Energy-Efficient Interactive Beam Alignment for Millimeter-Wave Networks

Muddassar Hussain and Nicolo Michelusi⬢, *Senior Member, IEEE*

Abstract—Millimeter-wave will be a key technology in next-generation wireless networks thanks to abundant bandwidth availability. However, the use of large antenna arrays with beamforming demands precise beam alignment between the transmitter and the receiver and may entail huge overhead in mobile environments. This paper investigates the design of an optimal interactive beam alignment and data communication protocol, with the goal of minimizing power consumption under a minimum rate constraint. The base station selects beam alignment or data communication and the beam parameters, based on the feedback from the user end. Based on the sectored antenna model and uniform prior on the angles of departure and arrival (AoD/AoA), the optimality of a fixed-length beam-alignment phase followed by a data-communication phase is demonstrated. Moreover, a decoupled fractional beam-alignment method is shown to be optimal, which decouples the alignment of AoD and AoA over time, and iteratively scans a fraction of their region of uncertainty. A heuristic policy is proposed for non-uniform prior on AoD/AoA, with provable performance guarantees, and it is shown that the uniform prior is the worst-case scenario. The performance degradation due to detection errors is studied analytically and via simulation. The numerical results with analog beams depict up to 4dB, 7.5dB, and 14dB gains over a state-of-the-art bisection method and conventional and interactive exhaustive search policies, respectively, and demonstrate that the sectored model provides valuable insights for beam-alignment design.

Index Terms—Millimeter-wave, beam-alignment, initial access, Markov decision process.

I. Introduction

MOBILE traffic has witnessed a tremendous growth over the last decade, 18-folds over the past five years alone, and is expected to grow with a compound annual growth rate of 47% from 2016 to 2021 [2]. This rapid increase poses a severe burden to current systems operating below 6 GHz, due to limited bandwidth availability. Millimeter-wave (mm-wave) is emerging as a promising solution to enable multi-Gbps communication, thanks to abundant bandwidth availability [3]. However, high isotropic path loss and sensitivity to blockages pose challenges in supporting high capacity and mobility [4]. To overcome the path loss, mm-wave systems will thus lever-age narrow-beams, by using large antenna arrays at both base stations (BSs) and user-ends (UEs).

Nonetheless, narrow transmission and reception beams are susceptible to frequent loss of alignment, due to mobility or blockage, which necessitate the use of beam-alignment protocols. Maintaining beam-alignment between transmitter and receiver can be challenging, especially in mobile scenarios, and may entail significant overhead, thus potentially offsetting the benefits of mm-wave directionality. Therefore, it is imperative to design schemes to mitigate its overhead.

To address this challenge, in our previous work [5]–[8], we address the optimal design of beam-alignment protocols. In [5], we optimize the trade-off between data communication and beam-sweeping, under the assumption of an exhaustive search method, in a mobile scenario where the BS widens its beam to mitigate the uncertainty on the UE position. In [6] and [7], we design a throughput-optimal beam-alignment scheme for one and two UEs, respectively, and we prove the optimality of a *bisection search*. However, the model therein does not consider the energy cost of beam-alignment, which may be significant when targeting high detection accuracy. It is noteworthy that, if the energy consumption of beam-alignment is small, bisection search is the best policy since it is the fastest way to reduce the uncertainty region of the angles of arrival (AoA) and departure (AoD). For this reason, it has been employed in previous works related to multi-resolution codebook design, such as [9]. In [1] and [8], we incorporate the energy cost of beam-alignment, and prove the optimality of a *fractional search* method. Yet, in [5]–[8], optimal design is carried out under restrictive assumptions that the UE receives isotropically, and that the duration of beam-alignment is fixed. In practice, the BS may switch to data transmission upon finding a strong beam, as in [10], and *both* BS and UE may use narrow beams to fully leverage the beamforming gain.

To the best of our knowledge, the optimization of *interactive* beam-alignment, *jointly* at both BS and UE, is still an open problem. Therefore, in this paper, we consider a more flexible model than our previous papers [5]–[8], by allowing dynamic switching between beam-alignment and data-communication and joint optimization over BS-UE beams, BS transmission power and rate. *Indeed, we prove that a fixed-length beam-alignment scheme followed by data communication is optimal, and we prove the optimality of a decoupled fractional search method, which decouples over time the alignment of AoD and AoA, and iteratively scans a fraction of their region of uncertainty.* Using Monte-Carlo simulation with analog beams, we demonstrate superior performance, with up to 4dB, 7.5dB, and 14dB power gains over the state-of-the-art bisection method [9], conventional exhaustive, and

interactive exhaustive search policies, respectively. Compared to our recent paper [1], the system model adopted in this paper is more realistic since it captures the effects of fading and resulting outages, non-uniform priors on AoD/AoA, and detection errors. Additionally, the model in [1] is restricted to a two-phase protocol with deterministic beam-alignment duration. In this paper, we show that this is indeed optimal.

A. Related Work

Beam-alignment has been a subject of intense research due to its importance in mm-wave communications. The research in this area can be categorized into beam-sweeping [5]–[9], [11]–[13], data-assisted schemes [14]–[17], and AoD/AoA estimation [18], [19]. The simplest and yet most popular beam-sweeping scheme is *exhaustive search* [11], which sequentially scans through all possible BS-UE beam pairs and selects the one with maximum signal power. A version of this scheme has been adopted in existing mm-wave standards including IEEE 802.15.3c [20] and IEEE 802.11ad [21]. An interactive version of exhaustive search has been proposed in [10], wherein the beam-alignment phase is terminated once the power of the received beacon is above a certain threshold. The second popular scheme is *iterative search* [12], where scanning is first performed using wider beams followed by refinement using narrow beams. A variant of *iterative search* is studied in [22], where the beam sequence is chosen adaptively from a pre-designed multi-resolution codebook. However, this codebook is designed independently of the beam-alignment protocol, thereby potentially resulting in suboptimal design. In [13], the authors consider the design of a beamforming vector sequence based on a partially observable (PO-) Markov decision processes (MDPs). However, POMDPs are generally not amenable to closed-form solutions, and have high complexity. To reduce the computational overhead, the authors focus on a greedy algorithm, which yields a sub-optimal policy.

Data-aided schemes utilize the information from sensors to aid beam-alignment and reduce the beam-sweeping cost (e.g., from radar [14], lower frequencies [15], position information [16], [17]). AoD/AoA estimation schemes leverage the sparsity of mm-wave channels, and include compressive sensing schemes [18] or approximate maximum likelihood estimators [19]. In [23], the authors compare different schemes and conclude that the performance of beam-sweeping is comparable with the best performing estimation schemes based on compressed sensing. Yet, beam-sweeping has the added advantage of low complexity over compressed sensing schemes, which often involve solving complex optimization problems, and is more amenable to analytical insights on the beam-alignment process. For these reasons, in this paper we focus on beam-sweeping, and derive insights on its optimal design.

All of the aforementioned schemes choose the beam-alignment beams from pre-designed codebooks, use heuristic protocols, or are not amenable to analytical insights. By choosing the beams from a restricted beam-space or a predetermined protocol, optimality may not be achieved. Moreover, all of these papers do not consider the
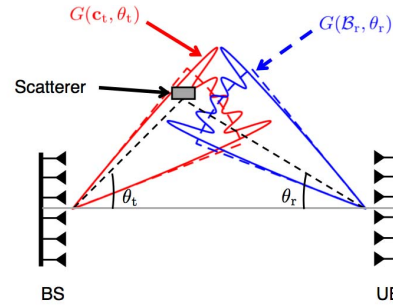

Fig. 1. Actual beam pattern $G(\mathbf{c}_\mathrm{x}, \theta_\mathrm{x})$ generated using the algorithm in [22] with $M_\mathrm{t} = M_\mathrm{r} = 128$ antennas. (solid lines) versus sectored model $G(\mathcal{B}_\mathrm{x}, \theta_\mathrm{x})$ (dashed lines) [24], on a linear scale. Sidelobes are not visible due to their small magnitude.

energy and/or time overhead of beam-alignment as part of their design. In this paper, we address these open challenges by optimizing the beam-alignment protocol to maximize the communication performance.

B. Our Contributions

1) Based on a MDP formulation, under the *sectored antenna model* [24], uniform AoD/AoA prior, and small detection error assumptions, we prove the optimality of a *fixed-length* two-phase protocol, with a beam-alignment phase of fixed duration followed by a data communication phase. We provide an algorithm to compute the optimal duration.

2) We prove the optimality of a *decoupled fractional search* method, which scans a fixed fraction of the region of uncertainty of the AoD/AoA in each beam-alignment slot. Moreover, the beam refinements over the AoD and AoA dimensions are decoupled over time, thus proving the sub-optimality of *exhaustive search* methods.

3) Inspired by the decoupled fractional search method, we propose a heuristic scheme for the case of non-uniform prior on AoD/AoA with provable performance, and prove that the uniform prior is indeed the worst-case scenario.

4) We analyze the effect of detection errors on the performance of the proposed protocol.

5) We evaluate its performance via simulation using analog beams, and demonstrate up to $4\mathrm{dB}$, $7.5\mathrm{dB}$, and $14\mathrm{dB}$ power gains compared to the state-of-the-art bisection scheme [9], conventional and interactive exhaustive search policies, respectively. Remarkably, the sectored model provides valuable insights for beam-alignment design.

The rest of this paper is organized as follows. In Secs. II, we describe the system model. In Sec. III, we formulate the optimization problem. In Secs. IV-VI, we provide the analysis for the case of uniform and non-uniform priors on AoD/AoA. In Sec. VII, we analyze the effects of false-alarm and misdetection errors. The numerical results are provided in Sec. VIII, followed by concluding remarks in Sec. IX. The main analytical proofs are provided in the Appendix.

II. SYSTEM MODEL

We consider a downlink scenario in a mm-wave cellular system with one base-station (BS) and one mobile user (UE)

at distance d from the BS, both equipped with uniform linear arrays (ULAs) with M_t and M_r antennas, respectively, depicted in Fig. 1. Communication occurs over frames of fixed duration T_{fr}, each composed of N slots indexed by $\mathcal{I} \equiv \{0, 1, \ldots, N-1\}$ of duration $T = T_{fr}/N$, each carrying S symbols of duration $T_{sy} = T/S$. Let s be the transmitted symbol, with $\mathbb{E}[|s|^2] = 1$. Then, the signal received at the UE is

$$y = \sqrt{P} \mathbf{c}_r^H \mathbf{H} \mathbf{c}_t s + \mathbf{c}_r^H \mathbf{w}, \qquad (1)$$

where P is the average transmit power of the BS; $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ is the channel matrix; $\mathbf{c}_t \in \mathbb{C}^{M_t}$ is the BS beam-forming vector; $\mathbf{c}_r \in \mathbb{C}^{M_r}$ is the UE combining vector; $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, N_0 W_{tot} \mathbf{I})$ is additive white Gaussian noise (AWGN). The symbols N_0 and W_{tot} denote the one-sided power spectral density of AWGN and the system bandwidth, respectively. By assuming analog beam-forming at both BS and UE, $\mathbf{c}_t$ and $\mathbf{c}_r$ satisfy the unit norm constraints $\|\mathbf{c}_t\|_2^2 = \|\mathbf{c}_r\|_2^2 = 1$. The channel matrix $\mathbf{H}$ follows the extended Saleh-Valenzuela geometric model [25],

$$\mathbf{H} = \sqrt{\frac{M_t M_r}{K}} \sum_{\ell=1}^{K} h_\ell \, \mathbf{d}_r(\theta_{r,\ell}) \mathbf{d}_t^H(\theta_{t,\ell}), \qquad (2)$$

where $h_\ell \in \mathbb{C}$, $\theta_{t,\ell}$ and $\theta_{r,\ell}$ denote the small scale fading coefficient, AoD and AoA of the ℓth cluster, respectively. The terms $\mathbf{d}_r(\theta_{r,\ell}) \in \mathbb{C}^{M_r}$ and $\mathbf{d}_t(\theta_{t,\ell}) \in \mathbb{C}^{M_t}$ are the UE and BS array response vectors, respectively. For ULAs, $\theta_{t,\ell}$ (respectively, $\theta_{r,\ell}$) is the angle formed between the outgoing (incoming) rays of the lth channel cluster and the perpendicular to the BS (UE) antenna array, as represented in Fig. 1, so that

$$\mathbf{d}_x(\theta_x) = \frac{1}{\sqrt{M_x}} \left[1, e^{j\frac{2\pi d_x}{\lambda} \sin \theta_x}, \ldots, e^{j(M_x-1)\frac{2\pi d_x}{\lambda} \sin \theta_x} \right]^\top,$$

where $x \in \{t, r\}$, d_t and d_r are the antenna spacing of the BS and UE arrays, respectively, λ is the wavelength of the carrier signal. In (2), $K \geq \text{rank}(\mathbf{H})$ is the total number of clusters. Note that $\mathbf{H}$ has low-rank if $K \ll \min\{M_t, M_r\}$. In this paper, we assume that there is a single dominant cluster ($K = 1$). This assumption has been adopted in several previous works (e.g., see [26], [27]), and is motivated by channel measurements and modeling works such as [3], where it is shown that, in *dense urban environments*, with high probability the mm-wave channel exhibits only one or two clusters, with the dominant one containing most of the signal energy. While our analysis is based on a single cluster model, in Sec. VIII we demonstrate by simulation that the proposed scheme is robust also against multiple clusters. For the single cluster model, we obtain

$$\mathbf{H} = \sqrt{M_t M_r} h \, \mathbf{d}_r(\theta_r) \mathbf{d}_t^H(\theta_t), \qquad (3)$$

where $\mathbb{E}[|h|^2] = 1/\ell(d)$, $\ell(d)$ denotes the path loss between BS and UE as a function of distance d, and $\boldsymbol{\theta} = (\theta_t, \theta_r)$ is the single-cluster AoD/AoA pair. We assume that $\boldsymbol{\theta}$ has prior joint distribution $f_0(\boldsymbol{\theta})$ with support $\text{supp}(f_0) = \mathcal{U}_{t,0} \times \mathcal{U}_{r,0}$, which reflects the availability of prior AoD/AoA information acquired from previous beam-alignment phases, or based on geometric constraints (*e.g.*, presence of buildings blocking the signal in certain directions). We assume that h and $\boldsymbol{\theta}$ do not change over a frame, whose duration T_{fr} is chosen

based upon the channel and beam coherence times T_c and T_b (time duration over which the AoD/AoA do not change appreciably) to satisfy this property. In [28], it has been reported that $T_c \ll T_b$. In the numerical values given below, $T_b \sim 100 T_c$. Therefore, by choosing $T_{fr} \leq T_c$, we ensure that the variations in h and $\boldsymbol{\theta}$ over the frame duration T_{fr} are small and can be ignored. For example, using the relationships of T_c and T_b in [28], we obtain $T_c \simeq 10$[ms] and $T_b \simeq 1$[s] for a UE velocity of 100[km/h]. In our numerical evaluations, we will therefore use $T_{fr} = 10$[ms]. It is noteworthy that this assumption has also been used extensively in previous beam-alignment works, such as [18], [19], and [26].

We assume that blockage occurs at longer time-scales than the frame duration, determined by the geometry of the environment and mobility of users, hence we neglect blockage dynamics within a frame duration [29]. By replacing (3) into (1), and defining the BS and UE beam-forming gains $G_x(\mathbf{c}_x, \theta_x) = M_x |\mathbf{d}_x^H(\theta_x) \mathbf{c}_x|^2$, $x \in \{t, r\}$, we get

$$y = h \sqrt{P G_t(\mathbf{c}_t, \theta_t) \cdot G_r(\mathbf{c}_r, \theta_r)} e^{j\Psi(\boldsymbol{\theta})} s + \hat{w}, \qquad (4)$$

where $\hat{w} \triangleq \mathbf{c}_r^H \mathbf{w} \sim \mathcal{CN}(0, N_0 W_{tot})$ is the noise component and $\Psi(\boldsymbol{\theta}) = \angle \mathbf{d}_t^H(\theta_t) \mathbf{c}_t - \angle \mathbf{d}_r^H(\theta_r) \mathbf{c}_r$ is the phase.

In this paper, we use the *sectored antenna* model [24] to approximate the BS and UE beam-forming gains, represented in Fig. 1. Under this model,

$$G_x(\mathbf{c}_x, \theta_x) \approx G_x(\mathcal{B}_x, \theta_x) = \frac{2\pi}{|\mathcal{B}_x|} \chi_{\mathcal{B}_x}(\theta_x), \ x \in \{t, r\}, \qquad (5)$$

where $\mathcal{B}_t \subseteq (-\pi, \pi]$ is the range of AoD covered by $\mathbf{c}_t$, $\mathcal{B}_r \subseteq (-\pi, \pi]$ is the range of AoA covered by $\mathbf{c}_r$, $\chi_{\mathcal{A}}(\theta)$ is the indicator function of the event $\theta \in \mathcal{A}$, and $|\mathcal{A}| = \int_{\mathcal{A}} d\theta$ is the measure of the set $\mathcal{A}$. Hereafter, the two sets $\mathcal{B}_t$ and $\mathcal{B}_r$ will be referred to as BS and UE beams, respectively. Additionally, we define $\mathcal{B}_k = \mathcal{B}_{t,k} \times \mathcal{B}_{r,k}$ as the 2-dimensional (2D) AoD/AoA support defined by the BS-UE beams. Note that the sectored model is used as an abstraction of the real model, which applies a precoding vector $\mathbf{c}_t$ at the transmitter and a beamforming vector $\mathbf{c}_r$ at the receiver. This abstraction, shown in Fig. 1, is adopted since direct optimization of $\mathbf{c}_t$ and $\mathbf{c}_r$ is not analytically tractable, due to the high dimensionality of the problem. In Sec. VIII we show via Monte-Carlo simulation that, by appropriate design of $\mathbf{c}_t$ and $\mathbf{c}_r$ to approximate the sectored model, our scheme attains near-optimal performance, and outperforms a state-of-the-art bisection search scheme [9]; thus, the sectored antenna model provides a valuable abstraction for practical design. Following the sectored antenna model, we obtain the received signal by replacing $G_x(\mathbf{c}_x, \theta_x)$ with $G_x(\mathcal{B}_x, \theta_x)$ in (4), yielding

$$y = h \sqrt{P G_t(\mathcal{B}_t, \theta_t) \cdot G_r(\mathcal{B}_r, \theta_r)} e^{j\Psi(\boldsymbol{\theta})} s + \hat{w}. \qquad (6)$$

Although the analysis in this paper is presented for ULAs (2D beamforming), the proposed scheme can be extended to the case of uniform planar arrays with 3D beamforming, by interpreting $\theta_x, x \in \{t, r\}$ as a vector denoting the azimuth and elevation pair in $(-\pi, \pi]^2$ and the beam $\mathcal{B}_x \subseteq (\pi, \pi]^2$. For notational convenience and ease of exposition, in this paper we focus on the 2D beamforming case (also adopted in, e.g., [9], [18], [22], [23]).

The entire frame duration is split into two, possibly interleaved phases: a beam-alignment phase, whose goal is to detect the best beam to be used in the data communication phase. To this end, we partition the slots $\mathcal{I}$ in each frame into the indices in the set $\mathcal{I}_s$, reserved for beam-alignment, and those in the set $\mathcal{I}_d$, reserved for data communication, where $\mathcal{I}_s \cap \mathcal{I}_d = \emptyset$ and $\mathcal{I}_s \cup \mathcal{I}_d = \mathcal{I}$. The optimal frame partition and duration of beam-alignment are part of our design. In the sequel, we describe the operations performed in the beam-alignment and data communication slots, and characterize their energy consumption.

A. Beam-Alignment

At the beginning of each slot $k \in \mathcal{I}_s$, the BS sends a beacon signal $\mathbf{s}$ of duration $T_B < T$ using the transmit beam $\mathcal{B}_{t,k}$ with power P_k,[1] and the UE receives the signal using the receive beam $\mathcal{B}_{r,k}$. Note that $\mathcal{B}_k = \mathcal{B}_{t,k} \times \mathcal{B}_{r,k}$ and P_k are design parameters. If the UE detects the beacon (*i.e.*, the AoD/AoA $\boldsymbol{\theta}$ is in $\mathcal{B}_k$, or a false-alarm occurs, see [31]), then, in the remaining portion of the slot of duration $T - T_B$, it transmits an acknowledgment (ACK) packet to the BS, denoted as $C_k = \text{ACK}$. Otherwise (the UE does not detect the beacon due to either mis-alignment or misdetection error), it transmits $C_k = \text{NACK}$. We assume that the ACK/NACK signal C_k is received perfectly and within the end of the slot by the BS (for instance, by using a conventional microwave technology as a control channel [32]).

As a result of (6), the UE attempts to detect the beam, and generates the ACK/NACK signal based on the following hypothesis testing problem,

$$\mathcal{H}_1 : \mathbf{y}_k = \sqrt{N_0\,W_{\text{tot}}\nu_k}\,h e^{j\Psi_k(\boldsymbol{\theta})}\mathbf{s} + \hat{\mathbf{w}}_k, \quad \text{(alignment, } \boldsymbol{\theta} \in \mathcal{B}_k)$$
$$\mathcal{H}_0 : \mathbf{y}_k = \hat{\mathbf{w}}_k, \quad \text{(misalignment, } \boldsymbol{\theta} \notin \mathcal{B}_k)$$

where $\mathbf{y}_k$ is the received signal vector, $\mathbf{s}$ is the transmitted symbol sequence, $\hat{\mathbf{w}}_k \sim \mathcal{CN}(\mathbf{0}, N_0\,W_{\text{tot}}\mathbf{I})$ is the AWGN vector, and ν_k is related to the beam-forming gain in slot k,

$$\nu_k = \frac{(2\pi)^2 \, P_k}{N_0\,W_{\text{tot}}|\mathcal{B}_k|}. \tag{7}$$

The optimal detector depends on the availability of prior information on h. We assume that an estimate of the channel gain $\gamma = |h|^2$ is available at the BS and UE at the beginning of each frame, denoted as $\hat{\gamma} = |\hat{h}|^2$, where $\hat{h} = h + e$ and $e \sim \mathcal{CN}(0, \sigma_e^2)$ denotes the estimation noise. A Neyman-Pearson threshold detector is optimal in this case,

$$\frac{|\mathbf{s}^H \mathbf{y}_k|^2}{N_0\,W_{\text{tot}}\|\mathbf{s}\|_2^2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \tau_{\text{th}}. \tag{8}$$

The detector's threshold τ_{th} and the transmission power P_k are designed based on the channel gain estimate $\hat{\gamma}$, so as to satisfy constraints on the false-alarm and misdetection probabilities, $p_{\text{fa}}, p_{\text{md}} \leq p_e$. We now compute these probabilities under the simplifying assumption that $\hat{h}$ and e are independent, so that $h|\hat{h} \sim \mathcal{CN}(\hat{h}, \sigma_e^2)$. Let $z_k \triangleq \frac{\mathbf{s}^H \mathbf{y}_k}{\sqrt{N_0\,W_{\text{tot}}}\|\mathbf{s}\|_2}$, so that $|z_k|^2$ is the

decision variable. We observe that

$$z_k = \begin{cases} \sqrt{\nu_k}\,h e^{j\Psi_k(\boldsymbol{\theta})}\|\mathbf{s}\|_2 + \dfrac{\mathbf{s}^H \hat{\mathbf{w}}_k}{\sqrt{N_0\,W_{\text{tot}}}\|\mathbf{s}\|_2}, & \text{if } \mathcal{H}_1 \text{ is true;} \\[2mm] \dfrac{\mathbf{s}^H \hat{\mathbf{w}}_k}{\sqrt{N_0\,W_{\text{tot}}}\|\mathbf{s}\|_2}, & \text{if } \mathcal{H}_0 \text{ is true.} \end{cases}$$

Since $\hat{h}$ and $\hat{\mathbf{w}}_k$ are independent and $h = \hat{h} - e$, we obtain

$$f(z_k|\hat{h}, \mathcal{H}_1, \boldsymbol{\theta}) = \mathcal{CN}\left(\sqrt{\nu_k}\,\hat{h} e^{j\Psi_k(\boldsymbol{\theta})}\|\mathbf{s}\|_2, 1 + \nu_k\|\mathbf{s}\|_2^2\sigma_e^2\right), \tag{9}$$
$$f(z_k|\hat{h}, \mathcal{H}_0) = \mathcal{CN}(0, 1), \tag{10}$$

so that $[|z_k|^2|\hat{h}, \mathcal{H}_0] \sim \text{Exponential}(1)$, and the false-alarm probability can be expressed as

$$p_{\text{fa}}(\tau_{\text{th}}) \triangleq \mathbb{P}\left(|z_k|^2 > \tau_{\text{th}}|\hat{h}, \mathcal{H}_0\right) = \exp(-\tau_{\text{th}}). \tag{11}$$

Similarly, the misdetection probability is found to be

$$p_{\text{md}}(\nu_k, \tau_{\text{th}}, \hat{\gamma}) \triangleq \mathbb{P}\left(|z_k|^2 < \tau_{\text{th}}|\hat{h}, \mathcal{H}_1\right)$$
$$= 1 - Q_1\left(\sqrt{\frac{2\hat{\gamma}\nu_k\|\mathbf{s}\|_2^2}{1 + \nu_k\|\mathbf{s}\|_2^2\sigma_e^2}}, \sqrt{\frac{2\tau_{\text{th}}}{1 + \nu_k\|\mathbf{s}\|_2^2\sigma_e^2}}\right), \tag{12}$$

where $Q_1(\cdot)$ is the first-order Marcum's Q function [33]. In fact, $z_k|(\hat{h}, \mathcal{H}_1)$ is complex Gaussian as in (9), so that, given $(\hat{\gamma}, \mathcal{H}_1)$, $\frac{2|z_k|^2}{1 + \nu_k\|\mathbf{s}\|_2^2\sigma_e^2}$ follows non-central chi-square distribution with 2 degrees of freedom and non-centrality parameter $\frac{2\nu_k\hat{\gamma}\|\mathbf{s}\|_2^2}{1 + \nu_k\|\mathbf{s}\|_2^2\sigma_e^2}$.

Herein, we design τ_{th} and P_k to achieve $p_{\text{fa}}, p_{\text{md}} \leq p_e$. To satisfy $p_{\text{fa}}(\tau_{\text{th}}) \leq p_e$ we need

$$\tau_{\text{th}} \geq -\ln(p_e). \tag{13}$$

Since $Q_1(a, b)$ is an increasing function of $a \geq 0$ and a decreasing function of $b \geq 0$, it follows that $p_{\text{md}}(\nu_k, \tau_{\text{th}})$ is a decreasing function of $\nu_k \geq 0$ and an increasing function of $\tau_{\text{th}} \geq 0$. Then, to guarantee $p_{\text{md}}(\nu_k, \tau_{\text{th}}, \hat{\gamma}) \leq p_e$, (13) should be satisfied with equality to attain the smallest p_{md}; additionally, there exists $\nu^* > 0$, determined as the unique solution of $p_{\text{md}}(\nu^*, \tau_{\text{th}}, \hat{\gamma}) = p_e$ and independent of the beam shape $\mathcal{B}_k$, such that $p_{\text{md}}(\nu_k, \tau_{\text{th}}, \hat{\gamma}) \leq p_e$ iff (if and only if) $\nu_k \geq \nu^*$. Then, using (7) and letting $E_k \triangleq P_k T_{\text{sy}}\|\mathbf{s}\|_2^2$ be the energy incurred for the transmission of the beacon $\mathbf{s}$ in slot k, E_k should satisfy

$$E_k \geq \phi_s(p_e)|\mathcal{B}_k|, \tag{14}$$
$$\text{where } \phi_s(p_e) \triangleq N_0\,W_{\text{tot}}\nu^* T_{\text{sy}}\|\mathbf{s}\|_2^2/(2\pi)^2 \tag{15}$$

is the energy/rad^2 required to achieve false-alarm and misdetection probabilities equal to p_e.

Note that false-alarm and misdetection errors are deleterious to performance, since they result in mis-alignment and outages during data transmission. Therefore, they should be minimized. For this reason, in the first part of this paper we assume that $p_e \ll 1$, and neglect the impact of these errors on beam-alignment. Thus, we let $E_k \geq \phi_s|\mathcal{B}_k|$ be the energy required in each beam-alignment slot to guarantee successful detection with high probability, where ϕ_s is computed under some small $p_e \ll 1$. We will consider the impact of these errors in Sec. VII.[2]

---

[1]In practice, there are limits on how small the beacon duration can be made, due to peak power constraints [30], beacon synchronization errors [4], and auto-correlation properties of the beacon sequence [4].

[2]The design of beam-alignment schemes robust to errors when $p_e \not\ll 1$ has been considered in [34]. Its analysis is outside the scope of this paper.

B. Data Communication

In the communication slots indexed by $k \in \mathcal{I}_d$, the BS uses $\mathcal{B}_{\mathrm{t},k}$, rate R_k, and transmit power P_k, while the UE processes the received signal using the beam $\mathcal{B}_{\mathrm{r},k}$. Therefore, letting $\gamma = |h|^2$ and ν_k as in (7), the instantaneous SNR can be expressed as

$$\mathsf{SNR}_k = \frac{\gamma P_k G_{\mathrm{t}}(\mathcal{B}_{\mathrm{t},k}, \theta_{\mathrm{t}}) G_{\mathrm{r}}(\mathcal{B}_{\mathrm{r},k}, \theta_{\mathrm{r}})}{N_0 \, W_{\mathrm{tot}}}$$
$$= \nu_k \gamma \chi_{\mathcal{B}_{\mathrm{t},k}}(\theta_{\mathrm{t}}) \chi_{\mathcal{B}_{\mathrm{r},k}}(\theta_{\mathrm{r}}). \qquad (16)$$

Outage occurs if $W_{\mathrm{tot}} \log_2(1+\mathsf{SNR}_k) < R_k$ due to either mis-alignment between transmitter and receiver, or low channel gain γ. The probability of this event, p_{out}, can be inferred from the posterior probability distribution of the AoD/AoA pair $\boldsymbol{\theta}$ and the channel gain γ, given its estimate $\hat{\gamma}$, and the history of BS-UE beams and feedback until slot k, denoted as $\mathcal{H}^k \triangleq \{(\mathcal{B}_0, C_0), \ldots, (\mathcal{B}_{k-1}, C_{k-1})\}$. Thus, $p_{out} \triangleq \mathbb{P}(W_{\mathrm{tot}} \log_2(1 + \mathsf{SNR}_k) < R_k | \hat{\gamma}, \mathcal{H}^k)$, yielding

$$p_{out} \overset{(a)}{=} \mathbb{P}\left(\mathsf{SNR}_k < 2^{\frac{R_k}{W_{\mathrm{tot}}}} - 1 | \hat{\gamma}, \boldsymbol{\theta} \in \mathcal{B}_k\right) \mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k)$$
$$+ \, \mathbb{P}(\boldsymbol{\theta} \notin \mathcal{B}_k | \mathcal{H}^k)$$
$$\overset{(b)}{=} 1 - \bar{F}_\gamma \left(\frac{2^{\frac{R_k}{W_{\mathrm{tot}}}} - 1}{\nu_k} \Big| \hat{\gamma} \right) \mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k), \qquad (17)$$

where (a) follows from the law of total probability and $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k)$ denotes the probability of correct beam-alignment; (b) follows by substituting $\bar{F}_\gamma(x|\hat{\gamma}) \triangleq \mathbb{P}(\gamma \geq x|\hat{\gamma})$ into (a), given as

$$\bar{F}_\gamma(x|\hat{\gamma}) = Q_1\left(\sqrt{2\hat{\gamma}/\sigma_e^2}, \sqrt{2x/\sigma_e^2}\right). \qquad (18)$$

Herein, we use the notion of ϵ-outage capacity to design R_k, defined as the largest transmission rate such that $p_{out} \leq \epsilon$, for a target outage probability $\epsilon < 1$. This can be expressed as

$$\mathcal{C}_\epsilon(P_k, \mathcal{B}_k | \mathcal{H}^k, \hat{\gamma})$$
$$\triangleq W_{\mathrm{tot}} \log_2 \left(1 + \nu_k \bar{F}_\gamma^{-1} \left(\frac{1-\epsilon}{\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k)} \Big| \hat{\gamma} \right)\right), \qquad (19)$$

where $\bar{F}_\gamma^{-1}(\cdot|\hat{\gamma})$ denotes the inverse posterior CCDF of γ, conditional on $\hat{\gamma}$. In other words, if $R_k \leq \mathcal{C}_\epsilon(P_k, \mathcal{B}_k | \mathcal{H}^k, \hat{\gamma})$, then the transmission is successful with probability at least $1-\epsilon$, and the average rate is at least $(1-\epsilon)R_k$. Note that, in order to achieve the target $p_{out} \leq \epsilon$, the probability of correct beam-alignment must satisfy $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k) \geq 1-\epsilon$. This can be achieved with a proper choice of $\mathcal{B}_k$, as discussed next.

Since the ACK/NACK feedback after data communication is generated by higher layers (*e.g.,* network or transport layer), we do not use it to improve beam-alignment. We define $C_k = \mathrm{NULL}$, $\forall k \in \mathcal{I}_d$, to distinguish it from the ACK/NACK feedback signal in the beam-alignment slots.

III. PROBLEM FORMULATION

In this section, we formulate the optimization problem, and characterize it as a Markov decision process (MDP). The goal is to minimize the power consumption at the BS over a frame duration, while achieving the quality of service (QoS)

requirements of the UE (rate and delay). Therefore, the objective function of the following optimization problem captures the beam-alignment and data communication energy costs; the QoS requirements are specified in the constraints through a rate requirement $R_{\min}$ of the UE along with an outage probability of ϵ; additionally, the frame duration T_{fr} represents a delay guarantee on data transmission. The design variables in slot k are denoted by the 4-tuple $\mathbf{a}_k = (\xi_k, P_k, \mathcal{B}_k, R_k)$, where ξ_k corresponds to the decision of whether to perform beam-alignment ($\xi_k = 1$) or data communication ($\xi_k = 0$); we let $R_k = 0$ for beam-alignment slots ($\xi_k = 1$). With this choice of $\mathbf{a}_k$, we aim to optimally select the beam-alignment slots $\mathcal{I}_s$ and data communication slots $\mathcal{I}_d$. If a slots is selected for beam-alignment ($\xi_k = 1$), we aim to optimize the associated power P_k and 2D beam $\mathcal{B}_k$. Likewise, if a slot is selected for data communication ($\xi_k = 0$), we aim to optimize the associated power P_k, data rate R_k, and 2D beam $\mathcal{B}_k$. Mathematically, the optimization problem is stated as

$$\mathrm{P}_1 : \quad \bar{P} \triangleq \min_{\mathbf{a}_0, \ldots, \mathbf{a}_{N-1}} \frac{1}{T_{\mathrm{fr}}} \mathbb{E}\left[\sum_{k=0}^{N-1} E_k \Big| f_0 \right] \qquad (20)$$

s.t. $\mathbf{a}_k = (\xi_k, P_k, \mathcal{B}_k, R_k), \quad \forall k,$

$$\mathcal{B}_k = \mathcal{B}_{\mathrm{t},k} \times \mathcal{B}_{\mathrm{r},k} \subseteq [-\pi, \pi]^2, \quad \forall k, \qquad (21)$$

$$E_k \geq \phi_s |\mathcal{B}_k|, \ \forall k \in \mathcal{I}_s, \qquad (22)$$

$$\frac{1}{N} \sum_k R_k \geq R_{\min}, \ R_k \leq \mathcal{C}_\epsilon(P_k, \mathcal{B}_k | \mathcal{H}^k, \hat{\gamma}), \forall k \in \mathcal{I}_d, (23)$$

$$P_k = E_k / [\xi_k T_B + (1 - \xi_k) T], \ \forall k, \qquad (24)$$

where f_0 in (20) denotes the prior belief over $\boldsymbol{\theta}$; (21) defines the 2D beam $\mathcal{B}_k$; (22) gives the energy consumption in the beam-alignment slots; (23) ensures the rate requirement $R_{\min}$ over the frame, and that R_k is within the ϵ-outage capacity, see (19); (24) gives the relation between energy and power.[3] Since the cost is the average BS power consumption, the inequality constraints (22)-(23) must be tight, i.e., we replace them with

$$E_k = \xi_k \phi_s |\mathcal{B}_k| + (1 - \xi_k) \frac{\psi_d(R_k) |\mathcal{B}_k|}{\bar{F}_\gamma^{-1}\left(\frac{1-\epsilon}{\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k)} | \hat{\gamma} \right)}, \qquad (25)$$

$$\frac{1}{N} \sum_k R_k = R_{\min}, \qquad (26)$$

where (25) when $\xi_k = 0$ is obtained by inverting (23) via (19) and (7) (with equality) to find P_k and $E_k = P_k T$, and we have defined the energy/rad^2 required to achieve the rate R

$$\psi_d(R) \triangleq (2\pi)^{-2} N_0 \, W_{\mathrm{tot}} T (2^{\frac{R}{W_{\mathrm{tot}}}} - 1).$$

Hereafter, we exclude P_k from the design space, since it is uniquely defined by the set of equality constraints (24)-(25). Thus, we simplify the design variable to $\mathbf{a}_k = (\xi_k, \mathcal{B}_k, R_k)$.

We pose P_1 as an MDP [35] over the time horizon $\mathcal{I}$. The state at the start of slot k is (f_k, D_k), where f_k is the probability distribution over the AoD/AoA pair $\boldsymbol{\theta}$, given the

---

[3]Data communication takes the entire slot, whereas beam-alignment occurs over a portion $T_B < T$ of the slot to allow for the time to receive the ACK/NACK feedback from the receiver.

history $\mathcal{H}^k$ up to slot k, denoted as *belief*; D_k is the backlog (untransmitted data bits). Initially, f_0 is the prior belief and $D_0 \triangleq R_{\min} T_{\text{fr}}$. Given (f_k, D_k), the BS and UE select $\mathbf{a}_k = (\xi_k, \mathcal{B}_k, R_k)$.[4] Then, the UE generates the feedback signal: if $\xi_k = 0$ (data communication), then $C_k = \text{NULL}$; if $\xi_k = 1$ (beam-alignment), then $C_k = \text{ACK}$ if $\boldsymbol{\theta} \in \mathcal{B}_k$, with probability

$$\mathbb{P}(C_k = \text{ACK}|f_k, \mathbf{a}_k) = \int_{\mathcal{B}_k} f_k(\boldsymbol{\theta}) d\boldsymbol{\theta}, \qquad (27)$$

and $C_k = \text{NACK}$ otherwise. Upon receiving C_k, the new backlog in slot $k+1$ becomes[5]

$$D_{k+1} = \max\{D_k - R_k T, 0\}, \qquad (28)$$

and the new belief f_{k+1} is computed via Bayes' rule, as given in the following lemma.

Lemma 1: Let f_0 be the prior belief on $\boldsymbol{\theta}$ with support $\text{supp}(f_0) = \mathcal{U}_0$. Then,

$$f_k(\boldsymbol{\theta}) = \frac{f_0(\boldsymbol{\theta})}{\int_{\mathcal{U}_k} f_0(\tilde{\boldsymbol{\theta}}) \mathrm{d}\tilde{\boldsymbol{\theta}}} \chi_{\mathcal{U}_k}(\boldsymbol{\theta}), \qquad (29)$$

where $\mathcal{U}_k \triangleq \text{supp}(f_k)$ is updated recursively as

$$\mathcal{U}_{k+1} = \begin{cases} \mathcal{U}_k \cap \mathcal{B}_k, & k \in \mathcal{I}_{\text{s}}, C_k = \text{ACK} \\ \mathcal{U}_k \setminus \mathcal{B}_k, & k \in \mathcal{I}_{\text{s}}, C_k = \text{NACK} \\ \mathcal{U}_k, & k \in \mathcal{I}_{\text{d}}. \end{cases} \qquad (30)$$

Proof: The proof follows by induction using Bayes' rule. In fact, if $C_k = \text{ACK}$ in a beam-alignment slot, then it can be inferred that $\boldsymbol{\theta} \in \mathcal{U}_k \cap \mathcal{B}_k$; otherwise ($C_k = \text{NACK}$) the UE lies outside $\mathcal{B}_k$, but within the support of f_k, i.e., $\boldsymbol{\theta} \in \mathcal{U}_k \setminus \mathcal{B}_k$. In the data communication slots, no feedback is generated, hence $f_{k+1} = f_k$ and $\mathcal{U}_{k+1} = \mathcal{U}_k$. A detailed proof is given in [36]. ∎

Lemma 1 implies that $\mathcal{U}_k$ is a sufficient statistic for decision making in slot k, and is updated recursively via (30). Accordingly, the state space is defined as

$$\mathcal{S} \equiv \{(\mathcal{U}, D) : \mathcal{U} \subseteq \mathcal{U}_0, 0 \le D \le D_0\}. \qquad (31)$$

Given the data backlog $D_k = D$, the action space is expressed as[6]

$$\mathcal{A}(D) \equiv \{(0, \mathcal{B}, R) : \mathcal{B} \equiv \mathcal{B}_{\text{t}} \times \mathcal{B}_{\text{r}} \subseteq [-\pi, \pi]^2, 0 < R \le D/T\}$$
$$\cup \{(1, \mathcal{B}, 0) : \mathcal{B} \equiv \mathcal{B}_{\text{t}} \times \mathcal{B}_{\text{r}} \subseteq [-\pi, \pi]^2\}. \qquad (32)$$

Given $(\mathcal{U}_k, D_k) \in \mathcal{S}$, the action $\mathbf{a}_k \in \mathcal{A}(D_k)$ is chosen based on policy μ_k, which determines the BS-UE beam $\mathcal{B}_k$ and whether to perform beam-alignment ($\xi_k = 1$, $R_k = 0$) or data communication ($\xi_k = 0$, $R_k > 0$), with energy cost E_k given by (25). With this notation, we can express the problem P_1 as that of finding the policy μ^* which minimizes the power consumption under rate requirement and outage probability constraints,

$$\text{P}_2: \ \bar{P} \triangleq \min_{\mu} \frac{1}{T_{\text{fr}}} \mathbb{E}_{\mu} \left[\sum_{k=0}^{N-1} c(\mathbf{a}_k; \mathcal{U}_k, D_k) \middle| \mathcal{U}_0, D_0, f_0 \right],$$
$$\text{s.t. } D_{k+1} = D_k - T R_k, \quad \forall k \in \mathcal{I}, \ D_N = 0, \qquad (33)$$

---

[4]Since feedback is error-free, both BS and UE have the same information to generate the action $\mathbf{a}_k$ and their beams.

[5]If $D_{k+1} \le 0$, all bits have been transmitted.

[6]Note that, for a data communication action $(0, \mathcal{B}, R)$, we assume that $R > 0$; in fact, data communication with zero rate is equivalent to a beam-alignment action $(1, \emptyset, 0)$ with empty beam.

where we have defined the cost per stage in state $(\mathcal{U}_k, D_k)$ under action $\mathbf{a}_k$ as

$$c(\mathbf{a}_k; \mathcal{U}_k, D_k) = \left[\xi_k \phi_s + \frac{(1 - \xi_k) \psi_d(R_k)}{\bar{F}_\gamma^{-1} \left(\frac{1-\epsilon}{\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{U}_k)} |\hat{\gamma} \right)} \right] |\mathcal{B}_k|, \quad (34)$$

and we used the sufficient statistic (Lemma 1) to express $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{H}^k) = \mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{U}_k)$ in (25). P_2 can be solved via dynamic programming (DP): the value function in state $(\mathcal{U}_k, D_k)$ under action $\mathbf{a}_k \in \mathcal{A}(D_k)$, $V_k(\mathbf{a}_k; \mathcal{U}_k, D_k)$, and the optimal value function, $V_k^*(\mathcal{U}_k, D_k)$, are expressed as

$$V_k(\mathbf{a}_k; \mathcal{U}_k, D_k) = c(\mathbf{a}_k; \mathcal{U}_k, D_k)$$
$$+ \mathbb{E}\left[V_{k+1}^*(\mathcal{U}_{k+1}, D_{k+1}) \middle| \mathcal{U}_k, D_k; \mathbf{a}_k \right],$$
$$V_k^*(\mathcal{U}_k, D_k) = \min_{\mathbf{a}_k \in \mathcal{A}(D_k)} V_k(\mathbf{a}_k; \mathcal{U}_k, D_k), \qquad (35)$$

where the minimum is attained by the optimal policy. To enforce $D_N = 0$, we initialize it as

$$V_N^*(\mathcal{U}_N, D_N) = \begin{cases} 0, & D_N = 0 \\ \infty, & D_N > 0. \end{cases} \qquad (36)$$

Further analysis is not doable for a generic prior f_0. To unveil structural properties, we proceed as follows:

1) We optimize over the extended action space

$$\mathcal{A}_{\text{ext}}(D) \equiv \{(0, \mathcal{B}, R) : \mathcal{B} \subseteq [-\pi, \pi]^2, 0 < R \le D/T\}$$
$$\cup \{(1, \mathcal{B}, 0) : \mathcal{B} \subseteq [-\pi, \pi]^2\}, \quad (37)$$

obtained by removing the "rectangular beam" constraint $\mathcal{B} \equiv \mathcal{B}_{\text{t}} \times \mathcal{B}_{\text{r}}$ in (32). Thus, $\mathcal{B} \in \mathcal{A}_{\text{ext}}(D)$ can be any subset of $[-\pi, \pi]^2$, not restricted to a "rectangular" shape $\mathcal{B} \equiv \mathcal{B}_{\text{t},k} \times \mathcal{B}_{\text{r},k}$. By optimizing over an extended action space, a lower bound to the value function is obtained, denoted as $\hat{V}_k^*(\mathcal{U}_k, D_k) \le V_k^*(\mathcal{U}_k, D_k)$, possibly not achievable by a "rectangular" beam.

2) In Sec. IV, we find structural properties under such extended action space, for the case of a uniform belief f_0. In this setting, we prove the optimality of a *fractional search* method, which selects $\mathcal{B}_k$ as $\mathcal{B}_k \subseteq \mathcal{U}$ with $|\mathcal{B}_k| = \rho_k |\mathcal{U}_k|$ (beam-alignment) or $|\mathcal{B}_k| = \vartheta |\mathcal{U}_k|$ (data communication), for appropriate *fractional parameters* ρ_k and ϑ; additionally, we prove the optimality of a *deterministic* duration of the beam-alignment phase (Theorems 1 and 3).

3) In Sec. V, we prove that such lower bound is indeed achievable by a *decoupled fractional search* method, which decouples the BS and UE beam-alignment over time using rectangular beams, hence it is optimal.

4) In Sec. VI, we use these results to design a heuristic policy with performance guarantees for the case of non-uniform prior f_0, and show that the uniform prior is the worst case.

IV. UNIFORM PRIOR

We denote the beam $\mathcal{B}$ taking value from the extended action space $\mathcal{A}_{\text{ext}}(D)$ as "2D beam", to distinguish it from $\mathcal{B} \in \mathcal{A}(D)$, that obeys a "rectangular" constraint. Additionally, since the goal is to minimize the energy consumption, we restrict $\mathcal{B} \subseteq \mathcal{U}$

during data communication and $\mathcal{B} \subset \mathcal{U}$ during beam-alignment, yielding the following extended action space in state $(\mathcal{U}, D)$[7]:

$$\mathcal{A}_{\text{ext}}(\mathcal{U}, D) \equiv \{(0, \mathcal{B}, R) : \mathcal{B} \subseteq \mathcal{U}, 0 < R \le D/T\}$$
$$\cup \{(1, \mathcal{B}, 0) : \mathcal{B} \subset \mathcal{U}\}. \quad (38)$$

In this section, we consider the independent uniform prior on $\boldsymbol{\theta} = (\theta_t, \theta_r)$, i.e.,

$$f_0(\boldsymbol{\theta}) = f_{r,0}(\theta_r) \cdot f_{t,0}(\theta_t), \ f_{x,0}(\theta_x) = \frac{\chi_{\mathcal{U}_{x,0}}(\theta_x)}{|\mathcal{U}_{x,0}|}. \quad (39)$$

From Lemma 1, it directly follows that f_k is *uniform* in its support $\mathcal{U}_k$, and the state transition probabilities from state $(\mathcal{U}_k, D_k)$ under the beam-alignment action $(1, \mathcal{B}_k, 0) \in \mathcal{A}_{\text{ext}}(\mathcal{U}, D)$, given in (27) for the general case, can be specialized as $D_{k+1} = D_k$ and

$$\mathcal{U}_{k+1} = \begin{cases} \mathcal{B}_k, & \text{w.p. } \dfrac{|\mathcal{B}_k|}{|\mathcal{U}_k|}, \\ \mathcal{U}_k \setminus \mathcal{B}_k, & \text{w.p. } 1 - \dfrac{|\mathcal{B}_k|}{|\mathcal{U}_k|}, \end{cases} \quad (40)$$

where "w.p." abbreviates "with probability". On the other hand, under the data communication action $(0, \mathcal{B}_k, R_k)$, the new state becomes $\mathcal{U}_{k+1} = \mathcal{U}_k$, and $D_{k+1} = D_k - R_k T$.

In order to determine the optimal policy with extended action set, we proceed as follows:

1) In Sec. IV-A, we find the structure of the optimal data communication beam, as a function of the transmit rate R_k and support $\mathcal{U}_k$, and investigate its energy cost;
2) Next, in Sec. IV-B, we prove that it is suboptimal to perform beam-alignment *after* data communication within the frame. Instead, it is convenient to narrow down the beam as much as possible via beam-alignment, to achieve the most energy-efficient data communication;
3) Finally, in Sec. IV-C, we investigate the structure of the value function, to prove the optimality of a *fixed-length* beam-alignment and of a *fractional-search method*.

A. Optimal Data Communication Beam

In the following theorem, we find the optimal 2D beam for data communication.

Theorem 1: *In any communication slot* $k \in \mathcal{I}_d$, *the 2D beam* $\mathcal{B}_k$ *is optimal iff*

$$\mathcal{B}_k \subseteq \mathcal{U}_k |\mathcal{B}_k| = \vartheta |\mathcal{U}_k|, \quad (41)$$

where $\vartheta = (1-\epsilon)/q^*$, *with* $q^* = \arg\max_{q \in [1-\epsilon, 1]} q \bar{F}_\gamma^{-1}(q|\hat\gamma)$.

Proof: The proof is provided in Appendix A. ∎

The significance of this result is that the optimal beam in the data communication phase is a *fraction* ϑ of the region of uncertainty $\mathcal{U}_k$, with ϑ reflecting the desired outage constraint. By substituting (41) into (34), and letting

$$\phi_d(R, \epsilon) \triangleq \frac{\psi_d(R)(1-\epsilon)}{q^* \bar{F}_\gamma^{-1}(q^*|\hat\gamma)} \quad (42)$$

be the energy/rad^2 to achieve transmission rate R with outage probability ϵ, the cost per stage of a data communication action

with beam given by Theorem 1 can be expressed as

$$c(\mathbf{a}_k; \mathcal{U}_k, D_k) = \phi_d(R_k, \epsilon)|\mathcal{U}_k|. \quad (43)$$

B. Beam-Alignment Before Data Communication Is Optimal

In Theorem 2, we prove that it is suboptimal to precede data communication to beam-alignment. Instead, it is more energy efficient to narrow down the beam as much as possible via beam-alignment, before switching to data communication.

Theorem 2: *Let* μ *be a policy and* $\{(\mathcal{U}_k, D_k), k \in \mathcal{I}\}$ *be a realization of the state process under* μ *such that* $\exists j : \xi_j(\mathcal{U}_j, D_j) = 0$ *and* $\xi_{j+1}(\mathcal{U}_{j+1}, D_{j+1}) = 1$ *(beam-alignment is followed by data communication, for some slot* j*). Then,* μ *is suboptimal.*

Proof: The theorem is proved in two parts using contradiction. The first part deals with the case when a data communication slot is followed by a beam-alignment slot having non-zero beam-width. The second part deals with the case when a data communication slot is followed by a beam-alignment slot having zero beam-width. Let μ be a policy such that, for some state $(\mathcal{U}_j, D_j)$ and slot index j, $\mu_j(\mathcal{U}_j, D_j) = (0, \mathcal{B}_j, R_j)$, satisfying the conditions of Theorem 1 (data communication action); thus, the state at $j + 1$ is $(\mathcal{U}_{j+1}, D_{j+1}) = (\mathcal{U}_j, D_j - TR_j)$. Further, assume that, in this state, $\mu_{j+1}(\mathcal{U}_j, D_j - TR_j) = (1, \mathcal{B}_{j+1}, 0)$ (beam-alignment), with $\mathcal{B}_{j+1} \subset \mathcal{U}_j$ (strict subset, see (38)), so that the state in slot $j+2$ is either $(\mathcal{B}_{j+1}, D_j - TR_j)$ with probability $|\mathcal{B}_{j+1}|/|\mathcal{U}_j|$ (ACK), or $(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j - TR_j)$ otherwise (NACK). This policy follows beam-alignment to data communication, and we want to prove that it is suboptimal. We use (35) to get the cost-to-go function in slot j under policy μ as

$$V_j^\mu(\mathcal{U}_j, D_j) = \phi_d(R_j, \epsilon)|\mathcal{U}_j| + V_{j+1}^\mu(\mathcal{U}_j, D_j - TR_j)$$
$$= \phi_d(R_j, \epsilon)|\mathcal{U}_j| + \phi_s|\mathcal{B}_{j+1}| + \frac{|\mathcal{B}_{j+1}|}{|\mathcal{U}_j|}V_{j+2}^\mu(\mathcal{B}_{j+1}, D_j - TR_j)$$
$$+ \frac{|\mathcal{U}_j \setminus \mathcal{B}_{j+1}|}{|\mathcal{U}_j|}V_{j+2}^\mu(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j - TR_j). \quad (44)$$

We consider the two cases $|\mathcal{B}_{j+1}| > 0$ and $|\mathcal{B}_{j+1}| = 0$ separately. In both cases, we will construct a new policy $\tilde\mu$ and compare the cost-to-go function at j under the two policies μ and $\tilde\mu$.

$|\mathcal{B}_{j+1}| > 0$: We define $\tilde\mu$ as being equal to μ except for the following: $\tilde\mu_j(\mathcal{U}_j, D_j) = (1, \mathcal{B}_{j+1}, 0)$, so that $\tilde\mu$ executes the beam-alignment action in slot j, instead of $j+1$. It follows that

$$V_j^{\tilde\mu}(\mathcal{U}_j, D_j) = \phi_s|\mathcal{B}_{j+1}| + \frac{|\mathcal{B}_{j+1}|}{|\mathcal{U}_j|}V_{j+1}^{\tilde\mu}(\mathcal{B}_{j+1}, D_j)$$
$$+ \frac{|\mathcal{U}_j \setminus \mathcal{B}_{j+1}|}{|\mathcal{U}_j|}V_{j+1}^{\tilde\mu}(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j). \quad (45)$$

Furthermore, we design $\tilde\mu$ such that $\tilde\mu_{j+1}(\mathcal{B}_{j+1}, D_j) = (0, \tilde{\mathcal{B}}'_{j+1}, R_j)$ and $\tilde\mu_{j+1}(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j) = (0, \tilde{\mathcal{B}}''_{j+1}, R_j)$, so that $\tilde\mu$ executes the data communication action in slot $j + 1$, instead of j, with beams $\tilde{\mathcal{B}}'_{j+1}$ and $\tilde{\mathcal{B}}''_{j+1}$ satisfying the conditions of Theorem 1. It follows that the system moves from state $(\mathcal{B}_{j+1}, D_j)$ to $(\mathcal{B}_{j+1}, D_j - TR_j)$, and from

[7] In fact, the AoD/AoA lie within the belief support $\mathcal{U}_k$; projecting a "2D beam" outside of $\mathcal{U}_k$ is suboptimal, since it yields an unnecessary energy cost. Additionally, choosing $\mathcal{B}_k = \mathcal{U}_k$ during beam-alignment is suboptimal, since it triggers an ACK with probability one, which is uninformative; we thus restrict $\mathcal{B}_k \subset \mathcal{U}_k$. A formal proof is provided in [36].

$(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j)$ to $(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j - TR_j)$ under policy $\tilde{\mu}$, yielding

$$V_{j+1}^{\tilde{\mu}}(\mathcal{B}_{j+1}, D_j) \overset{(a)}{=} \phi_d(R_j, \epsilon)|\mathcal{B}_{j+1}| + V_{j+2}^{\tilde{\mu}}(\mathcal{B}_{j+1}, D_j - TR_j),$$

$$V_{j+1}^{\tilde{\mu}}(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j) \overset{(b)}{=} \phi_d(R_j, \epsilon)|\mathcal{U}_j \setminus \mathcal{B}_{j+1}| + V_{j+2}^{\tilde{\mu}}(\mathcal{U}_j \setminus \mathcal{B}_{j+1}, D_j - TR_j).$$

By substituting (a),(b) into (45), and using the fact that $\tilde{\mu}_k$ and μ_k are identical for $k \geq j+2$ (hence $V_{j+2}^{\tilde{\mu}} = V_{j+2}^{\mu}$), it follows that

$$V_j^{\tilde{\mu}}(\mathcal{U}_j, D_j) - V_j^{\mu}(\mathcal{U}_j, D_j) \overset{(a)}{=} -2\phi_d(R_j, \epsilon)\frac{|\mathcal{B}_{j+1}||\mathcal{U}_j \setminus \mathcal{B}_{j+1}|}{|\mathcal{U}_j|}$$

$$\overset{(b)}{<} 0, \qquad (46)$$

where (a) follows from $|\mathcal{U}_j \setminus \mathcal{B}_{j+1}| = |\mathcal{U}_j| - |\mathcal{B}_{j+1}|$; (b) follows from $|\mathcal{B}_{j+1}| > 0$ and $\mathcal{B}_{j+1} \subset \mathcal{U}_j$.

$\underline{|\mathcal{B}_{j+1}| = 0}$: In this case, we design $\tilde{\mu}$ equal to μ except for the following: $\tilde{\mu}_j(\mathcal{U}_j, D_j) = (0, \tilde{\mathcal{B}}_j', R_j/2)$, with $\tilde{\mathcal{B}}_j'$ satisfying the conditions of Theorem 1, so that state $(\mathcal{U}_j, D_j)$ transitions to state $(\mathcal{U}_j, D_j - TR_j/2)$. Moreover $\tilde{\mu}_{j+1}(\mathcal{U}_j, D_j - TR_j/2) = (0, \tilde{\mathcal{B}}_j'', R_j/2)$, with $\tilde{\mathcal{B}}_j''$ satisfying the conditions of Theorem 1, so that the system moves to state $(\mathcal{U}_j, D_j - TR_j)$ in slot $j+2$. Under this new policy, the BS performs data communication in both slots, with rate $R_j/2$. Thus, the cost-to-go function under $\tilde{\mu}$ in slot j is given as

$$V_j^{\tilde{\mu}}(\mathcal{U}_j, D_j) = \phi_d\left(\frac{R_j}{2}, \epsilon\right)|\mathcal{U}_j| + V_{j+1}^{\hat{\mu}}\left(\mathcal{U}_j, D_j - T\frac{R_j}{2}\right)$$

$$= 2\phi_d\left(\frac{R_j}{2}, \epsilon\right)|\mathcal{U}_j| + V_{j+2}^{\hat{\mu}}(\mathcal{U}_j, D_j - TR_j). \quad (47)$$

By comparing (47) and (44) and using the fact that μ and $\tilde{\mu}$ are identical for $k \geq j+2$, we get

$$V_j^{\tilde{\mu}}(\mathcal{U}_j, D_j) - V_j^{\mu}(\mathcal{U}_j, D_j)$$

$$\overset{(a)}{=} \left[2\phi_d\left(\frac{R_j}{2}, \epsilon\right) - \phi_d(R_j, \epsilon)\right]|\mathcal{U}_j| \overset{(b)}{<} 0, \quad (48)$$

where (a) follows from $|\mathcal{B}_{j+1}| = 0$; (b) follows from the strict convexity of $\phi_d(R, \epsilon)$ over $R > 0$, implying that $2\phi_d\left(\frac{R_j}{2}, \epsilon\right) < \phi_d(R_j, \epsilon)$. (46) and (48) imply that μ does not satisfy Bellman's optimality equation, hence it is suboptimal, yielding a contradiction. The theorem is proved. ∎

From Theorem 2, we infer that:

Corollary 1: Under an optimal policy μ^*, the frame can be split into a beam-alignment phase, followed by a data communication phase until the end of the frame. The duration $L^* \in \mathcal{I}$ of beam-alignment is, possibly, a random variable, function of the realization of the beam-alignment process.

To capture this phase transition, we introduce the state variable $\nabla \in \{\text{BA}, \text{DC}\}$, denoting that the system is operating in the beam-alignment phase ($\nabla = \text{BA}$) or switched to data communication ($\nabla = \text{DC}$). The extended state is denoted as $(\mathcal{U}_k, D_k, \nabla_k)$, with the following DP updates. If $\nabla_k = \text{DC}$, then the system remains in the data communication phase until the end of the frame, and $\nabla_j = \text{DC}, \forall j \geq k$, yielding

$$\hat{V}_k^*(\mathcal{U}_k, D_k, \text{DC}) \qquad (49)$$

$$= \min_{0 < R \leq D_k/T} \left\{\phi_d(R, \epsilon)|\mathcal{U}_k| + \hat{V}_{k+1}^*(\mathcal{U}_{k+1}, D_k - TR, \text{DC})\right\}.$$

Using the convexity of $\phi_d(R, \epsilon)$ with respect to R, it is straightforward to prove the following.

Lemma 2:

$$\hat{V}_k^*(\mathcal{U}_k, D_k, \text{DC}) = (N-k)\phi_d\left(\frac{D_k}{T(N-k)}, \epsilon\right)|\mathcal{U}_k|.$$

That is, it is optimal to transmit with constant rate $\frac{D_k}{T(N-k)}$ in the remaining $(N-k)$ slots until the end of the frame. On the other hand, if $\nabla_k = \text{BA}$, then $\nabla_j = \text{BA}, \forall j \leq k$ and $D_k = D_0$, since no data has been transmitted yet. Then,

$$\hat{V}_k^*(\mathcal{U}_k, D_0, \text{BA}) = \min\left\{(N-k)\phi_d\left(\frac{NR_{\min}}{N-k}, \epsilon\right)|\mathcal{U}_k|, \right.$$

$$\min_{\mathcal{B}_k \subset \mathcal{U}_k}\left[\phi_s|\mathcal{B}_k| + \frac{|\mathcal{B}_k|}{|\mathcal{U}_k|}\hat{V}_{k+1}^*(\mathcal{B}_k, D_0, \text{BA})\right.$$

$$\left.\left. + \left(1 - \frac{|\mathcal{B}_k|}{|\mathcal{U}_k|}\right)\hat{V}_{k+1}^*(\mathcal{U}_k \setminus \mathcal{B}_k, D_0, \text{BA})\right]\right\}, \quad (50)$$

where the outer minimization reflects an optimization over the actions "switch to data communication in slot k with rate $R_k = \frac{NR_{\min}}{N-k}$," or "perform beam-alignment." The inner minimization represents an optimization over the 2D beam $\mathcal{B}_k$ used for beam-alignment.

C. Optimality of Deterministic Beam-Alignment Duration With Fractional-Search Method

It is important to observe that the proposed protocol is *interactive*, so that the duration of the beam-alignment phase, $L^* \in \mathcal{I}$, is possibly a random variable, function of the realization of the beam-alignment process. For example, if it occurs that the AoD/AoA is identified with high accuracy, the BS may decide to switch to data communication to achieve energy-efficient transmissions until the end of the frame. Although it may seem intuitive that L^* should indeed be random, in this section we will show that, instead, L^* is *deterministic*. Additionally, we prove the optimality of a *fractional search method*, which dictates the optimal beam design.

To unveil these structural properties, we define $v_k^*(\mathcal{U}_k) \triangleq \frac{\hat{V}_k^*(\mathcal{U}_k, D_0, \text{BA})}{|\mathcal{U}_k|}$. Then, (50) yields

$$v_k^*(\mathcal{U}_k) = \min\left\{(N-k)\phi_d\left(\frac{NR_{\min}}{N-k}, \epsilon\right), \right.$$

$$\left.\min_{\rho \in [0,1)} \phi_s\rho + \rho^2\, v_{k+1}^*(\mathcal{B}_{t,k}) + (1-\rho)^2\, v_{k+1}^*(\mathcal{U}_k \setminus \mathcal{B}_k)\right\}, \quad (51)$$

where $v_N^*(\mathcal{U}_N) = \infty$ and we used ρ in place of $\frac{|\mathcal{B}_k|}{|\mathcal{U}_k|}$, with $\rho < 1$ since $\mathcal{B}_k \subset \mathcal{U}_k$. Using this fact, we find that $v_{N-1}^*(\mathcal{U}_{N-1}) = \phi_d(NR_{\min}, \epsilon)$ is *independent* of $\mathcal{U}_{N-1}$. By induction on k, it is then straightforward to see that $v_k^*(\mathcal{U}_k)$ is *independent* of $\mathcal{U}_k, \forall k$. We thus let $v_k^* \triangleq v_k^*(\mathcal{U}_k), \forall \mathcal{U}_k$ to capture this independence, which is then defined recursively as

$$v_k^* = \min\left\{(N-k)\phi_d\left(\frac{NR_{\min}}{N-k}, \epsilon\right), \right.$$

$$\left.\min_{\rho \in [0,1)} \phi_s\rho + \left[\rho^2 + (1-\rho)^2\right]v_{k+1}^*\right\}. \quad (52)$$

The value of ρ achieving the minimum in (52) is $\rho_k = \frac{|\mathcal{B}_k|}{|\mathcal{U}_k|} = \frac{1}{2}\left(1 - \frac{\phi_s}{2v_{k+1}^*}\right)^+$, yielding

$$v_k^* = \min\left\{\underbrace{(N-k)\phi_d\left(\frac{NR_{\min}}{N-k}, \epsilon\right)}_{\Gamma_k \text{ (data communication)}}, \underbrace{v_{k+1}^* - \frac{[(2v_{k+1}^* - \phi_s)^+]^2}{8v_{k+1}^*}}_{\Lambda_k \text{ (beam-alignment)}}\right\}.$$

From this decomposition, we infer important properties:

1) Given v_k^*, the original value function is obtained as $\hat{V}_k^*(\mathcal{U}_k, D_0, \text{BA}) = v_k^*|\mathcal{U}_k|$. If, at time k, $\Gamma_k < \Lambda_k$, then it is optimal to switch to data communication in the remaining $N - k$ slots, with constant rate $\frac{NR_{\min}}{N-k}$.
2) Otherwise, it is optimal to perform beam-alignment, with beam $\mathcal{B}_k \subset \mathcal{U}_k$, $|\mathcal{B}_k| = \rho_k|\mathcal{U}_k|$.
3) Finally, since the time to switch to data communication is solely based on $\{v_k^*\}$, but not on $\mathcal{U}_k$, it follows that *fixed-length* beam-alignment is optimal, with duration

$$L^* = \min\{k : \Gamma_k < \Lambda_k\}. \tag{53}$$

These structural results are detailed in the following theorem.

Theorem 3: *Let*

$$L_{\min} = \underset{L \in \{0, \ldots, N-1\}}{\arg\min}\left\{L : (N-L)\phi_d\left(\frac{NR_{\min}}{N-L}, \epsilon\right) > \frac{\phi_s}{2}\right\} \tag{54}$$

and, for $L_{\min} \leq L < N$,

$$\begin{cases} v_L^{(L)} = (N-L)\phi_d\left(\frac{NR_{\min}}{N-L}, \epsilon\right), \\ v_k^{(L)} = v_{k+1}^{(L)} - \frac{(2v_{k+1}^{(L)} - \phi_s)^2}{8v_{k+1}^{(L)}}, \quad k < L. \end{cases} \tag{55}$$

Then, the beam-alignment phase has deterministic *duration*

$$L^* = \arg\min_{L \in \{0\} \cup \{L_{\min}, \ldots, N-1\}} v_0^{(L)}. \tag{56}$$

For $0 \leq k < L^*$ (beam-alignment phase), $\mathcal{B}_k$ is optimal iff

$$\mathcal{B}_k \subset \mathcal{U}_k, \quad |\mathcal{B}_k| = \rho_k|\mathcal{U}_k|, \tag{57}$$

where ρ_k is the fractional search *parameter, defined as*

$$\begin{cases} \rho_{L^*-1} = \frac{1}{2} - \frac{\phi_s}{4(N-L^*)\phi_d\left(\frac{NR_{\min}}{N-L^*}, \epsilon\right)}, \\ \rho_k = \frac{1 - \rho_{k+1}}{1 - 2\rho_{k+1}^2}\rho_{k+1}, \quad k < L^* - 1. \end{cases} \tag{58}$$

Moreover, $\rho_k \in (0, 1/2)$, strictly increasing in k. For $k \geq L^*$, the data communication phase occurs with rate $\frac{NR_{\min}}{N-L^*}$, and 2D beam given by Theorem 1.

Proof: Since the optimal duration of the beam-alignment phase is deterministic, as previously discussed, we consider a fixed beam-alignment duration L, and then optimize over L to achieve minimum energy consumption. Let $L \in \mathcal{I}$. Then, the DP updates are obtained by adapting (52) to this case (so

that the outer minimization disappears for $k < L$), yielding

$$\begin{cases} v_L^{(L)} = (N-L)\phi_d\left(\frac{NR_{\min}}{N-L}, \epsilon\right), \\ v_k^{(L)} = g_k(\rho_k), \quad k < L, \text{ where} \\ g_k(\rho) \triangleq \phi_s\rho + \left[\rho^2 + (1-\rho)^2\right]v_{k+1}^{(L)}, \\ \rho_k = \arg\min_{\rho \in [0,1]} g_k(\rho) = \frac{1}{2}\left(1 - \frac{\phi_s}{2v_{k+1}^{(L)}}\right)^+. \end{cases} \tag{59}$$

Since the goal is to minimize the energy consumption, the optimal L is obtained as $L^* = \arg\min_L v_0^{(L)}$. We now prove that $0 < L < L_{\min}$ is suboptimal, so that this optimization can be restricted to $L \in \{0\} \cup \{L_{\min}, \cdots, N-1\}$, as in (56). Let $0 < L < L_{\min}$, so that $v_L^{(L)} \leq \phi_s/2$, as can be seen from the definition of $L_{\min}$ in (54). Note that $v_k^{(L)}$ is a non-decreasing function of k. In fact, $v_k^{(L)} \leq g_k(0) = v_{k+1}^{(L)}$. Then, it follows that $v_k^{(L)} \leq \phi_s/2, \forall k$, hence $\rho_k = 0, \forall k$, yielding $v_0^{(L)} = v_L^{(L)}$ by induction. However, $v_L^{(L)}$ is an increasing function of L (it is more energy efficient to spread transmissions over a longer interval), hence $v_0^{(L)} > v_0^{(0)}$ and such L is suboptimal. This proves that any $0 < L < L_{\min}$ is suboptimal.

We now prove the updates for $L \geq L_{\min}$, i.e., $v_L^{(L)} > \phi_s/2$. By induction, we have that $v_k^{(L)} > \phi_s/2, \forall k$. In fact, this condition trivially holds for $k = L$, by hypothesis. Now, assume $v_{k+1}^{(L)} > \phi_s/2$ for some $k < L$. Then, $v_k^{(L)} = \min_{\rho \in [0,1]} g_k(\rho)$, minimized at $\rho_k = \frac{1}{2}\left(1 - \frac{\phi_s}{2v_{k+1}^{(L)}}\right)$, so that $v_k^{(L)} = g_k(\rho_k)$, yielding (55). This recursion is an increasing function of $v_{k+1}^{(L)}$, yielding $v_k^{(L)} > \phi_s/2$, thus proving the induction. It follows that $\rho_k = \frac{1}{2}\left(1 - \frac{\phi_s}{2v_{k+1}^{(L)}}\right), \forall k$, yielding the recursion given by (55). The fractional search parameter ρ_k is finally obtained by substituting $v_{k+1}^{(L)} = \frac{\phi_s}{2(1-2\rho_k)}$ into the recursion (55) to find a recursive expression of ρ_k from ρ_{k+1}, yielding (58). These fractional values are used to obtain $\mathcal{B}_k$ in (57).

To conclude, we show by induction that $\rho_k \in (0, 1/2)$, strictly increasing in k. This is true for $k = L-1$ since $\rho_{L-1} \in (0, 1/2)$. Assume that $\rho_{k+1} \in (0, 1/2)$, for some $k \leq L-2$. Then, by inspection of (58), it follows that $0 < \rho_k < \rho_{k+1} < 1/2$. The theorem is thus proved. ∎

V. DECOUPLED BS AND UE BEAM-ALIGNMENT

In the previous section, we proved the optimality of a fractional search method, based on an extended action space that uses the 2D beam $\mathcal{B}_k \in [-\pi, \pi]^2$, which may take any shape. However, actual beams should satisfy the rectangular constraint $\mathcal{B}_k = \mathcal{B}_{t,k} \times \mathcal{B}_{r,k}$, and therefore, it is not immediate to see that the optimal scheme outlined in Theorem 3 is attainable in practice. Indeed, in this section we prove that there exists a feasible beam design attaining optimality. The proposed beam design decouples over time the beam-alignment of the AoD at the BS (*BS beam-alignment*) and of the AoA at the UE (*UE beam-alignment*). To explain this approach, we define the support of the marginal belief with respect to $\theta_x, x \in \{t, r\}$ as $\mathcal{U}_{x,k} \equiv \text{supp}(f_{x,k})$. In *BS beam-alignment*, indicated with

$\beta_k = 1$, the 2D beam is chosen as $\mathcal{B}_k = \mathcal{B}_{\mathrm{t},k} \times \mathcal{U}_{\mathrm{r},k}$, where $\mathcal{B}_{\mathrm{t},k} \subset \mathcal{U}_{\mathrm{t},k}$, so that the BS can better estimate the support of the AoD, whereas the UE receives over the entire support of the AoA. On the other hand, in *UE beam-alignment*, indicated with $\beta_k = 2$, the 2D beam is chosen as $\mathcal{B}_k = \mathcal{U}_{\mathrm{t},k} \times \mathcal{B}_{\mathrm{r},k}$, where $\mathcal{B}_{\mathrm{r},k} \subset \mathcal{U}_{\mathrm{r},k}$, so that the UE can better estimate the support of the AoA, whereas the BS transmits over the entire support of the AoD. We now define a policy μ that uses this principle, and then prove its optimality.

Definition 1 (Decoupled Fractional Search Policy): Let $L^*, \vartheta, \{\rho_k : k = 0, \dots, L^*-1\}$ as in Theorems 1, 3. In slots $k = L^*, \dots, N-1$, data communication occurs with rate $R_k = \frac{N R_{\min}}{N - L^*}$ and beams

$$\mathcal{B}_{\mathrm{t},k} \subseteq \mathcal{U}_{\mathrm{t},k}, \ \mathcal{B}_{\mathrm{r},k} \subseteq \mathcal{U}_{\mathrm{r},k}, \ |\mathcal{B}_{\mathrm{t},k}||\mathcal{B}_{\mathrm{r},k}| = \vartheta |\mathcal{U}_{\mathrm{t},k}||\mathcal{U}_{\mathrm{r},k}|. \quad (60)$$

In slots $k = 0, 1, \dots, L^*-1$, $\beta_k \in \{1, 2\}$ is chosen arbitrarily and beam-alignment occurs with beams

$$\begin{cases} \mathcal{B}_{\mathrm{t},k} \subset \mathcal{U}_{\mathrm{t},k}, \ \mathcal{B}_{\mathrm{r},k} = \mathcal{U}_{\mathrm{r},k}, \ |\mathcal{B}_{\mathrm{t},k}| = \rho_k |\mathcal{U}_{\mathrm{t},k}|, & \text{if } \beta_k = 1 \\ \mathcal{B}_{\mathrm{t},k} = \mathcal{U}_{\mathrm{t},k}, \mathcal{B}_{\mathrm{r},k} \subset \mathcal{U}_{\mathrm{r},k}, \ |\mathcal{B}_{\mathrm{r},k}| = \rho_k |\mathcal{U}_{\mathrm{r},k}|, & \text{if } \beta_k = 2. \end{cases}$$

Theorem 4: *The decoupled fractional search policy is optimal, with minimum power consumption*

$$\bar{P}_{\mathrm{u}} = \frac{v_0^{(L^*)}}{T_{\mathrm{fr}}} |\mathcal{U}_0|. \quad (61)$$

Proof: The proof is provided in Appendix B. ∎

The intuition behind this result is that, by decoupling the beam-alignment of the AoD and AoA over time, the proposed method maintains a rectangular support $\mathcal{U}_k = \mathcal{U}_{\mathrm{t},k} \times \mathcal{U}_{\mathrm{r},k}$, so that no loss of optimality is incurred by using a rectangular beam $\mathcal{B}_k = \mathcal{B}_{\mathrm{t},k} \times \mathcal{B}_{\mathrm{r},k}$. Additionally, we can infer that the *exhaustive search* method is suboptimal, since it searches over the AoD/AoA space in an exhaustive manner, rather than by decoupling this search over time.

VI. NON-UNIFORM PRIOR

In this section, we investigate the case of non-uniform prior f_0. We use the previous analysis to design a heuristic scheme with performance guarantees. We consider the decoupled fractional search policy (Definition 1), with the following additional constraints: in the beam-alignment phase $k < L^*$, if $\beta_k^* = 1$ (BS beam-alignment), then

$$\mathcal{B}_{\mathrm{t},k}^* = \arg \max_{\mathcal{B}_{\mathrm{t},k} \subset \mathcal{U}_{\mathrm{t},k}} \int_{\mathcal{B}_{\mathrm{t},k}} f_{\mathrm{t},k}(\theta_{\mathrm{t}}) \mathrm{d}\theta_{\mathrm{t}}, \ \text{s.t. } |\mathcal{B}_{\mathrm{t},k}| = \rho_k |\mathcal{U}_{\mathrm{t},k}|; \quad (62)$$

if $\beta_k^* = 2$ (UE beam-alignment), then

$$\mathcal{B}_{\mathrm{r},k}^* = \arg \max_{\mathcal{B}_{\mathrm{r},k} \subset \mathcal{U}_{\mathrm{r},k}} \int_{\mathcal{B}_{\mathrm{r},k}} f_{\mathrm{r},k}(\theta_{\mathrm{r}}) \mathrm{d}\theta_{\mathrm{r}}, \ \text{s.t. } |\mathcal{B}_{\mathrm{r},k}| = \rho_k |\mathcal{U}_{\mathrm{r},k}|. \quad (63)$$

Hence, the probability of ACK can be bounded as

$$\left. \begin{array}{l} \text{Case } \beta_k^* = 1: \int_{\mathcal{B}_{\mathrm{t},k}^*} f_{\mathrm{t},k}(\theta_{\mathrm{t}}) \mathrm{d}\theta_{\mathrm{t}} \geq \frac{|\mathcal{B}_{\mathrm{t},k}^*|}{|\mathcal{U}_{\mathrm{t},k}|} \\ \text{Case } \beta_k^* = 2: \int_{\mathcal{B}_{\mathrm{r},k}^*} f_{\mathrm{r},k}(\theta_{\mathrm{r}}) \mathrm{d}\theta_{\mathrm{r}} \geq \frac{|\mathcal{B}_{\mathrm{r},k}^*|}{|\mathcal{U}_{\mathrm{r},k}|} \end{array} \right\} = \rho_k. \quad (64)$$

In other words, such choice of the BS-UE beam maximizes the probability of successful beam-detection, so that the resulting probability of ACK is at least as good as in the uniform case.

Similarly, in the data communication phase $k \geq L^*$, the BS transmits with rate $R_k = \frac{N R_{\min}}{N - L^*}$, and the beams are chosen as in Definition 1, with the additional constraint

$$(\mathcal{B}_{\mathrm{t},k}^*, \mathcal{B}_{\mathrm{r},k}^*) = \arg \max_{\mathcal{B}_{\mathrm{t},k} \times \mathcal{B}_{\mathrm{r},k} \subseteq \mathcal{U}_k} \int_{\mathcal{B}_{\mathrm{t},k} \times \mathcal{B}_{\mathrm{r},k}} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta},$$
$$\text{s.t. } |\mathcal{B}_{\mathrm{t},k}||\mathcal{B}_{\mathrm{t},k}| = \vartheta |\mathcal{U}_{\mathrm{t},k}||\mathcal{U}_{\mathrm{r},k}|. \quad (65)$$

Under this choice, the energy consumption per data communication slot is obtained from (34),

$$E_k = \psi_d(R_k) \frac{|\mathcal{B}_k|}{\bar{F}_\gamma^{-1} \left(\frac{1-\epsilon}{\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{U}_k)} \right)} \quad (66)$$

$$\overset{(a)}{\leq} \psi_d(R_k) \frac{|\mathcal{B}_k|}{\bar{F}_\gamma^{-1} \left(\frac{(1-\epsilon)|\mathcal{U}_k|}{|\mathcal{B}_k|} \right)} \overset{(b)}{=} \phi_d(R_k, \epsilon) |\mathcal{U}_k|, \quad (67)$$

where (a) follows from $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{U}_k) \geq |\mathcal{B}_k|/|\mathcal{U}_k|$, and (b) from $|\mathcal{B}_{\mathrm{t},k}||\mathcal{B}_{\mathrm{r},k}| = \vartheta |\mathcal{U}_{\mathrm{t},k}||\mathcal{U}_{\mathrm{r},k}|$, and from (42) with $\vartheta = (1-\epsilon)/q^*$ (Theorem 1). This result implies that data communication is more energy efficient than in the uniform case, see (43). These observations suggest that the uniform prior yields the worst performance, as confirmed by the following theorem.

Theorem 5: *The minimum power consumption for the non-uniform prior is upper bounded by $\bar{P}_{\mathrm{nu}} \leq \bar{P}_{\mathrm{u}}$, with equality when f_0 is uniform.*

Proof: We denote the value function of the non-uniform case under such policy as $V_{\mathrm{nu},k}(\mathcal{U}_k, D_k)$. Additionally, we let $\bar{P}_{\mathrm{nu}}$ be the corresponding minimum power consumption, solution of problem P_2 in (33), to distinguish it from the minimum power consumption in the uniform case, given by (61). For $k = L^*$ (data communication begins), (66) implies that

$$V_{\mathrm{nu},k}(\mathcal{U}_k, D_0) \leq (N - L^*)\phi_d \left(\frac{N R_{\min}}{N - L^*}, \epsilon \right) |\mathcal{U}_k|. \quad (68)$$

For $k < L^*$ (beam-alignment phase), it can be expressed as

$$V_{\mathrm{nu},k}(\mathcal{U}_k, D_0) = \phi_s |\mathcal{B}_k^*| + \int_{\mathcal{B}_k^*} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} V_{\mathrm{nu},k+1}(\mathcal{B}_k^*, D_0)$$
$$+ \left(1 - \int_{\mathcal{B}_k^*} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \right) V_{\mathrm{nu},k+1}(\mathcal{U}_k \setminus \mathcal{B}_k^*, D_0), \quad (69)$$

where $\mathcal{B}_k^*$ is given by (62) or (63). The minimum power consumption is given by $\bar{P}_{\mathrm{nu}} = V_{\mathrm{nu},0}(\mathcal{U}_0, D_0)/T_{\mathrm{fr}}$, so that $\bar{P}_{\mathrm{nu}} \leq \bar{P}_{\mathrm{u}}$ is equivalent to $V_{\mathrm{nu},k}(\mathcal{U}_k, D_k) \leq v_k^{(L^*)} |\mathcal{U}_k|$ when $k = 0$. We prove this inequality for general k by induction. The induction hypothesis holds for $k = L^*$, see (68) with $v_{L^*}^{(L^*)}$ given in (55). Assume it holds for $k + 1$, where $k \leq L^* - 1$. Then, (69) can be expressed as

$$V_{\mathrm{nu},k}(\mathcal{U}_k, D_0) \leq \phi_s |\mathcal{B}_k^*| + \int_{\mathcal{B}_k^*} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} v_{k+1}^{(L^*)} |\mathcal{B}_k^*|$$
$$+ \left(1 - \int_{\mathcal{B}_k^*} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \right) v_{k+1}^{(L^*)} |\mathcal{U}_k \setminus \mathcal{B}_k^*|$$
$$\overset{(a)}{=} \left[\phi_s \rho_k + v_{k+1}^{(L^*)} \left(1 - 2\rho_k + 2\rho_k^2 \right) \right] |\mathcal{U}_k|$$
$$- \left(\int_{\mathcal{B}_k^*} f_k(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} - \rho_k \right) v_{k+1}^{(L^*)} |\mathcal{U}_k| (1 - 2\rho_k),$$

where (a) follows from (62)-(63) and $|\mathcal{U}_k \setminus \mathcal{B}_k^*| = |\mathcal{U}_k| - |\mathcal{B}_k^*|$. Finally, the bound (64) yields

$$V_{\mathrm{nu},k}(\mathcal{U}_k, D_0) \leq \left[\phi_s \rho_k + v_{k+1}^{(L^*)} \left(1 - 2\rho_k + 2\rho_k^2 \right) \right] |\mathcal{U}_k|$$
$$= v_k^{(L^*)} |\mathcal{U}_k|,$$

where the last equality is obtained by using the recursion (55) and the fact that $\rho_k = \frac{1}{2} - \frac{\phi_s}{4v_{k+1}^{(L^*)}}$ (see proof of Theorem 3). This proves the induction step. Clearly, equality is attained in the uniform case. The theorem is thus proved. ∎

This result is in line with the fact that one can leverage the structure of the joint distribution over $\boldsymbol{\theta}$ to improve the beam-alignment algorithm. However, for the first time to the best of our knowledge, this result provides a heuristic scheme with provable performance guarantees.

VII. IMPACT OF FALSE-ALARM AND MISDETECTION

In this section, we analyze the impact of false-alarm and misdetection on the performance of the decoupled fractional search policy (Definition 1). For simplicity, we focus only on the uniform prior case. Under false-alarm and misdetection, the MDP introduced in Sec. III does not follow the Markov property. To overcome this problem, we augment it with the state variable $e_k \in \{0,1\}$, with $e_k = 0$ iff no errors have been introduced up to slot k. Note that, if errors have been introduced ($e_k = 1$), then necessarily $\boldsymbol{\theta} \notin \mathcal{U}_k$, so that we can write $e_k = 1 - \chi(\boldsymbol{\theta} \in \mathcal{U}_k)$. It should be noted that e_k is not observable in reality and is considered for the purpose of analysis only (indeed, the policy under analysis does not use such information). We thus define the state as $(\mathcal{U}_k, e_k)$,[8] and study the transition probabilities during the beam-alignment phase $k < L^*$. From state $(\mathcal{U}_k, 0)$ (no errors have been introduced), the transitions are

$$(\mathcal{U}_{k+1}, e_{k+1}) = \begin{cases} (\mathcal{B}_k, 0), & \text{w.p. } \rho_k(1 - p_{\mathrm{md}}) \\ (\mathcal{B}_k, 1), & \text{w.p. } (1 - \rho_k) p_{\mathrm{fa}} \\ (\mathcal{U}_k \setminus \mathcal{B}_k, 0), & \text{w.p. } (1 - \rho_k)(1 - p_{\mathrm{fa}}) \\ (\mathcal{U}_k \setminus \mathcal{B}_k, 1), & \text{w.p. } \rho_k \, p_{\mathrm{md}}, \end{cases} \tag{70}$$

where p_{fa} and p_{md} denote the false-alarm and misdetection probabilities, respectively. In fact, if no errors occur, then $\boldsymbol{\theta} \in \mathcal{B}_k$ with probability $\frac{|\mathcal{B}_k|}{|\mathcal{U}_k|} = \rho_k$ and $\boldsymbol{\theta} \notin \mathcal{B}_k$ otherwise, yielding the first and third cases; if a false-alarm or misdetection error is introduced, then the BS infers incorrectly that $\boldsymbol{\theta} \in \mathcal{B}_k$ (second case) or $\boldsymbol{\theta} \notin \mathcal{B}_k$ (fourth case), respectively, and the new state becomes $e_{k+1} = 1$. Once errors have been introduced (state $(\mathcal{U}_k, 1)$), it follows that $\boldsymbol{\theta} \notin \mathcal{B}_k$, so that $\mathcal{U}_{k+1} = \mathcal{B}_k$ iff a false-alarm error occurs, and the transitions are

$$(\mathcal{U}_{k+1}, e_{k+1}) = \begin{cases} (\mathcal{B}_k, 1), & \text{w.p. } p_{\mathrm{fa}} \\ (\mathcal{U}_k \setminus \mathcal{B}_k, 1), & \text{w.p. } 1 - p_{\mathrm{fa}}. \end{cases} \tag{71}$$

The average throughput and power are given by

$$\bar{T}_{\mathrm{err}} = \mathbb{E}\left[(1 - e_{L^*})(1 - \epsilon) R_{\min} | \mathcal{U}_0, e_0 = 0 \right], \tag{72}$$

---

[8] The backlog D_k is removed from the state space, since no data is transmitted during the beam-alignment phase.

$$\bar{P}_{\mathrm{err}} = \frac{1}{T_{\mathrm{fr}}} \mathbb{E}\left[\phi_s \sum_{k=0}^{L^*-1} \rho_k |\mathcal{U}_k| \right.$$
$$\left. + (N - L^*)\phi_d \left(\frac{NR_{\min}}{N - L^*}, \epsilon \right) |\mathcal{U}_{L^*}| \right| \mathcal{U}_0, e_0 = 0 \right]. \tag{73}$$

In fact, a rate equal to $R_{\min}$ is sustained if: (1) no outage occurs in the data communication phase, with probability $1 - \epsilon$; (2) no errors occur during the beam-alignment phase, $e_{L^*} = 0$.

The analysis of the underlying Markov chain $\{(\mathcal{U}_k, e_k), k \geq 0\}$ yields the following theorem.

Theorem 6: *Under the decoupled fractional search policy,*

$$\bar{T}_{\mathrm{err}} = (1 - \epsilon) R_{\min} \prod_{k=0}^{L^*-1} \left[(1 - \rho_k)(1 - p_{\mathrm{fa}}) + \rho_k(1 - p_{\mathrm{md}}) \right], \tag{74}$$

$$\bar{P}_{\mathrm{err}} = \bar{P}_{\mathrm{u}} + \frac{h_0 + u_0}{T_{\mathrm{fr}}} |\mathcal{U}_0|, \tag{75}$$

where $\bar{P}_{\mathrm{u}}$ in (61) is the error-free case, and we have defined $h_{L^*} = u_{L^*} = 0$ and, for $k < L^*$,

$$h_k = \phi_s \frac{\rho_k - p_{\mathrm{fa}}}{2} + \left[\rho_k p_{\mathrm{fa}} + (1 - \rho_k)(1 - p_{\mathrm{fa}}) \right] h_{k+1}, \tag{76}$$

$$u_k = \left[\rho_k^2 (1 - p_{\mathrm{md}}) + (1 - \rho_k)^2 (1 - p_{\mathrm{fa}}) \right] u_{k+1}$$
$$- (1 - p_{\mathrm{fa}} - p_{\mathrm{md}})\rho_k \left[\frac{\phi_s}{2} + h_{k+1}(1 - 2\rho_k) \right]. \tag{77}$$

Proof: The proof is provided in Appendix C. ∎

VIII. NUMERICAL RESULTS

In this section, we demonstrate the performance of the proposed *decoupled fractional search* (DFS) scheme and compare it with the *bisection search* algorithm developed in [9] and two variants of *exhaustive search*. In the bisection algorithm [9] (BiS), in each beam-alignment slot the uncertainty region is divided into two regions of equal width, scanned in sequence by the BS by transmitting beacons corresponding to each region. Then, the UE compares the signal power (the strongest indicating alignment) and transmits the feedback to the BS. Since in each beam-alignment slot two sectors are scanned (each of duration T_B), the total duration of the beam-alignment phase is $(2 T_B + T_F)L$ [s], where T_F is the feedback time.

In *conventional exhaustive search* (CES), the BS-UE scan exhaustively the entire beam space. In the BS beam-alignment sub-phase, the BS searches over $N_B^{(BS)}$ beams covering the AoD space, while the UE receives isotropically; in the second UE beam-alignment sub-phase, the BS transmits using the best beam found in the first sub-phase, whereas the UE searches exhaustively over $N_B^{(UE)}$ beams covering the AoA space. Since the UE reports the best beam at the end of each sub-phase, the total duration of the beam-alignment phase is $[N_B^{(BS)} + N_B^{(UE)}]T_B + 2T_F$. On the other hand, in the *interactive exhaustive search* (IES) method, the UE reports the feedback at the end of each beam-alignment slot, and each beam-alignment sub-phase terminates upon receiving an ACK from the UE. Since the BS awaits for feedback at the
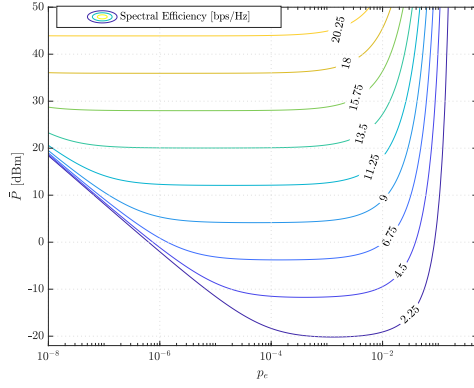
Fig. 2. Spectral efficiency versus beam-alignment error probability p_e for DFS.

end of each beam, the duration of the beam-alignment phase is $(T_B + T_F)[\hat{N}_B^{(BS)} + \hat{N}_B^{(UE)}]$, where $\hat{N}_B \leq N_B$ is the number of beams scanned until receiving an ACK; assuming the AoD/AoA is uniformly distributed over the beam space, the expected duration of the beam-alignment phase is then $\frac{1}{2}(T_B + T_F)\left[N_B^{(BS)} + N_B^{(UE)} + 2\right]$.

We use the following parameters: [carrier frequency]= 30GHz, $d = 10$m, [path loss exponent]= 2, $T_{\mathrm{fr}} = 20$ms, $T_B = 50\mu$s, $T_F = 50\mu$s, $|\mathcal{U}_0| = [\pi]^2$, $N_0 = -173$dBm, $W_{\mathrm{tot}} = 500$MHz, $M_{\mathrm{t}} = M_{\mathrm{r}} = 128$.

In Fig. 2, we depict the average power vs the probability of false-alarm and misdetection p_e for different values of the spectral efficiency using expressions (74) and (75). We use $\epsilon = 0.01$, and consider Rayleigh fading with no CSI at BS, corresponding to $h \sim \mathcal{CN}(0, 1/\ell(d))$ with $\hat{h} = 0$ and $\sigma_e^2 = 1/\ell(d)$. We restrict the optimization of L over $L \in \{0, \ldots, L_{\max}\}$, to capture a maximum resolution constraint for the antenna array, where we chose $L_{\max} = 14$. From the figure, we observe that, for a given p_e, as the spectral efficiency increases so does the average power consumption due to increase in the energy cost of data communication. Moreover, the figure reveals that, for a given value of spectral efficiency, there exists an optimal range of p_e, where power consumption is minimized. The performance degrades for p_e above the optimal range due to false-alarm and misdetection errors during beam-alignment, causing outage in data communication; similarly, it degrades for p_e below the optimal range due to an increased power consumption of beam-alignment.

In Fig. 3, we plot the results of a Monte-Carlo simulation with analog beams generated using the algorithm in [22]. In this case, we obtain $\phi_s = -94$dBm with $p_{\mathrm{fa}} = p_{\mathrm{md}} = 10^{-5}$. For BiS and DFS we set $L_{\max} = 10$ to capture a maximum resolution constraint for the antenna array; for the exhaustive search methods, we choose $N_B^{(BS)} = N_B^{(UE)} = 32$. The performance gap between the analytical and the simulation-based curves for DFS is attributed to the fact that the beams used in the simulation have non-zero side-lobe gain and non-uniform main-lobe gain, as opposed to the "sectored" beams used in the analytical model. This results in false-alarm, misdetection errors, and leakage, which lead to some performance degradation. *However, the simulation is in line with the analytical curve, and exhibits*
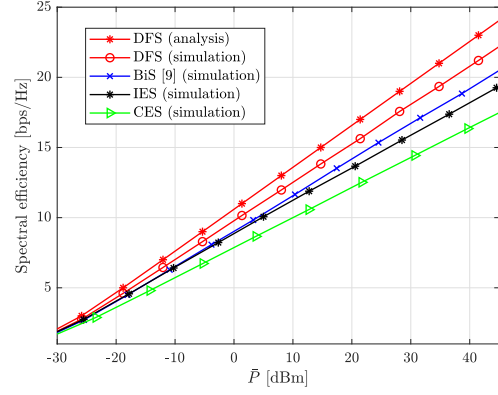

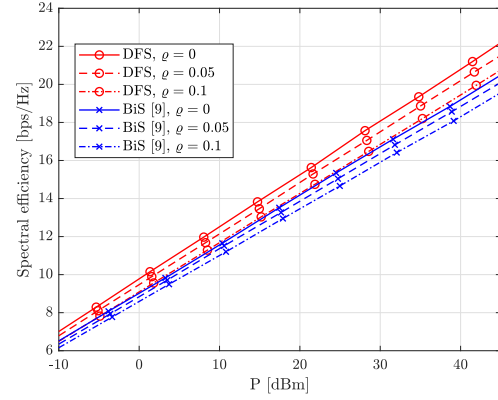Fig. 3. Spectral efficiency versus average power consumption.

Fig. 4. Performance degradation with multi-cluster channel ($K = 2$).

superior performance compared to the other schemes, thus demonstrating that the analysis using the sectored gain model provides useful insights for practical design.

For instance, to achieve a spectral efficiency of 15bps/Hz, BiS [9] requires 4dB more average power than DFS, mainly due to the time and energy overhead of scanning two sectors in each beam-alignment slot, whereas IES and CES require 7.5dB and 14dB more power, respectively. The performance degradation of IES and CES is due to the exhaustive search of the best sector, which demands a huge time overhead. Indeed, IES outperforms CES since it stops beam-alignment once a strong beam is detected.

So far in our analysis, we assumed a channel with a single cluster of rays, see (3). In Fig. 4, we depict the performance of DFS and BiS [9] in a multi-cluster channel ($K = 2$ in (2)), with the weakest cluster having a fraction ϱ of the total energy, $0 \leq \varrho \leq 0.1$. It can be seen that the performance of both DFS and BiS degrade as ϱ increases, since a portion of the energy is lost in the weaker cluster, and the algorithms may misdetect the weaker cluster instead of the strongest one. For example, for spectral efficiency of 15bps/Hz, both schemes exhibit ~ 2dB and ~ 5dB performance loss at $\varrho = 5\%$ and $\varrho = 10\%$, respectively, compared to $\varrho = 0$ (single cluster). However, DFS consistently outperforms BiS, with a gain of ~ 3.5dB. This evaluation demonstrates the robustness of the proposed algorithm in multi-cluster scenarios.

IX. CONCLUSIONS

In this paper, we designed an optimal interactive beam-alignment scheme, with the goal of minimizing power

consumption under a rate constraint. For the case of perfect detection and uniform prior on AoD/AoA, we proved that the optimal beam-alignment protocol has fixed beam-alignment duration, and that a *decoupled fractional search* method is optimal. Inspired by this scheme, we proposed a heuristic policy for the case of a non-uniform prior, and showed that the uniform prior is the worst-case scenario. Furthermore, we investigated the impact of beam-alignment errors on the average throughput and power consumption. The numerical results depicted the superior performance of our proposed scheme, with up to 4dB, 7.5dB, and 14dB gain compared to a state-of-the-art bisection search, conventional exhaustive search and interactive exhaustive search policies, respectively, and robustness against multi-cluster channels.

APPENDIX A: PROOF OF THEOREM 1

Proof: For a data communication action $\mathbf{a}_k \in \mathcal{A}_{\text{ext}}(\mathcal{U}, D)$, the state transition is independent of $\mathcal{B}_k$ since $\mathcal{U}_{k+1} = \mathcal{U}_k$ and $D_{k+1} = D_k - R_k T$. Hence, the optimal beam given R_k is obtained by minimizing $c(\mathbf{a}_k; \mathcal{U}_k, D_k)$ in (34), yielding

$$c(\mathbf{a}_k; \mathcal{U}_k, D_k) \overset{(a)}{=} \psi_d(R_k) \frac{|\mathcal{B}_k|}{\bar{F}_\gamma^{-1}\left(\frac{(1-\epsilon)|\mathcal{U}_k|}{|\mathcal{B}_k|}|\hat{\gamma}\right)}$$
$$\overset{(b)}{\geq} \psi_d(R_k)(1-\epsilon)|\mathcal{U}_k| \frac{1}{q^* \bar{F}_\gamma^{-1}(q^*|\hat{\gamma})}, \quad (78)$$

where (a) follows from $\mathbb{P}(\boldsymbol{\theta} \in \mathcal{B}_k | \mathcal{U}_k, \mathbf{a}_k) = \frac{|\mathcal{B}_k|}{|\mathcal{U}_k|}$, with $q \triangleq (1-\epsilon)\frac{|\mathcal{U}_k|}{|\mathcal{B}_k|} \leq 1$ to enforce the ϵ-outage constraint; (b) follows by maximizing $q\bar{F}_\gamma^{-1}(q|\hat{\gamma})$ over $q \in [1-\epsilon, 1]$. Equality holds in (b) if $|\mathcal{B}_k| = \vartheta |\mathcal{U}_k|$, with $\vartheta = (1-\epsilon)/q^*$ and q^* as in the statement. The theorem is thus proved. ∎

APPENDIX B: PROOF OF THEOREM 4

Proof: Note that, if this policy satisfies $\mathcal{B}_k \equiv \mathcal{B}_{\text{t},k} \times \mathcal{B}_{\text{r},k} \subseteq \mathcal{U}_k \equiv \text{supp}(f_k)$, along with the appropriate fractional values $|\mathcal{B}_k|/|\mathcal{U}_k|$, then it is optimal since it satisfies all the conditions of Theorems 1 and 3. We now verify these conditions. Since $\mathcal{B}_{\text{t},k} \subseteq \mathcal{U}_{\text{t},k}$ and $\mathcal{B}_{\text{r},k} \subseteq \mathcal{U}_{\text{r},k}$, it is sufficient to prove that $\mathcal{U}_{\text{t},k} \times \mathcal{U}_{\text{r},k} \equiv \mathcal{U}_k, \forall k$. Indeed, $\mathcal{U}_0 \equiv \mathcal{U}_{\text{t},0} \times \mathcal{U}_{\text{r},0}$. By induction, assume that $\mathcal{U}_k \equiv \mathcal{U}_{\text{t},k} \times \mathcal{U}_{\text{r},k}$. Then, for $\beta_k = 1$ (a similar result holds for $\beta_k = 2$), using (30) we get

$$\mathcal{U}_{k+1} = \begin{cases} (\mathcal{U}_{\text{t},k} \cap \mathcal{B}_{t,k}) \times \mathcal{U}_{r,k}, & \text{if } C_k = \text{ACK}, \\ (\mathcal{U}_{\text{t},k} \setminus \mathcal{B}_{t,k}) \times \mathcal{U}_{r,k}, & \text{if } C_k = \text{NACK}. \end{cases} \quad (79)$$

By letting $\mathcal{U}_{\text{r},k} \equiv \mathcal{U}_{\text{r},k-1}$, $\mathcal{U}_{\text{t},k} \equiv \mathcal{U}_{\text{t},k-1} \cap \mathcal{B}_{t,k-1}$ if $C_k = \text{ACK}$ and $\mathcal{U}_{\text{t},k} \equiv \mathcal{U}_{\text{t},k-1} \setminus \mathcal{B}_{t,k-1}$ if $C_k = \text{NACK}$, we obtain $\mathcal{U}_k \equiv \mathcal{U}_{\text{t},k} \times \mathcal{U}_{\text{r},k}$. This policy is then optimal. Finally, (61) is obtained by using the relation between power consumption and value function. Thus, we have proved the theorem. ∎

APPENDIX C: PROOF OF THEOREM 6

Proof: We prove it by induction using the DP updates. Let $\bar{T}_k(\mathcal{U}_k, e_k)$ be the *throughput-to-go* function from state $(\mathcal{U}_k, e_k)$ in slot $k \leq L^*$. We prove by induction that

$$\bar{T}_k(\mathcal{U}_k, e_k) = (1 - e_k)(1 - \epsilon) R_{\min}$$
$$\times \prod_{j=k}^{L^*-1} [(1 - \rho_k)(1 - p_{\text{fa}}) + \rho_k(1 - p_{\text{md}})]. \quad (80)$$

Then, (75) follows from $\bar{T}_{\text{err}} = \bar{T}_0(\mathcal{U}_0, 0)$. The induction hypothesis holds at $k = L^*$, since $\bar{T}_{L^*}(\mathcal{U}_{L^*}, e_{L^*}) = (1 - e_{L^*})(1 - \epsilon) R_{\min}$, see (72). Now, assume it holds for some $k + 1 \leq L^*$. Using the transition probabilities from state $(\mathcal{U}_k, 1)$ and the induction hypothesis, we obtain $\bar{T}_k(\mathcal{U}_k, 1) = 0$. Instead, from state $(\mathcal{U}_k, 0)$ we obtain

$$\bar{T}_k(\mathcal{U}_k, 0) = \rho_k(1 - p_{\text{md}})\bar{T}_{k+1}(\mathcal{B}_k, 0)$$
$$+ (1 - \rho_k)(1 - p_{\text{fa}})\bar{T}_{k+1}(\mathcal{U}_k \setminus \mathcal{B}_k, 0)$$
$$= (1 - \epsilon) R_{\min} \prod_{j=k}^{L^*-1} \Big[(1 - \rho_k)(1 - p_{\text{fa}}) + \rho_k(1 - p_{\text{md}})\Big],$$

which readily follows by applying the induction hypothesis. The induction step is thus proved.

Let $\bar{E}_k(\mathcal{U}_k, e_k)$ be the *energy-to-go* from state $(\mathcal{U}_k, e_k)$ in slot $k \leq L^*$. We prove that

$$\bar{E}_k(\mathcal{U}_k, e_k) = \left[v_k^{(L^*)} + h_k + u_k(1 - e_k)\right] |\mathcal{U}_k|. \quad (81)$$

Then, (75) follows from $\bar{P}_{\text{err}} = \frac{1}{T_{\text{fr}}} \bar{E}_0(\mathcal{U}_0, 0)$, and by noticing that $v_0^{(L^*)}/T_{\text{fr}}$ is the power consumption in the error-free case, given in Theorem 4. The induction hypothesis holds at $k = L^*$, since $\bar{E}_{L^*}(\mathcal{U}_{L^*}, e_{L^*}) = (N - L^*)\phi_d\left(\frac{N R_{\min}}{N - L^*}, \epsilon\right)|\mathcal{U}_{L^*}| = v_{L^*}^{(L^*)} + h_{L^*} + u_{L^*}(1 - e_{L^*})$, with $v_{L^*}^{(L^*)}$ given by (55), $h_{L^*} = u_{L^*} = 0$, see (72). Now, assume it holds for some $k + 1 \leq L^*$. Using the transition probabilities from state $(\mathcal{U}_k, e_k)$, the induction hypothesis, and the fact that $|\mathcal{B}_k| = \rho_k|\mathcal{U}_k|$ and $|\mathcal{U}_k \setminus \mathcal{B}_k| = (1 - \rho_k)|\mathcal{U}_k|$, we obtain

$$\bar{E}_k(\mathcal{U}_k, 1) = \phi_s \rho_k|\mathcal{U}_k| + p_{\text{fa}}\bar{E}_{k+1}(\mathcal{B}_k, 1)$$
$$+ (1 - p_{\text{fa}})\bar{E}_{k+1}(\mathcal{U}_k \setminus \mathcal{B}_k, 1)$$
$$= \Big\{\phi_s \rho_k + \big(v_{k+1}^{(L^*)} + h_{k+1}\big)[p_{\text{fa}}\rho_k + (1 - p_{\text{fa}})(1 - \rho_k)]\Big\}|\mathcal{U}_k|;$$
$$\bar{E}_k(\mathcal{U}_k, 0) = \phi_s \rho_k|\mathcal{U}_k| + \rho_k(1 - p_{\text{md}})\bar{E}_{k+1}(\mathcal{B}_k, 0)$$
$$+ (1 - \rho_k)p_{\text{fa}}\bar{E}_{k+1}(\mathcal{B}_k, 1)$$
$$+ (1 - \rho_k)(1 - p_{\text{fa}})\bar{E}_{k+1}(\mathcal{U}_k \setminus \mathcal{B}_k, 0)$$
$$+ \rho_k p_{\text{md}}\bar{E}_{k+1}(\mathcal{U}_k \setminus \mathcal{B}_k, 1)$$
$$= \Bigg\{\phi_s \rho_k + \big(v_{k+1}^{(L^*)} + h_{k+1} + u_{k+1}\big)$$
$$\times \Big[\rho_k^2(1 - p_{\text{md}}) + (1 - \rho_k)^2(1 - p_{\text{fa}})\Big]$$
$$+ \big(v_{k+1}^{(L^*)} + h_{k+1}\big)\rho_k(1 - \rho_k)(p_{\text{fa}} + p_{\text{md}})\Bigg\}|\mathcal{U}_k|.$$

The induction step $\bar{E}_k(\mathcal{U}_k, e_k) = (v_k^{(L^*)} + h_k + u_k(1 - e_k))|\mathcal{U}_k|$ can be finally proved by expressing $v_k^{(L^*)} = g_k(\rho_k)$ and $\rho_k = \frac{1}{2} - \frac{\phi_s}{4 v_{k+1}^{(L^*)}}$ using (59), and using (76)-(77). ∎

REFERENCES

[1] M. Hussain and N. Michelusi, "Optimal interactive energy efficient beam-alignment for millimeter-wave networks," in *Proc. 52st Asilomar Conf. Signals, Syst., Comput.*, 2018, pp. 1–5.

[2] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2016–2021," Cisco, San Jose, CA, USA, White Paper 1454457600805266. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[3] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[4] T. S. Rappaport, R. W. Heath, Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2014.

[5] N. Michelusi and M. Hussain, "Optimal beam-sweeping and communication in mobile millimeter-wave networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jan. 2018, pp. 1–6.

[6] M. Hussain and N. Michelusi, "Throughput optimal beam alignment in millimeter wave networks," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2017, pp. 1–6.

[7] R. A. Hassan and N. Michelusi, "Multi-user beam-alignment for millimeter-wave networks," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2018, pp. 1–7.

[8] M. Hussain and N. Michelusi, "Energy efficient beam-alignment in millimeter wave networks," in *Proc. 51st Asilomar Conf. Signals, Syst., Compute.*, Oct./Nov. 2017, pp. 1219–1223.

[9] J. Zhang, Y. Huang, Q. Shi, J. Wang, and L. Yang, "Codebook design for beam alignment in millimeter wave communication systems," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4980–4995, Nov. 2017.

[10] S. Haghighatshoar and G. Caire, "The beam alignment problem in mmwave wireless networks," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 741–745.

[11] C. Jeong, J. Park, and H. Yu, "Random access in millimeter-wave beamforming cellular networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 180–185, Jan. 2015.

[12] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, "Initial beamforming for mmWave communications," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 1926–1930.

[13] J. Seo, Y. Sung, G. Lee, and D. Kim, "Training beam sequence design for millimeter-wave MIMO systems: A POMDP framework," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1228–1242, Mar. 2016.

[14] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath, Jr., "Radar aided beam alignment in mmWave V2I communications supporting antenna diversity," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Jan./Feb. 2016, pp. 1–7.

[15] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-Wave beam steering without in-band measurement," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr./May 2015, pp. 2416–2424.

[16] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, Jr., "Beam design for beam switching based millimeter wave vehicle-to-infrastructure communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[17] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, Jr., "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4042–4058, May 2018.

[18] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[19] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive channel estimation and tracking for large arrays in mm-Wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, Apr. 2016.

[20] IEEE Standard 802.15.3c-2009, Oct. 2009, pp. 1–200.

[21] IEEE Standard 802.11ad-2012, Dec. 2012, pp. 1–628.

[22] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sep. 2017.

[23] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

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

[25] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[26] C. N. Barati *et al.*, "Initial access in millimeter wave cellular systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 7926–7940, Dec. 2016.

[27] Y. Li, J. G. Andrews, F. Baccelli, T. D. Novlan, and J. C. Zhang, "Design and analysis of initial access in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6409–6425, Oct. 2017.

[28] V. Va, J. Choi, and R. W. Heath, Jr., "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5014–5029, Jun. 2017.

[29] M. Gapeyenko *et al.*, "On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10124–10138, Nov. 2017.

[30] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 37, no. 3, pp. 379–423, 1948.

[31] M. Hussain, D. J. Love, and N. Michelusi, "Neyman-pearson codebook design for beam alignment in millimeter-wave networks," in *Proc. 1st ACM Workshop Millim.-Wave Netw. Sens. Syst.*, Oct. 2017, pp. 17–22.

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

[33] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*. New York, NY, USA: Springer, 2002.

[34] M. Hussain and N. Michelusi, "Coded energy-efficient beam-alignment for millimeter-wave networks," in *Proc. 56th Annu. Allerton Conf. Commun., Control, Comput.*, 2018, pp. 1–6.

[35] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA, USA: Athena Scientific, 2005.

[36] M. Hussain and N. Michelusi. (2018). "Energy-efficient interactive beam-alignment for millimeter-wave networks." [Online]. Available: https://arxiv.org/abs/1805.06089

Muddassar Hussain received the bachelor's degree in electrical engineering from the School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2013. He is currently pursuing the Ph.D. degree with Purdue University, West Lafayette, IN, USA. He was a Research Assistant with the Information Processing and Transmission Lab, SEECS, NUST, from 2013 to 2016. His research interests lie in the areas of 5G wireless communication systems, millimeter-wave communications, and stochastic optimal control.

Nicolo Michelusi (S'09–M'13–SM'18) received the B.Sc. (Hons.), M.Sc. (Hons.), and Ph.D. degrees from the University of Padua, Italy, in 2006, 2009, and 2013, respectively, and the M.Sc. degree in telecommunications engineering from the Technical University of Denmark in 2009, as part of the T.I.M.E. Double Degree Program. He was a Post-Doctoral Research Fellow with the Ming-Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA, from 2013 to 2015. He is currently an Assistant Professor with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. His research interests lie in the areas of 5G wireless networks, millimeter-wave communications, stochastic optimization, and distributed optimization. He serves as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and as a reviewer for several IEEE TRANSACTIONS.