AoA pair. As in the SA-BT case, this can be done by repeating the training procedure over the number of paths L_{est} to be estimated and removing, at each iteration, the contribution of the already estimated paths.

C. Beam training overhead

Both the SA-BT and PA-BT strategies can be implemented using the frame structure proposed in Fig. 4. In case of reciprocal channel, at each stage, the BS transmits training sequences to the MS through one (SA-BT) or more sectors simultaneously (PA-BT) in each time slot. At the same time, the MS performs power measurements according to the adopted SA-BT or PA-BT strategy and feeds the results back to the BS. For simplicity, this paper will always assume a reciprocal channel between BS and MS — channel non-reciprocity can be easily handled by repeating the beam training algorithm with $\mathbf{H}$ replaced by the uplink channel and the roles of the BS and MS reversed.

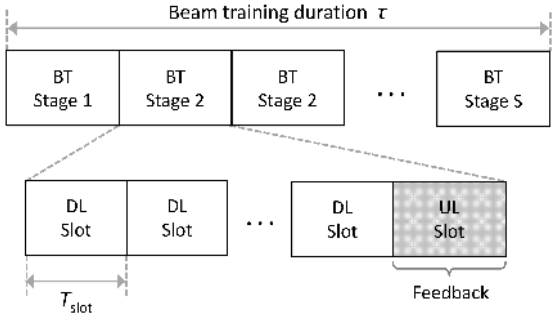


Fig. 4. Frame structure used during the beam training phase. At each stage, the BS transmits training sequences through different sectors depending on the adopted SA-BT or PA-BT strategy (DL slots) while the MS performs power measurements and feeds the result back to the BS (UL slot).

As shown in [17], the number of beam training stages required to estimate the L_{est} most dominant AoD/AoA pairs of the channel with angular resolution $2\pi/N$ is

$$S = \max\left[\log_{K_{\text{BS}}} N, \log_{K_{\text{MS}}} N\right], \tag{8}$$

With time-slot duration T_{slot}, the SA-BT strategy requires that K_{BS} precoders are sequentially used at the BS and, for each of them, K_{MS} combiners are sequentially employed at the MS. The total beam training time, including feedback slots, is therefore:

$$\tau_{\text{SA-BT}} = L_{\text{est}} S (K_{\text{BS}} K_{\text{MS}} + 1) T_{\text{slot}}, \tag{9}$$

In the PA-BT case, instead, since $\lceil K_{\text{BS}}/D_{\text{BS}} \rceil$ precoding vectors are used at the BS at each stage and, for each of them, $\lceil K_{\text{MS}}/D_{\text{MS}} \rceil$ combining vectors are employed at the MS, the total beam training time, including feedback slots, becomes:

$$\tau_{\text{PA-BT}} = L_{\text{est}} S \left(\left\lceil \frac{K_{\text{BS}}}{D_{\text{BS}}} \right\rceil \left\lceil \frac{K_{\text{MS}}}{D_{\text{MS}}} \right\rceil + 1 \right) T_{\text{slot}}, \tag{10}$$

As is evident from Eq. 9 and Eq. 10, the PA-BT strategy provides a significant speed up compared to the SA-BT strategy. Numerical experiments in §VI will demonstrate how a reduced time devoted to beam training leads to an increased data transmission rate.

The AoD/AoA estimation accuracy achieved at the end of beam training critically depends on how well the adopted beam patterns approximate the ideal sectors in Fig. 2 and Fig. 3. In the next section, we present an algorithmic strategy to accurately approach fully-digital sector beams by means of a hybrid analog-digital architecture requiring very low complexity hardware.

V. Hybrid analog-digital beam patterns

The problem of designing the best hybrid analog-digital precoder/combiner at the BS and MS can be translated into the one of designing, at each stage of the beam training protocol, beam patterns able to optimally cover the sectors involved

in the transmission/reception of training sequences. This can be translated into the synthesis of beam patterns with (1) limited overlap with adjacent beams, (2) flat-top shape over the covered angular region, and (3) reduced SLL.

In this section, we present a practical codebook design which relies on the availability of a hybrid analog-digital architecture with number of RF chains much lower than the number of antenna elements, and RF phase shifters with only 2-bit resolution, i.e., they can be set with just four phase values (0, $\pm\pi/2$, π), without amplitude adjustment.[2]

We refer to the mm-wave system in Fig. 1 where the BS transmits D_{BS} parallel streams of training sequences, i.e., it employs beam patterns able to cover D_{BS} sectors simultaneously as depicted in Fig. 3. On the other side, the MS also adopts parallel reception using beam patterns able to cover D_{MS} sectors simultaneously. For the sake of simplicity, we focus on the codebook design at the MS, but the same formulation holds also for the BS.

A. Problem formulation

At each stage of the adaptive beam training protocol illustrated in §IV, the set of concurrent beam patterns — in the sequel also referred to as measurement vectors or combining vectors — used by the MS to receive beam training data can be arranged into a $M_{\text{MS}} \times D_{\text{MS}}$ matrix $\mathbf{C} = \mathbf{C}_{\text{RF}}\mathbf{C}_{\text{BB}}$, where $D_{\text{MS}} \leq N_{\text{MS}}$ is the number of sectors which the MS is able to cover simultaneously, $\mathbf{C}_{\text{RF}}$ is the RF combining matrix, and $\mathbf{C}_{\text{BB}}$ is the digital baseband combining matrix. The problem of designing a hybrid analog-digital codebook with multiple simultaneous receive beams can be modeled as finding the optimal RF combiner $\mathbf{C}_{\text{RF}}^{\text{opt}}$ and the optimal baseband combiner $\mathbf{C}_{\text{BB}}^{\text{opt}}$ such that:

$$\begin{aligned} \mathbf{C}_{\text{RF}}^{\text{opt}} &= \underset{\mathbf{C}_{\text{RF}}}{\arg\min} \left\| \mathbf{C}_{\text{bsl}} - \mathbf{C}_{\text{RF}}\hat{\mathbf{C}}_{\text{BB}} \right\|_F \\ &\text{s.t. } [\mathbf{C}_{\text{RF}}]_{:,i} \in \mathcal{C}_{\text{RF}},\ i = 1, 2, ..., N_{\text{MS}} \\ \hat{\mathbf{C}}_{\text{BB}} &= \mathbf{C}_{\text{RF}}^{+}\mathbf{C}_{\text{bsl}} \\ \left[\mathbf{C}_{\text{BB}}^{\text{opt}}\right]_{:,i} &= \frac{[\hat{\mathbf{C}}_{\text{BB}}]_{:,i}}{\|\mathbf{C}_{\text{RF}}[\hat{\mathbf{C}}_{\text{BB}}]_{:,i}\|_2},\ i = 1, 2, ..., N_{\text{MS}}\,, \end{aligned} \tag{11}$$

where $\mathbf{C}_{\text{RF}}^{+}$ is a left pseudo-inverse of $\mathbf{C}_{\text{RF}}$ and $\mathbf{C}_{\text{bsl}}$ is the matrix that contains, in its columns, the baseline array vector weights that are to be approximated. In this paper, we refer to baseline beam patterns as the almost ideal sector beam patterns that can be synthesized by exploiting a fully digital BF architecture where the availability of a dedicated RF chain for each antenna element enables precise control of both phase and amplitude of the mm-wave signals. As in [20], [27], we choose the sector beam design based on the Fourier Series Method with Kaiser Window (FSM-KW) to synthesize beam patterns confined to a desired angular region [28, Chapter 21, pp. 946-949]. Compared to other windows such as Hamming, Blackman, etc., the Kaiser window has more design flexibility since the trade-off between the main lobe width and the sidelobe ripple amplitude can be accurately set by adjusting some window parameters.

The finite set of possible analog/RF beamforming vectors is encompassed by the set $\mathcal{C}_{\text{RF}} \in \mathbb{C}^{M_{\text{MS}}}$ of vectors with components $\{\pm 1, \pm j\}$, which is also referred to as dictionary in the sequel. Such constraint on $\mathbf{C}_{\text{RF}}$ translates into the hardware constraint of using RF phase shifters with only two quantization bits.

B. The greedy geometric approach

The problem stated in Eq. 11 can be handled using the OMP algorithm and its variants as proposed in most of the literature on hybrid analog-digital BF [14]–[17], [20]. Despite their simplicity and ease of implementation, OMP-based approaches share the disadvantage of high computational complexity (due to a matrix inversion at each step) and reduced accuracy since they rely on non-complete dictionaries made of basis subsets. In contrast, we propose a novel, dictionary-free approach based on geometric considerations and manipulations, to solve the problem in Eq. 11 with lower computational complexity and higher accuracy than OMP-based strategies. A comparative analysis on the computational complexity of our algorithm versus the conventional OMP algorithm is carried out in §VI-D via numerical experiments. It is worth noting that, *in contrast to the above mentioned literature which focuses on the design of hybrid codebooks covering only one sector at a time, our design enables BS and MS to simultaneously cover more than one sector during the beam training process. This translates into the problem of finding the optimal RF precoder* $\mathbf{P}_{\text{RF}}$ *(at the BS) and RF combiner* $\mathbf{C}_{\text{RF}}$ *(at the MS) for all the directions scanned simultaneously.*

Algorithm 1 outlines the pseudo-code of our greedy, geometric strategy to approximate the fully-digital codebook $\mathbf{C}_{\text{bsl}}$ by a hybrid analog-digital architecture with 2-bit RF phase shifters[3]. The algorithm takes as input parameter the baseline combiner to approximate, $\mathbf{C}_{\text{bsl}}$. Then, $\mathbf{C}_{\text{RF}}$ and the residual matrix $\mathbf{C}_{\text{res}}$ are initialized. The algorithm proceeds by selecting from $\mathbf{C}_{\text{res}}$ the residual column $\mathbf{c}_{\text{res}}$ with maximum norm and appending its transformed version $\mathcal{S}(\mathbf{c}_{\text{res}})$ to the current RF combiner $\mathbf{C}_{\text{RF}}$ (Lines 4 to 6). It is worth highlighting here the dictionary-free nature of our approach. While at this stage OMP-based strategies select the column vectors of the RF combiner, $\mathbf{C}_{\text{RF}}$, from the dictionary (usually, a small-sized dictionary is used to reduce the computational complexity), our algorithm computes the current $\mathbf{C}_{\text{RF}}$ column on-the-fly. To do this, we design a lightweight round operator $\mathcal{S}(\mathbf{c}_{\text{res}})$ to map the vector $\mathbf{c}_{\text{res}}$ into one close vector attainable with 2-bit RF phase shifters. In other words, $\mathcal{S}(\mathbf{c}_{\text{res}})$ maps each component $\mathbf{c}_{\text{res}}(\ell)$ to the nearest value $e^{j\phi_\ell}$, with $\phi_\ell \in \left\{-\frac{\pi}{2}, 0, \frac{\pi}{2}, \pi\right\}$. The advantage of the dictionary-free strategy is that, compared to OMP algorithms, we save one matrix-vector multiplication and, therefore, reduce the overall computational complexity.

Lines 7 to 14 refer to the residual update strategy which represents the core of the algorithm. Ideally, if each iteration were independent from the previous one, the objective would

[2]The solution we provide requires only minor modifications to accommodate RF phase shifters with a different number of quantization bits.

[3]The Matlab script implementing Algorithm 1 is available for download from http://wireless.networks.imdea.org/software/.

Algorithm 1 A geometric approach to design hybrid analog-digital multi-directional beam patterns

Require: $\mathbf{C}_{\text{bsl}}$
1: $\mathbf{C}_{\text{RF}} = \text{Empty matrix}$
2: $\mathbf{C}_{\text{res}} = \mathbf{C}_{\text{bsl}}$
3: **for** $n \leq N_{\text{MS}}$ **do**
4: $\quad k = \arg\max_k \|[\mathbf{C}_{\text{res}}]_{:,k}\|_2^2$
5: $\quad \mathbf{c}_{\text{res}} = [\mathbf{C}_{\text{res}}]_{:,k}$
6: $\quad \mathbf{C}_{\text{RF}} = [\mathbf{C}_{\text{RF}}, \mathcal{S}(\mathbf{c}_{\text{res}})]$
7: $\quad \mathcal{M} = \max_i |\mathbf{c}_{\text{res}}(i)|,\ m = \min_i |\mathbf{c}_{\text{res}}(i)|$
8: $\quad J = \text{find}\left[|(\mathbf{c}_{\text{res}}| \geq \frac{\mathcal{M}+m}{2}\right]$
9: $\quad \delta' = \text{mean}\left[\mathbf{c}_{\text{res}}(J)/\mathcal{S}(\mathbf{c}_{\text{res}})(J)\right]$
10: $\quad$ **if** $|\delta'| > \frac{\mathcal{M}+m}{2}$ **then**
11: $\quad\quad \delta = \frac{\delta'}{|\delta'|}\frac{\mathcal{M}+m}{2}$
12: $\quad$ **else**
13: $\quad\quad \delta = \delta'$
14: $\quad$ **end if**
15: $\quad \mathbf{c}_{\text{res}} = \mathbf{c}_{\text{res}} - \delta\mathcal{S}(\mathbf{c}_{\text{res}})$
16: $\quad [\mathbf{C}_{\text{res}}]_{:,k} = \mathbf{c}_{\text{res}}$
17: $\quad$ **for** $i \leq D_{\text{MS}},\ i \neq k$ **do**
18: $\quad\quad [\mathbf{C}_{\text{res}}]_{:,i} = [\mathbf{C}_{\text{res}}]_{:,i} - \mathcal{S}(\mathbf{c}_{\text{res}})\frac{[\mathcal{S}(\mathbf{c}_{\text{res}})]^H[\mathbf{C}_{\text{res}}]_{:,i}}{\|\mathcal{S}(\mathbf{c}_{\text{res}})\|_2^2}$
19: $\quad$ **end for**
20: **end for**
21: $\hat{\mathbf{C}}_{\text{BB}} = \left(\mathbf{C}_{\text{RF}}^H\mathbf{C}_{\text{RF}}\right)^{-1}\mathbf{C}_{\text{RF}}^H\mathbf{C}_{\text{bsl}}$
22: **for** $i \leq D_{\text{MS}}$ **do**
23: $\quad [\mathbf{C}_{\text{BB}}]_{:,i} = \frac{[\hat{\mathbf{C}}_{\text{BB}}]_{:,i}}{\|\mathbf{C}_{\text{RF}}\hat{\mathbf{C}}_{\text{BB}}\|_2}$
24: **end for**
25: **return** $\mathbf{C}_{\text{RF}}$, $\mathbf{C}_{\text{BB}}$

be to find $\delta \in \mathbb{C}$ such that $\mathbf{c}_{\text{res}} - \delta\mathcal{S}(\mathbf{c}_{\text{res}})$ is equal to zero (not in norm, but element by element). However, since the residual in the current iteration depends on the outcomes from the previous one, a good strategy is to not zero $\mathbf{c}_{\text{res}}$, but make it to assume a favorable value not only for the current iteration, but also for the successive ones[4]. According to Line 6 of our algorithm, in fact, the $\mathbf{C}_{\text{RF}}$ column selected at each iteration is the mapped version of the residual. This "altruistic" strategy for residual update tries to share the effort of approaching the baseline combiner among all the available RF chains. We geometrically envision this problem as equivalent to that of finding a good center for the set of points $\frac{\mathbf{c}_{\text{res}}(i)}{\mathcal{S}(\mathbf{c}_{\text{res}})(i)}$ which are distributed inside a 90° circular sector in the complex plane as illustratively shown in Fig. 5. In fact, since $\mathcal{S}(\mathbf{c}_{\text{res}})$ maps the components of $\mathbf{c}_{\text{res}}$ to the closest points with 2-bit quantized phase, each point $\frac{\mathbf{c}_{\text{res}}(i)}{\mathcal{S}(\mathbf{c}_{\text{res}})(i)}$ has a complex argument between $-\pi/4$ and $\pi/4$ and modulus $|\mathbf{c}_{\text{res}}(i)|$. In order to find a good center for this set of points, we first find the maximum $\mathcal{M} = \max_i |\mathbf{c}_{\text{res}}(i)|$ and the minimum $m = \min_i |\mathbf{c}_{\text{res}}(i)|$. Then, we divide the 90° circular sector into two sub-sectors, namely the north pole which includes the points with modulus greater than or equal to $(\mathcal{M}+m)/2$ and the south pole which includes the remaining points. Since, due to the shape of the circular sector, the points near the south pole have less variance compared to the points near the north pole, we focus on the latter, more representative points and calculate their mean value δ'. In Line 8 of Algorithm 1, J is the set of indices relative to the elements of $\mathbf{c}_{\text{res}}$ falling within the north pole of Fig. 5, where the operator "/" in Line 9 represents the element-wise division. If δ' (magenta triangle in Fig. 5) falls within the north pole, we set δ (green cross in Fig. 5) to be the "equator", i.e., $\delta = \frac{\mathcal{M}+m}{2}\frac{\delta'}{|\delta'|}$; otherwise, we set $\delta = \delta'$. The residual matrix is updated in Lines 15 to 19. The process continues until all N_{MS} beamforming vectors have been selected. Finally, the algorithm normalizes the digital baseband combining vector $\hat{\mathbf{C}}_{\text{BB}}$ to satisfy the constraint $\|\mathbf{C}_{\text{RF}}[\mathbf{C}_{\text{BB}}]_{:,i}\|_2^2 = 1$, for $i = 1, 2, ..., D_{\text{MS}}$, and outputs it along with the constructed RF precoding matrix $\mathbf{C}_{\text{RF}}$.

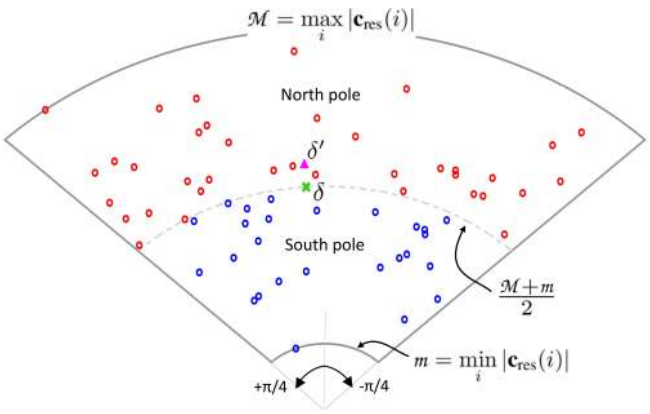


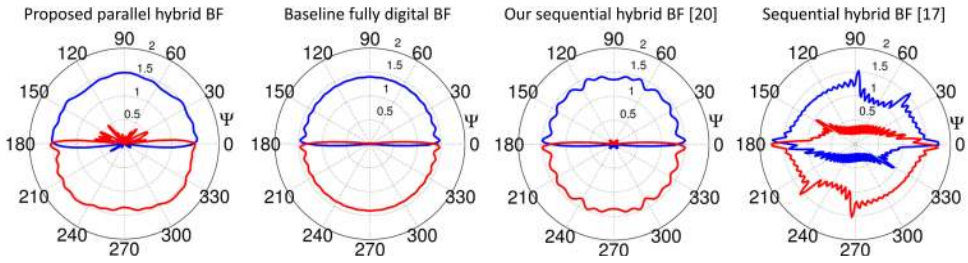
Fig. 5. Illustrative example of the geometric approach to find the parameter δ required for the residual update strategy in Algorithm 1 with 64 antennas.

An example implementation of the ideal case proposed in Fig. 3 is shown in Fig. 6, where hybrid beam patterns for PA-BT obtained with the proposed algorithm are compared against the baseline fully digital BF scheme based on FSM-KW and the hybrid beam patterns for SA-BT obtained in our prior work [20] and in [17]. All the patterns are generated by a device having 64, $\lambda/2$-spaced isotropic antenna elements. As for the hybrid analog-digital design proposed in this work and in [20], a transceiver architecture consisting of 2-bit RF phase shifters and 10 RF chains is considered, while 7-bit RF phase shifters and 32 RF chains are used for the hybrid BF algorithm in [17]. As is evident from the plots, the proposed strategy is able to synthesize multi-beam, multi-beamwidth radiation patterns with limited overlap between adjacent beams and excellent flatness over the covered sectors. Furthermore, it is worth noting that, even with very low complexity hardware compared to the state of the art, the designed hybrid beam patterns are significantly better than those found in the literature and almost indistinguishable from the baseline fully-digital ones.

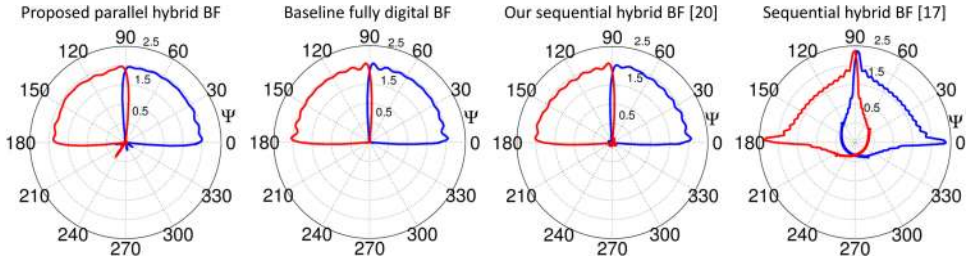
VI. Simulation Results

In this section, we carry out numerical experiments to assess the performance of the adaptive beam training protocols presented in §IV, based on the system model outlined in

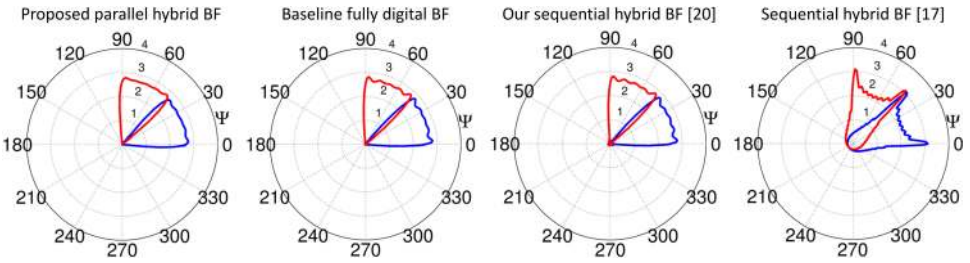
[4]MSE minimization is not a convenient way to attack this problem because: (i) it aims at zeroing the norm instead of the single elements of the residual and (ii) it finds a good solution for the current iteration, without also optimizing for later iterations.



(a) Beam patterns for Stage 1



(b) Beam patterns for Stage 2



(c) Beam patterns for Stage 3

Fig. 6. Beam patterns for the first three stages of adaptive beam training protocols. Comparisons among the shapes synthesized using the proposed hybrid BF algorithm for simultaneous multi-sector scanning, the ideal fully digital BF scheme based on FSM-KW, our hybrid BF algorithm in [20] for sequential single-sector scanning, and the hybrid BF design in [17] for sequential single-sector scanning.

§III and the hybrid analog-digital BF design proposed in §V. We split the problem into two phases as indicated by the frame structure in Fig. 7. In the first phase, BS and MS use Algorithm 1 to shape their beam patterns and perform either SA-BT or PA-BT to estimate the most powerful mm-wave channel paths, i.e., the AoD/AoA pairs exhibiting the higher beamformed channel gains. In the beam training phase, the BS-MS communication is organized according to the frame structure in Fig. 4. After beam training, in a second phase, BS and MS use again Algorithm 1 along with the estimated channel information to build their hybrid precoder $\mathbf{P}_{\text{DATA}}$ (at the BS) and combiner $\mathbf{C}_{\text{DATA}}$ (at the MS). Concretely, $\mathbf{P}_{\text{DATA}}$ and $\mathbf{C}_{\text{DATA}}$ are designed so as to give rise to multi-beam antenna patterns with the narrowest synthesizable beamwidth, pointing towards the L_{est} most reliable spatial directions (i.e., the ones resulting from beam training). As shown later, such precoders/combiners well approach the dominant singular vectors of the channel and hence can be used for data transmission via L_{est} parallel streams to achieve spatial multiplexing gain. Since we are interested in the achievable sum rate, the time slots in the data transmission phase can be indifferently either downlink or uplink slot. In summary, $\mathbf{P}$ (with dimensions $M_{\text{BS}} \times D_{\text{BS}}$) and $\mathbf{C}$ (with dimensions $M_{\text{MS}} \times D_{\text{MS}}$) are the matrices used respectively by BS and MS during beam training, while $\mathbf{P}_{\text{DATA}}$ (with dimensions $M_{\text{BS}} \times L_{\text{est}}$) and $\mathbf{C}_{\text{DATA}}$ (with dimensions $M_{\text{MS}} \times L_{\text{est}}$) are respectively the precoder and combiner matrices used for communication over L_{est} parallel data streams.

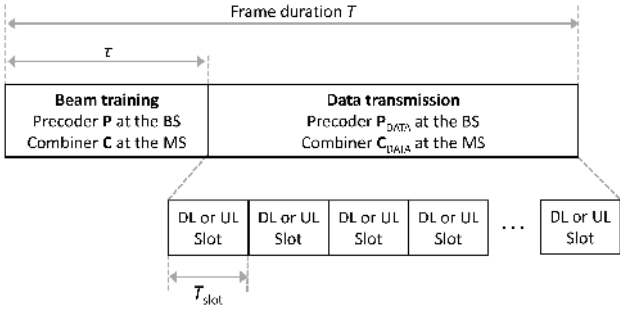


Fig. 7. Frame structure encompassing both the beam training phase and the data transmission phase. Since the achievable sum rate is considered in the simulations, the time slots in the data transmission phase can be indifferently either downlink or uplink slots.

In order to quantify the quality of the channel estimation, we consider the instantaneous spectral efficiency (also referred to as rate in the sequel) experienced in the data transmission period $T-\tau$, which is given by the Shannon sum-rate capacity

[17, Eq. 46]:

$$R = \log_2 \left| \mathbf{I}_{L_{\text{est}}} + \frac{P_t}{L_{\text{est}}} \mathbf{R}_{\text{n}}^{-1} \mathbf{C}_{\text{DATA}}^H \mathbf{H} \mathbf{P}_{\text{DATA}} \mathbf{P}_{\text{DATA}}^H \mathbf{H}^H \mathbf{C}_{\text{DATA}} \right|, \quad (12)$$

where L_{est} is the number of transmitted streams, which is equal to the number of paths estimated via beam training, $\mathbf{R}_{\text{n}}$ is the post-processing noise covariance matrix, i.e., $\mathbf{R}_{\text{n}} = \sigma^2 \mathbf{C}_{\text{DATA}}^H \mathbf{C}_{\text{DATA}}$, and σ^2 is the average noise power. The expression in Eq. 12 represents the mutual information under Gaussian signaling over the BS-MS mm-wave link in Fig. 1, which is maximum when $\mathbf{P}_{\text{DATA}}$ and $\mathbf{C}_{\text{DATA}}$ equal, respectively, the right and left singular vectors of the channel matrix $\mathbf{H}$. Based on this, Eq. 12 can be used as a valid indicator of how good the estimated beam steering directions (given by $\mathbf{P}_{\text{DATA}}$ and $\mathbf{C}_{\text{DATA}}$) are compared to the optimal ones.

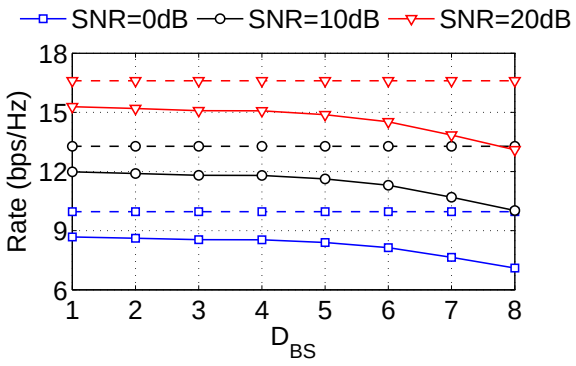
All the simulations in this section adopt the hybrid analog-digital system architecture in Fig. 1 with both BS and MS equipped with ULAs of $\lambda/2$-spaced isotropic antennas. The transmit power at the BS is set to 30 dBm and the system is assumed to operate at 28 GHz carrier frequency with 500-MHz bandwidth. The total training power is distributed over the adaptive estimation stages according to [17, Corollary 2]. For each simulation run, we generate the channel impulse response (CIR) using the mm-wave channel model[5] in Eq. 4. Unless otherwise mentioned, the AoD/AoA resolution parameter N is set to 256, while the parameters K_{BS}, K_{MS}, L_{est}, D_{BS}, and D_{MS} will be defined with each simulation. As for the training sequences, we adopt 128-length Golay sequences while the length of the Walsh code is selected for each simulation depending on the number of simultaneous transmit directions. Finallu, all the numerical results provided in this section are obtained from Monte Carlo simulations with 5000 independent channel realizations for each BS-MS configuration.

A. Impact of multi-directionality

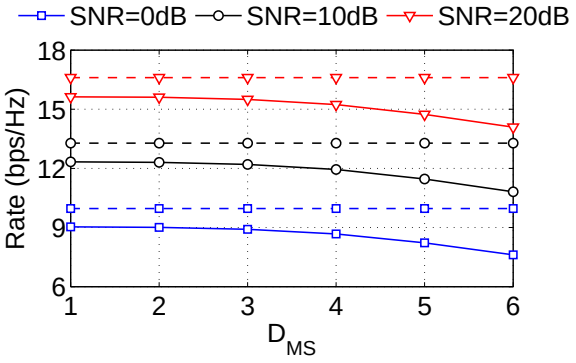
In a first set of simulations, we aim at analyzing the impact of multi-directional scanning on the beam training performance. We consider a BS with $M_{\text{BS}} = 64$ antennas and $N_{\text{BS}} = 10$ RF chains, and a MS with $M_{\text{MS}} = 24$ antennas and $N_{\text{MS}} = 6$ RF chains running the sequential SA-BT and parallel PA-BT protocols described in §IV with only one estimated path (L_{est}=1). There are two main effects of transmitting to and receiving from multiple sectors simultaneously during the PA-BT protocol:

1) the degradation of the multi-beam sector patterns compared to the single-beam sector shapes used in the SA-BT case;
2) the reduced power received by the MS due to the BS transmit power being split among multiple concurrent streams.

In order to isolate the first effect from the second one, we set the distance d between BS and MS to 30 m and tune the level of noise injected into the MS receive chain so that the system operates in a high SNR regime — the effect of power splitting is in fact relevant only at low SNR regimes. In such a scenario, the quality of the channel estimation is directly related to the quality of the beam patterns used during the beam training. In Fig. 8, we investigate the impact of increasing number of directions scanned simultaneously at both the BS and MS sides during the beam training phase. Specifically, in Fig. 8(a), we increase the number of simultaneous transmit directions at the BS, D_{BS}, while the MS is assumed to sequentially receive from one direction at a time (D_{MS}=1)[6]. The results indicate that, for a given SNR, comparable spectral efficiencies are achieved for $D_{\text{BS}} \leq 5$, while rate degradation is experienced for $D_{\text{BS}} > 5$. Similarly, in Fig. 8(b), we focus on the impact of simultaneous, multi-directional reception at the MS, while the BS is assumed to transmit training packets sequentially over one sector at a time (D_{BS}=1)[6]. In this case, for a given SNR, comparable rates are obtained for $D_{\text{MS}} \leq 3$, while performance degradation is experienced for $D_{\text{MS}} > 3$. The reason for this behavior is that, for a given number of RF chains, increasing the number of directions



(a) Multi-directional scanning at the BS



(b) Multi-directional scanning at the MS

Fig. 8. Spectral efficiency when varying the number of simultaneously scanned directions at the BS and MS sides during beam training. In (a), $D_{\text{MS}} = 1$ and K_{BS}=K_{MS}=8, while, in (b), $D_{\text{BS}} = 1$ and K_{BS}=K_{MS}=6. The figure compares the results for different levels of omnidirectional SNR (i.e., the SNR when omnidirectional antennas are used at both the BS and MS sides). Dashed curves refer to the optimum performance with perfect channel knowledge and SVD-based precoding/combining.

[5]The Matlab open-source code is available from http://wireless.engineering.nyu.edu/5g-millimeter-wave-channel-modeling-software/.

[6]Note that the results for D_{BS}=D_{MS}=1 are obtained by running the SA-BT protocol in §IV-A, while the results for any other combination of D_{BS} and D_{MS} are obtained by running the PA-BT protocol in §IV-B.

scanned simultaneously likewise increases the complexity of the multi-beam sector patterns that are to be approximated. Hence, the quality of the beam shapes, and consequently of the channel estimation, strongly depends on the number of RF chains at each device compared to the relative degree of multi-directionality. Exhaustive simulations with different number of antennas and RF chains reveal that, in general, keeping $D_{\text{BS}} \leq N_{\text{BS}}/2$ and $D_{\text{MS}} \leq N_{\text{MS}}/2$ represents the best choice to not incur performance degradation and, at the same time, benefit from the beam training acceleration provided by the simultaneous, multi-directional scanning. It is worth recalling that the rate in Fig. 8 is calculated by Eq. 12, which gives also a direct indication about the channel estimation quality. The BS data precoder $\mathbf{P}_{\text{DATA}}$ and MS data combiner $\mathbf{C}_{\text{DATA}}$ are in fact designed to point towards the dominant channel directions estimated via beam training. In that respect, Fig. 8 shows that, as long as $D_{\text{BS}} \leq N_{\text{BS}}/2$ and $D_{\text{MS}} \leq N_{\text{MS}}/2$, both SA-BT and PA-BT achieve a rate only 10% lower, at most, than that achieved with perfect channel knowledge and SVD-based precoding/combining (dashed lines).

The effect of splitting the BS transmit power among the multiple, simultaneous beams employed during PA-BT is particularly detrimental for-low SNR users (usually referred to as "cell edge users"). To highlight this effect, we run the PA-BT protocol for different values of increasing BS-MS distance d and compute, at each distance, the outage probability — we define as outage the event that the rate delivered to the MS is below a target rate value R_{th}. In Fig. 9, we plot the measured outage probability for R_{th}=0.1 bps/Hz as a function of the SNR at the MS and for different values of D=D_{BS}=D_{MS} (D=1 means that the SA-BT protocol is employed and, therefore, no power splitting occurs at the BS). As expected, the reduced power received by the MS, due to the power split among D concurrent sectors at each stage of the PA-BT protocol, leads to a slightly higher outage probability compared to the SA-BT case. As a solution to this problem, one could use D_{BS} times longer time slots in the first PA-BT stage (see Fig. 3). In fact, since the first stage is the most susceptible one to power splitting because of the low-gain beam patterns, increasing its time slot duration by a factor of D_{BS} provides the required SNR gap of resilience to cope with the power splitting problem[7]. The promising results of adopting this strategy are shown in Fig. 9, where we plot as red dashed line the outage probability achieved using PA-BT with D=3 and using three times longer slots for the first beam training stage. As is evident, the strategy, that we call "Resilient PA-BT", is effective in reducing the outage probability of cell edge users compared to the regular PA-BT approach and closely approaches the SA-BT performance. It is worth recalling that the rate considered in this subsection to determine an outage event is the instantaneous rate experienced in the data transmission phase, i.e., without taking into account the impact of the beam training overhead.

As discussed in §III, another problem arising from transmitting training packets over multiple sectors simultaneously is the inter-sector interference at the MS caused by the adverse multipath propagation. The solution we adopt in our PA-BT strategy is to encode the training sequences transmitted by the BS with orthogonal Walsh spreading codewords. In order to highlight the benefits of this solution, we plot in Fig. 10 the probability of selecting the wrong sector in the first (S1) and second (S2) beam training stage of the PA-BT protocol, with (solid lines) and without (dashed lines) orthogonal encoding. It is evident that orthogonal codes are effective in substantially reducing the estimation error probability. The plot shows also that the Resilient PA-BT strategy with orthogonal encoding (dashed red line without markers) achieves performance very close to that of SA-BT (solid blue line with square markers).

Taken collectively, the results in this sub-section demonstrate the feasibility of using simultaneous multi-directional scanning via hybrid beamforming for mm-wave beam training. In the next subsections, we focus on the effective rate in the frame — computed by normalizing the instantaneous rate to the actual data transmission duration — which provides more insights into the tradeoff between training overhead and overall system performance.

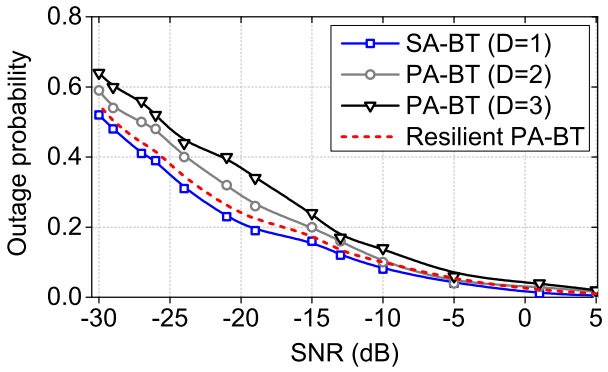


Fig. 9. Outage probability referred to a target rate R_{th}=0.1 bps/Hz versus omnidirectional SNR for different values of D=D_{BS}=D_{MS} with K_{BS}=K_{MS}=3.

B. Training overhead vs. achievable rate tradeoff

In a second set of simulations, we aim at observing and analyzing the well-known tradeoff between the time devoted to beam training and the achievable data transmission rate. As analytically demonstrated in [10], a narrower beamwidth introduces significant search overhead, since many training stages are required to converge to the final beam, but provides a higher transmission rate due to higher directivity gain. In contrast, larger beamwidths require few training stages at the expense of reduced transmission rate.

We consider a BS with $M_{\text{BS}} = 64$ and $N_{\text{BS}} = 10$, and a MS with $M_{\text{MS}} = 24$ and $N_{\text{MS}} = 6$ running the SA-BT and PA-BT protocols to estimate the most dominant path in the channel (L_{est}=1). The distance d between the two devices is set to 70 m, corresponding to an average omnidirectional SNR of 0 dB. The adopted frame structure, based on the work in [10], [29], [30], is depicted in Fig. 7. Each frame has duration $T = 10$ ms and is split into 100 slots with duration

[7]Note that this approach can be generalized by optimizing the slot length based on the beamforming gain at each stage. However, since changing the length of the first-stage slots only provides already sufficient resilience, we leave the more general optimization problem as future work.

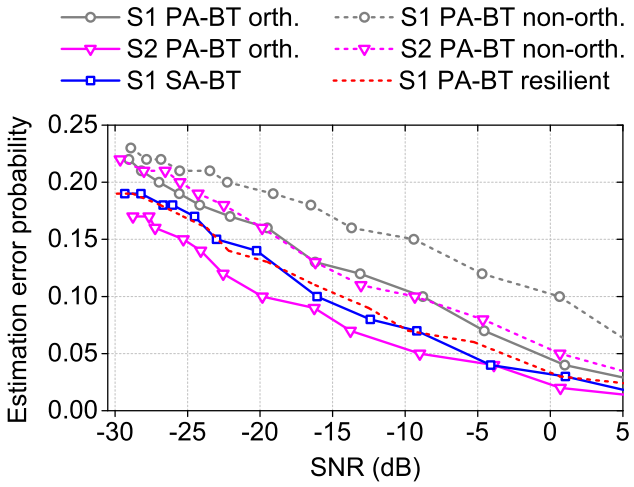


Fig. 10. Probability of selecting the wrong sector during the first (S1) and second (S2) stage of adaptive beam training protocols with K_{BS}=K_{MS}=2. As for the PA-BT strategy, we set D_{BS}=D_{MS}=2 and compare the performance with and without orthogonal encoding.

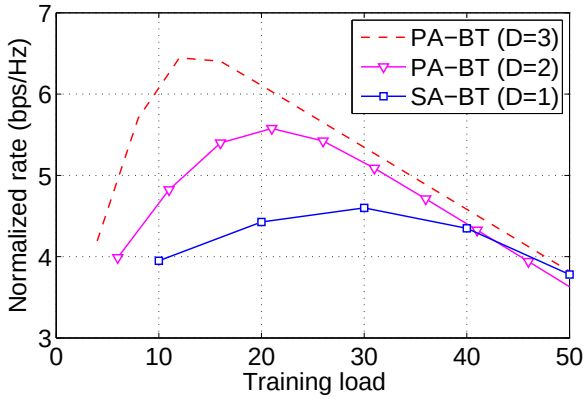


Fig. 11. Achieved normalized rate versus training overhead for different values of D=D_{BS}=D_{MS} with K_{BS}=K_{MS}=3.

$T_{\text{slot}} = 100$ μs, a sufficiently small value to ensure channel coherence at mm-waves.

Instead of fixing the maximum number of training stages S as in Eq. 8, we assume that, in order to speed up the search, BS and MS can choose to stop the beam training protocol at any stage and use the sector pair selected at such stage for data transmission. For example, referring to Fig. 2 and Fig. 3, BS and MS could decide to stop the beam search at Stage 1 and use respectively Sector #2 and Sector #1 for data transmission. Or, they could stop at Stage 2 and exchange data through Sector #3 and Sector #1 respectively.

The performance metric we consider is the normalized rate within a frame, which is obtained by multiplying the rate R in Eq. 12 by the term $(1 - \tau/T)$. Note that, since we have highlighted in the previous sub-section the benefits of using the Resilient PA-BT strategy (which differs from the regular PA-BT approach in having D_{BS} times longer slots in the first beam training stage), we hereafter consider always the Resilient PA-BT strategy in the simulations, referring to it simply as PA-BT. Moreover, although we generically refer to τ as the beam training time, we adopt Eq. 9 to calculate the SA-BT time $\tau_{\text{PA-BT}}$, while a modified version of Eq. 10, accounting for the different time slot duration in the first stage, is used to compute the PA-BT time $\tau_{\text{PA-BT}}$.

In Fig. 11, we plot the normalized rate when varying the training load, defined as the total number of training packets exchanged between BS and MS until the beam search is stopped. The plots show the tradeoff between the training overhead and the achievable rate. That is, increasing the training load results in narrower beamwidth and, therefore, higher transmission rate due to higher directivity gain. However, this is only true until an optimal balance point is reached. From there on, the beam training overhead τ dominates over the data transmission period $T - \tau$, thus resulting in rate degradation. Moreover, exploiting the simultaneous, multi-directional scanning capabilities of hybrid transceivers, more beam training stages within the same time budget can be executed. As a result, compared with the SA-BT strategy, the achievable rate using PA-BT protocol is 20% to 40% greater and the optimal balance point is reached with a 30% to 60% lower training load.

C. Beam training acceleration

In a third set of simulations, we analyze the beam training time and evaluate the impact of the speed-up yielded by the PA-BT strategy on the average data communication rate. To this end, similarly to what is done in the previous simulations, we consider the normalized spectral efficiency within a frame, which is obtained by multiplying the rate R in Eq. 12 by the term $1 - \tau/T$. We consider two different transceiver configurations for the mm-wave system in Fig. 1, namely:

1) Configuration 1: the BS is equipped with $M_{\text{BS}} = 64$ antennas and $N_{\text{BS}} = 10$ RF transceiver chains, while the MS has $M_{\text{MS}} = 24$ antennas and $N_{\text{MS}} = 6$ RF transceiver chains;
2) Configuration 2: the BS is equipped with $M_{\text{BS}} = 32$ antennas and $N_{\text{BS}} = 8$ RF transceiver chains, while the MS has $M_{\text{MS}} = 16$ antennas and $N_{\text{MS}} = 4$ RF transceiver chains.

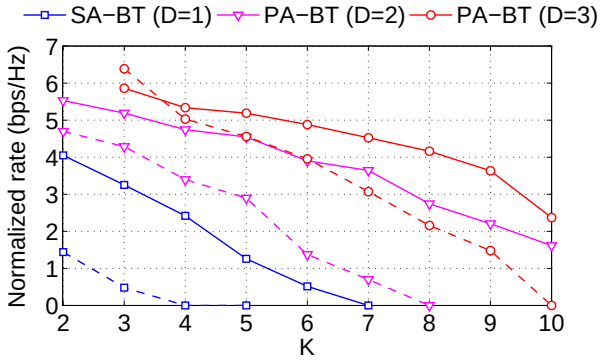
The devices are placed at a distance of d=70 m to each other (corresponding to an average omnidirectional SNR of 0 dB) and are configured to run either the SA-BT protocol or the PA-BT protocol with the same frame structure as in the previous simulations. For the sake of ease of exposition, we assume that the partitioning parameters (K_{BS} and K_{MS}) and the number of directions scanned simultaneously during the beam training phase (D_{BS} and D_{MS}) are the same for both BS and MS, i.e., we define K=K_{BS}=K_{MS} and D=D_{BS}=D_{MS}.

Table I shows, for different values of K, the time τ required to complete the beam training phase when L_{est}=1 — the beam training time for any $L_{\text{est}} > 1$ can be simply obtained by multiplying the table entries by L_{est}[8]. The results demonstrate

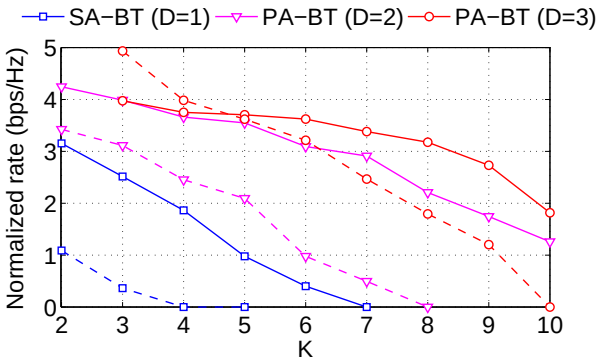
[8]Note that the important outcome in Table I is the relative, rather than the absolute performance. In fact, the reported times are only indicative, since we are considering beam training and data transmission slots with same duration. Note also that we do not provide the beam training times for D=1 and $K > 5$, since they exceed the frame duration T=10 ms, and for D=3 and $K = 2$, since the PA-BT protocol requires $D \leq K$.

TABLE I
BEAM TRAINING TIME (IN MILLISECONDS) FOR DIFFERENT COMBINATIONS OF THE PARAMETERS D=D_{BS}=D_{MS} AND K=K_{BS}=K_{MS}

D	K								
	2	3	4	5	6	7	8	9	10
1 (SA-BT)	4.0	5.0	6.8	7.8	-	-	-	-	-
2 (PA-BT)	1.7	2.9	2.4	3.9	3.9	5.0	5.0	7.7	7.7
3 (PA-BT)	-	1.2	2.8	2.3	2.3	3.8	3.8	3.8	6.6



(a) Configuration 1 with higher number of RF chains and antennas



(b) Configuration 2 with lower number of RF chains and antennas

Fig. 12. Achieved normalized rate versus K=K_{BS}=K_{MS} for different values of the parameter D=D_{BS}=D_{MS}. The solid-line and dashed-line plots refer to the case with L_{est}=1 and L_{est}=2, respectively.

that PA-BT is highly effective in reducing the beam training time by up to 70% compared to SA-BT.

The rate achieved during the data transmission phase using the two BS-MS configurations and normalized according to the beam training times in Table I is plotted in Fig. 12 for both L_{est}=1 (solid lines) and L_{est}=2 (dashed lines). We recall that, in all of the simulations in this paper, the rate for $L_{\text{est}}>1$ is calculated assuming data transmission through L_{est} concurrent streams, whatever beam training strategy is adopted. As expected, the achieved rate decreases by increasing K due to the higher beam training overhead (see Eq. 9 and Eq. 10). It is evident from Fig. 12 that the speed-up of PA-BT over SA-BT allows to significantly boost the achieved rate. Specifically, the maximum achieved rates are around 6.5 bps/Hz and 5 bps/Hz, respectively for Configuration 1 and Configuration 2, and are obtained using the PA-BT strategy with D=K=3 and L_{est}=2 parallel data streams (dashed lines with circle markers). Compared with the best rate achieved by SA-BT (around 4 bps/Hz and 3 bps/Hz, respectively for Configuration 1 and Configuration 2, with K=2 and only one data stream), such result represents, roughly, a 65% rate enhancement. Figure 12 provides another fundamental insight. That is, the long beam training time needed by the SA-BT protocol to estimate multiple AoD/AoA pairs in the mm-wave channel precludes the rate boosting enabled by the multi-stream communication (multiplexing gain). In fact, as is evident from the plot, the rate achieved by SA-BT with L_{est}=2 is significantly lower than that with L_{est}=1. The same behavior, but to less extent, is seen in the PA-BT case with D=2. On the contrary, the very fast PA-BT protocol with D=3 is able to take advantage from the multiplexing gain, yielding approximately 10% to 24% rate improvement when two, instead of only one, data streams are used in the data communication phase.

Finally, we show in Fig. 13 the normalized spectral efficiency when injecting different noise power levels so that the average omnidirectional SNR is varied. We consider the BS-MS Configuration 1 and the most performing setups for our SA-BT and PA-BT protocols, i.e., SA-BT with $K = 2$ and $L_{\text{est}} = 1$, and PA-BT with $K = 3$, $D = 3$, and $L_{\text{est}} = 3$. The performance is compared with our prior work [20] and the reference paper [17], both adopting a SA-BT strategy, and the algorithm in [13] which implements a PA-BT approach with simultaneous multi-directional reception (the number of directions scanned simultaneously is fixed and equal to the number of device's RF chains). Since the genetic algorithm used in [13] to synthesize the hybrid analog-digital radiation patterns is not made available by the authors, we re-implement their PA-BT strategy under ideal conditions of fully digital BF. As is evident from Fig. 13, the proposed PA-BT approach effectively improves the normalized spectral efficiency with respect to SA-BT and PA-BT strategies in the literature. More concretely, at an SNR regime of 0 dB, the average rate achieved by our PA-BT protocol is 66%, 82%, and 74% greater than that obtained by [13], [17], and [20], respectively. It is worth recalling that *our strategy relies on a simple hybrid analog-digital transceiver architecture with only 2-bit RF phase shifters in contrast to the 7-bit RF phase shifters considered in [17] and the fully digital BF architecture used in [13].*

D. Computational complexity

As pointed out in §V-B, one of the main problems of the hybrid BF algorithms in the literature is their large computational complexity. This aspect is crucial, for example, in applications where the transceiver configuration needs to be programmed dynamically based on the propagation environment and link performance to deliver the required quality of service. In such situations, the availability of lightweight, effective, and fast algorithms to synthesize the optimal hybrid precoders and combiners is essential, especially in the case of resource-constrained, battery-powered wireless mobile nodes.

As a final numerical experiment, we analyze the run time required by Algorithm 1 to compute hybrid analog-digital precoders and compare the performance against the OMP

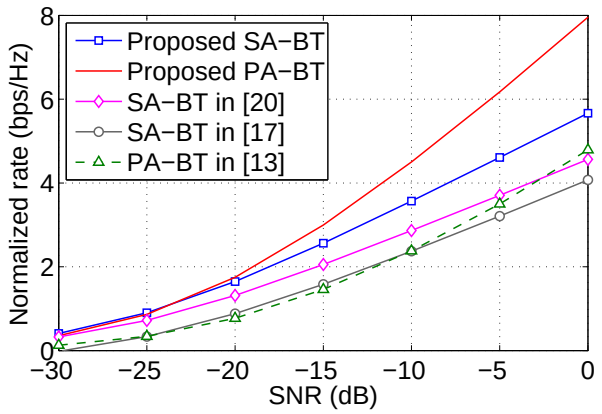
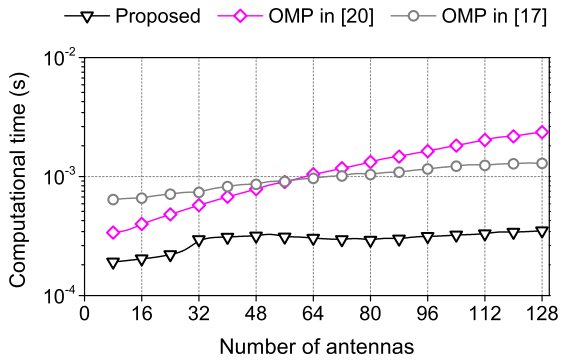
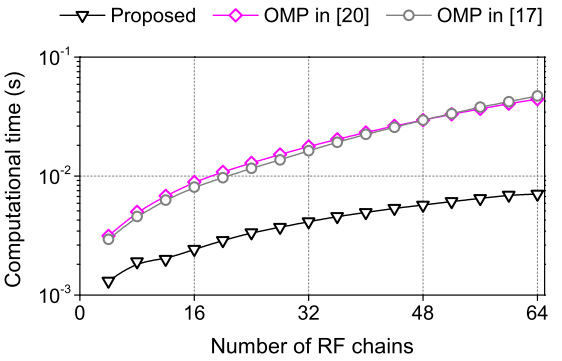


Fig. 13. Normalized rate when varying the average omnidirectional SNR: comparison with the literature. As for the PA-BT protocol in [13], $L_{\text{est}} = 2$ and the antenna patterns are synthesized using a fully digital BF architecture. As for the SA-BT protocols in [17] and [20], we consider the setup which provides the best performance, i.e., the one with K=2 and $L_{\text{est}} = 1$.

algorithms in [20] and [17]. In Fig. 14(a), we consider a hybrid BF architecture with four RF transceiver chains and plot, as a function of the number of antenna elements, the computational times averaged over 5000 Monte-Carlo simulations on a PC with quad-core Intel Core i7 CPU. In the same way, in Fig. 14(b), we consider a hybrid architecture with 64 antennas and plot the average computational times as a function of the number of RF chains. The results demonstrate the computational advantage of our greedy geometric approach which perform approximately five to ten times faster than state-of-the-art, OMP-based techniques.



(a) Hybrid BF architecture with 4 RF transceiver chains



(b) Hybrid BF architecture with 64 antennas

Fig. 14. Average computational time (on a semilogarithmic scale) required to synthesize hybrid analog-digital precoders as function of (a) the number of antennas and (b) the number of RF chains.

VII. Conclusion

In this paper, we investigated the problem of link establishment between mm-wave devices using adaptive beam training protocols. Specifically, we proposed and implemented a low-overhead beam training protocol which exploits the parallel, multi-directional scanning capabilities of hybrid analog-digital transceivers. To accomplish that, we developed an effective algorithm to synthesize multi-beamwidth, multi-beam antenna patterns to be used during the parallel adaptive beam training stages. In contrast to almost all of the approaches in the literature, which are based on computationally intensive techniques, our design relies on a lightweight, greedy geometric algorithm which was shown to be able to shape beam patterns almost indistinguishable from those attained by fully digital BF, yet requiring lower complexity hardware compared with the state of the art. Simulation results showed that the speed up of the beam training phase, enabled by the ability of hybrid transceivers to scan multiple directions simultaneously, provided up to 82% increase in spectral efficiency compared to state-of-the-art strategies that adopt sequential, single-directional scanning during beam training. Additionally, we considered the well-known tradeoff between the training overhead and the achieved directivity gain and showed that, in order to reach the optimal balance point, the proposed parallel protocol requires a training load 60% lower than that needed by sequential beam training strategies. As future work, we intend to extend the current setting to more complex scenarios with multiple mm-wave devices and to investigate the effect of user mobility on the efficiency of the beam training protocols.

Acknowledgments

The research leading to these results received funding from the European Commission H2020 programme under grant agreement no○ 671650 (5G PPP mmMAGIC project), the European Research Council grant ERC CoG 617721, the Madrid Regional Government through the TIGRE5-CM program (S2013/ICE-2919), and the Spanish Government through the Ramon y Cajal grant (RYC-2012-10788).

References

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[2] mmMAGIC. Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications, EU 5G-PPP H2020-ICT-2014-2 Project. [Online]. Available: http://5g-mmmagic.eu/

[3] J. Wang *et al.*, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[4] L. Chen, Y. Yang, X. Chen, and W. Wang, "Multi-stage beamforming codebook for 60GHz WPAN," in *2011 6th International ICST Conference on Communications and Networking in China (CHINACOM)*, Aug. 2011.

[5] J. Wang *et al.*, "Beamforming codebook design and performance evaluation for 60GHz wideband WPANs," in *2009 IEEE 70th Vehicular Technology Conference Fall*, Sept 2009.

[6] B. Li, Z. Zhou, H. Zhang, and A. Nallanathan, "Efficient beamforming training for 60-GHz millimeter-wave communications: A novel numerical optimization framework," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 703–717, Feb 2014.
[7] Y. Tsang, A. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmWave communication system," in *2011 IEEE Global Telecommunications Conference (GLOBECOM)*, Houston, TX, USA, Dec. 2011.
[8] A. Valdes-Garcia *et al.*, "A SiGe BiCMOS 16-element phased-array transmitter for 60Ghz communications," in *2010 IEEE Solid-State Circuits Conference (ISSCC)*, Feb. 2010.
[9] D. Pepe and D. Zito, "A novel phase shifter for 60 Ghz phased arrays," in *2015 Irish Signals and Systems Conf. (ISSC)*, Jun. 2015.
[10] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione, "Beam-searching and transmission scheduling in millimeter wave communications," in *2015 IEEE International Conference on Communications (ICC)*, June 2015.
[11] C. N. Barati *et al.*, "Directional cell discovery in millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6664–6678, Dec. 2015.
[12] M. C. Filippou *et al.*, "Throughput vs. latency: QoS-centric resource allocation for multi-user millimeter wave systems," in *2017 IEEE International Conference on Communications (ICC)*, May 2017.
[13] S. Payami, M. Shariat, M. Ghoraishi, and M. Dianati, "Effective RF codebook design and channel estimation for millimeter wave communication systems," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, Jun. 2015.
[14] J. Song, J. Choi, and D. J. Love, "Codebook design for hybrid beamforming in millimeter wave systems," in *2015 IEEE International Conference on Communications (ICC)*, London, UK, Jun. 2015.
[15] O. El Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *2012 IEEE International Conference on Communications (ICC)*, Jun. 2012.
[16] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[17] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
[18] M. E. Eltayeb, A. Alkhateeb, R. W. Heath, and T. Y. Al-Naffouri, "Opportunistic beam training with hybrid analog/digital codebooks for mmWave systems," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, USA, Dec. 2015.
[19] F. Sohrabi and Y. Wei, "Hybrid beamforming with finite-resolution phase shifters for large-scale MIMO systems," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2015.
[20] D. De Donno, J. Palacios, D. Giustiniano, and J. Widmer, "Hybrid analog-digital beam training for mmWave systems with low-resolution RF phase shifters," in *2016 IEEE ICC Workshop on 5G RAN Design*, May 2016.
[21] J. Palacios, D. De Donno, D. Giustiniano, and J. Widmer, "Speeding up mmWave beam training through low-complexity hybrid transceivers," in *27th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2016.
[22] K. H. Sayidmarie and Q. H. Sultan, "Synthesis of wide beam array patterns using quadratic-phase excitations," *International Journal of Electromagnetics and Applications*, vol. 3, no. 6, pp. 127–135, Jun. 2013.
[23] A. Yehezkely and S. Kupferman, "Modular millimeter-wave radio frequency system," https://www.google.com/patents/US20120309331, Dec. 2012, uS Patent App. 13/312,127.
[24] Y. Atesal, B. Cetinoneri, M. Chang, R. Alhalabi, and G. Rebeiz, "Millimeter-wave wafer-scale silicon BiCMOS power amplifiers using free-space power combining," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 4, pp. 954–965, Apr. 2011.
[25] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, June 2014.
[26] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Single-sided adaptive estimation of multi-path millimeter wave channels," in *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2014.
[27] J. Palacios, D. D. Donno, and J. Widmer, "Lightweight and effective sector beam pattern synthesis with uniform linear antenna arrays," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, no. 1, pp. 1–4, Dec 2017.
[28] S. J. Orfinidas, "Electromagnetic waves and antennas," http://www.ece.rutgers.edu/~orfanidi/ewa/.
[29] F. Khan and Z. Pi, "mmWave mobile broadband (MMB): Unleashing the 3–300Ghz spectrum," in *34th IEEE Sarnoff Symposium*, May 2011.
[30] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, and S. Rangan, "5G mmWave module for the ns-3 network simulator," in *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2015, pp. 283–290.

Danilo De Donno received the B.Sc. and M.Sc. degrees (cum laude) in telecommunication engineering from Politecnico di Milano, Italy, in 2005 and 2008, respectively, and the Ph.D. degree in Information Engineering from the University of Salento, Lecce, Italy, in 2012. He was a Post-Doctoral Fellow with the ElectroMagnetic Lab Lecce (EML^2) of the University of Salento from 2012 to 2015. In July 2015, he joined the Pervasive Wireless Systems Group and the Wireless Networking Group at the IMDEA Networks Institute, Madrid, Spain, as a Post-Doc Researcher. His areas of interest lie in mm-Wave communications and wireless LANs, with main research focus on the development of efficient schemes for beam training and tracking through hybrid analog-digital beamforming.

Joan Palacios received his B.Sc. degree in mathematics in 2015 from the Universitat de Valencia, Spain, and his M.Sc degree in multimedia and communications in 2016 from the Universidad Carlos III de Madrid, Spain. In September 2015, he joined the IMDEA Networks Institute, Madrid, Spain as a PhD student. His areas of interest lie in mm-Wave communications with main research focus on the development of efficient schemes for beam training and tracking, hybrid analog-digital beamforming, and Angle Difference of Arrival (ADoA) based localization for quasi-optical communication systems.

Joerg Widmer is Research Professor at IMDEA Networks Institute in Madrid, Spain. His research focuses primarily on wireless networks, ranging from extremely high frequency millimeter-wave communication and MAC layer design to mobile network architectures. From 2005 to 2010, he was manager of the Ubiquitous Networking Research Group at DOCOMO Euro-Labs in Munich, Germany, leading several projects in the area of mobile and cellular networks. Before, he worked as post-doctoral researcher at EPFL, Switzerland on ultra-wide band communication and network coding. He was a visiting researcher at the International Computer Science Institute in Berkeley, CA, USA and University College London, UK. Joerg Widmer authored more than 150 conference and journal papers and three IETF RFCs, holds 13 patents, serves on the editorial board of IEEE Transactions on Communications, and regularly participates in program committees of several major conferences. He was awarded an ERC consolidator grant, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation, as well as a Spanish Ramon y Cajal grant. He is senior member of IEEE and ACM.