The Trade-off between Scanning Beam Penetration and Transmission Beam Gain in mmWave Beam Alignment

Nariman Torkzaban
University of Maryland, College Park
College Park, MD
narimant@umd.edu

Mohammad A. (Amir) Khojastepour
NEC Laboratories, America
Princeton, NJ
amir@nec-labs.com

John S. Baras
University of Maryland, College Park
College Park, MD
baras@umd.edu

2022 56th Asilomar Conference on Signals, Systems, and Computers | 978-1-6654-5906-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IEEECONF56349.2022.10051971

Abstract—**Beam search algorithms have been proposed to align the beams from an access point to a user equipment. The process relies on sending beams from a set of scanning beams (SB) and tailoring a transmission beam (TB) using the received feedback.**

In this paper, we discuss a fundamental trade-off between the gain of SBs and TBs. The higher the gain of an SB, the better the penetration of the SB and the higher the gain of the TB the better the communication link performance. However, TB depends on the set of SBs and by increasing the coverage of each SB and in turn reducing its penetration, there is more opportunity to find a sharper TB to increase its beamforming gain. We define a quantitative measure for such trade-off in terms of a trade-off curve. We introduce SB set design namely *Tulip design* and formally prove it achieves this fundamental trade-off curve for channels with a single dominant path. We also find closed-form solutions for the trade-off curve for special cases and provide an algorithm with its performance evaluation results to find the trade-off curve revealing the need for further optimization on the SB sets in the state-of-the-art beam search algorithms.

I. Introduction

In pursuance of larger bandwidth that is required for realizing one of the main promises of 5G, i.e. enhanced mobile broadband (eMBB), millimeter wave (mmWave) communications is a key technology due to abundance of unused spectrum available at mmWave frequency ranges [1]. However, high path loss and poor scattering associated with mmWave communications leads to intense shadowing and severe blockage, especially in dense urban environments. These are among the major obstacles to increase data rate in such high frequency bands. To tackle these issues effective beamforming (BF) techniques are required to avoid the power leakage to undesired directions using directional transmission patterns, i.e., narrow beams [2][3]. Furthermore, several experimental results demonstrate that the mmWave channel usually consists of a few components (a.k.a spatial clusters) [4]. Therefore, it is essential to align the devised narrow transmission beams with the direction of the channel components. The problem of aligning the directions of the beams with the angle of departure (AoD) associated with clusters of the channel, is termed as the *beam alignment (BA)* problem. In the literature the beam alignment problem is also indexed as *beam training* or *beam search*. Devising effective beam alignment schemes is essential since a slight deviation of the transmitted beam AoDs from the mmWave channel clusters may result in a severe drop in the beamforming gain [5][6].

Beam alignment schemes may be categorized as *exhaustive search (ES)* and *hierarchical search (HS)*. Under the ES scheme, a.k.a beam sweeping, the angular search space is divided into multiple angular coverage intervals (ACIs) each covered by a beam. Then the beam with the highest received signal strength at the receiver is chosen [7][8]. To yield narrower beams in ES, the number of beams increases which results in larger beam sweeping overhead. The HS scheme lowers the overhead of the beam search by first scanning the angular search space by coarser beams and then gradually finer beams [9][10][11]. The BA procedure may happen in one of the two modes, i.e. *interactive* BA (I-BA), and *non-interactive* BA (NI-BA). In the NI-BA mode, the transmitter sends the scanning packets in the scanning phase and receives the feedback from the users after the scanning phase is over, while in the I-BA mode the transmitter receives the feedback for the previously transmitted scanning pilots during the scanning phase and can utilize this information in the rest of the scanning phase. Most of the prior art on I-BA are limited to single-user scenarios while NI-BA schemes can handle multi-user scenarios as the set of scanning beam does not change or depend on the received feedback from the users. The process of beam search relies on sending beams out of a set of scanning beams (SB) and tailor a data transmission beam (TB) based on the received feedback. In this paper, we discuss a fundamental trade-off between the gain of the SBs and TBs. The higher the gain of a SB, the better the penetration of the SB and the higher the gain of TB the better the communication link performance. However, TB depends on the set of SBs and by increasing the coverage of each SB, there is more opportunity to find a sharper TB to increase its beamforming gain. This means that the beamforming gain and hence the penetration of the SB is reduced. We define a quantitative measure for such trade-off in terms of a trade-off curve. We prove a fundamental result by finding the class of SB set designs namely *Tulip design* which achieves this fundamental trade-off curve for channels with single dominant path. We also find closed-form solutions for the trade-off curve for special cases and provide an algorithm with its performance evaluation results to find the trade-off curve in general. Our results reveal that the state-of-the-art beam search algorithm should further optimize their SB sets based on this fundamental trade-off between the SB penetration and TB gain.

The remainder of the paper is organized as follows. Section II describes the system model. In Section III we elaborate on the notion of trade-off curve and and propose our main contributions in section IV. Section V presents our evaluation results, and finally, in Section VI, we highlight our conclusions.

II. System model

We consider a mmWave communications scenario with a single base station (BS) and an arbitrary number of mobile users (MUs), say N, where prior knowledge on the value of N may or may not be available at the BS. The BA procedure aims at obtaining the accurate AoDs corresponding to the downlink mmWave channel from the BS to the users. Under the BA procedure, the BS transmits probing packets in different directions via various *scanning beams (SBs)* and receives feedback from all the users, based on which the BS computes a *transmission beam (TB)* for each user.

A. Channel Model

Let Ψ_j denote the random AoD vector corresponding to the channel between the BS and the j^{th} MU. Denote by f_{Ψ_j}, defined over $\mathcal{D} \subset (0, 2\pi]$, the probability density function (PDF) of Ψ_j. The PDF $f_{\Psi_j}(.)$ encapsulates the knowledge about the AoD of the j^{th} user prior to the BA procedure, or may act as a priority function over the angular search domain. Such information may be inferred from previous beam tracking, training, or alignment trials. A uniform distribution means lack of any prior knowledge over the search domain.

B. Beamforming Model

We consider a multi-antenna base station with an antenna array of large size realizing beams of high resolution. For power efficiency, we assuming hybrid beamforming techniques are in effect in the BS deploying only a few RF chains. Further, we adopt a *sectored antenna model* where each beam is modeled by the constant gain of its main lobe, and the angular coverage interval (ACI) it covers. Such models are widely adopted in the literature for modeling the beamforming gain and the directivity of mmWave transmitters.

C. Time-slotted System Model

We consider a system operating under the time division duplex (TDD) and the NI-BA schemes, with frames of length T. Each frame consists of T equal slots. In each frame, the first b slots are dedicated to the transmission of the probing packets, denoted by *scanning time-slots (STS)* and the next d slots denoted by *feedback time-slots (FTS)* are allocated to receiving the users feedback that may arrive through a side channel or according to any random access mode. Finally, the last $T + b + d$ slots are reserved for data transmission, namely *data transmission time-slots (DTS)*.

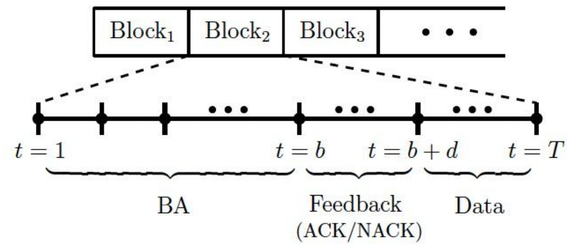

Fig. 1: Time-slotted System Model

D. Beam Alignment Model

The objective of the BA scheme is to generate narrowest possible TBs for the data transmission phase for each user to produce beams of higher gain and quality. In other words, utilizing the feedbacks provided by the users in the FTS to the SBs transmitted by the BS in the STS, the BS aims at localizing the AoD of each user to minimize the *uncertainty region (UR)* for each AoD. Let $\mathcal{B} = \{\Phi_i\}_{i=1}^{b}$ be the set of STS scanning beams where Φ_i denotes the ACI of the SB sent over time-slot $i \in [b]$. The feedback provided by each user to each SB is binary. If the AoD corresponding to at least one of the resolvable paths in the channel from the BS to the MU is within the ACI of the SB, then the MU will receive the probing packet sent via that SB and feeds back an acknowledgment (ACK). Otherwise, the feedback of the MU will be considered as a negative acknowledgment (NACK) indicating none of the user AoD's lie in the ACI of the SB. Once the FTS ends, the BS will determine the TBs using the SBs and the feedback sequences provided by the users according to the BA *policy*. The BA policy is formally defined as a function from the set of feedback sequences to the set of TBs.

III. Problem Formulation

To reach more distant users, in the scanning phase of the BA procedure, we are in need of a SB set with high-gain beams, i.e., beams as narrow as possible. However, there is a limit on sizes of the SBs in order to cover the entire angular domain $\mathcal{D}$ (say in azimuth angle). More precisely, the union of the set of SBs have to be at least equal to $\mathcal{D}$. Moreover, the intersections of the SBs contribute to the possibility of having additional TBs which in turn allows for optimizing the set of TBs. For example, using b SBs, a GES scheme has b TBs while a tulip design has $2b$ TBs. This would allow tulip design to have smaller TBs than GES in the expense of having larger SBs. In general, there exists a trade-off between the beamwidth of the beams in the SB and the TB sets. To formally define a quantitative measure for such tradeoff we first review some preliminaries.

A. Preliminaries

1) Beam Alignment Policies: The BA policy determines how the direction of the TBs is computed. This decision naturally considers the UR of the AoDs of each user channel. In the literature, two main policies have been considered frequently, i.e. i) *Spatial Diversity (SD)*, and ii) *Beamforming (BF)* Policy. The SD policy aims at generating TBs with minimal angular span that cover all the angular intervals that may contain a resolvable path, while the promise of the BF policy is to generate TBs that cover at least one resolvable path but further reduce the angular span of the TBs. The advantage of the SD policy to the BF policy is its resilience against the potential failure or blockage of one or some of the spatial clusters as long as at least one resolvable path remains, while the BF policy has the advantage of producing much higher beamforming gains compared to that of the SD policy but it is vulnerable to path blockage.

Let $B^j_{\mathcal{P}}(\mathcal{B}, \mathbf{s})$ denote the UR of the j^{th} user providing the feedback sequence $\mathbf{s}$ under the policy $\mathcal{P} \in \{\mathrm{SD}, \mathrm{BF}\}$ and the SB set $\mathcal{B}$. Further, let the *positivity set* $A^j(\mathbf{s}) \subseteq [b]$ be the set of all indices corresponding to the SBs that are acknowledged by the j^{th} user. Define the *negativity set* $N^j(\mathbf{s}) \subseteq [b]$ in a similar fashion for the not acknowledged SBs.

Note that if the j^{th} user sends an ACK in response to the SB Φ_i, this would mean that $\Theta_i(\mathbf{s}) \doteq \Phi_i$ has at least one resolvable path. On the other hand, a NACK would mean that no resolvable paths reside in Φ_i and therefore, any resolvable path should exist in $\Theta_i(\mathbf{s}) \doteq \mathcal{D} - \Phi_i$, and $B^j_{\mathcal{P}}(\mathcal{B}, \mathbf{s}) \in \mathcal{D} - \Phi_i$ for the above-mentioned policies. Having this in mind, we can explicitly express the UR for the mentioned policies as follows.

$$B^j_{\mathrm{SD}}(\mathbf{s}) = \left(\cup_{i \in A(\mathbf{s})} \Theta_i(\mathbf{s})\right) \cap \left(\cap_{i \in N(\mathbf{s})} \Theta_i(\mathbf{s})\right) \quad (1)$$

$$B^j_{\mathrm{BF}}(\mathbf{s}) = \Theta_k(\mathbf{s}) \cap \left(\cap_{i \in N(\mathbf{s})} \Theta_i(\mathbf{s})\right) \quad (2)$$

where $k = \arg\min_{l \in A(\mathbf{s})} \left|\Theta_l(\mathbf{s}) \cap \left(\cap_{i \in N(\mathbf{s})} \Theta_i(\mathbf{s})\right)\right|$.

2) Beam Alignment Formulation: . We assume there are N users that are prioritized according to the weight vector $\{c_j \geq 0\}_{j=1}^N, \sum_{j=1}^N c_j = 1$. Let $\mathcal{U} = \{u_k\}_{k=1}^M$ denote the range of the policy function $B^j_{\mathcal{P}}(\mathcal{B}, \mathbf{s})$. In other words, the TBs resulting from the BA scheme may take any value in the set $\mathcal{U}$. The expected value of the average beamwidth resulted from the BA scheme for policy $\mathcal{P}$ is

$$\bar{U}_{\mathcal{P}}(\mathcal{B}) = \sum_{j=1}^{N} c_j \mathbb{E}\left[|B_{\mathcal{P}}(\mathbf{s})|\right], \qquad \text{where,} \quad (3)$$

$$\mathbb{E}\left[|B_{\mathcal{P}}(\mathbf{s})|\right] = \sum_{k=1}^{M} |u_k| \, \mathbb{P}\left\{B_{\mathcal{P}}(\mathbf{s}) = u_k\right\} \quad (4)$$

and $|u_k|$ denotes the Lebesgue measure of the u_k. Note that u_k may be a finite union of multiple intervals in which case $|u_k|$ will be the sum of their widths. Given the value of b the objective of the BA scheme is to design $\{\Phi_i\}_{i=1}^b$ such that the expected average TB beamwidths as in (3) gets minimized. i.e.,

$$\{\Phi_i^*\}_{i=1}^b = \arg\min_{\{\Phi_i\}_{i=1}^b} \bar{U}_{\mathcal{P}}\left(\{\Phi_i\}_{i=1}^b\right) \quad (5)$$

The solution to the optimization problem is presented in great details in [12]. We summarize the main results as follows.

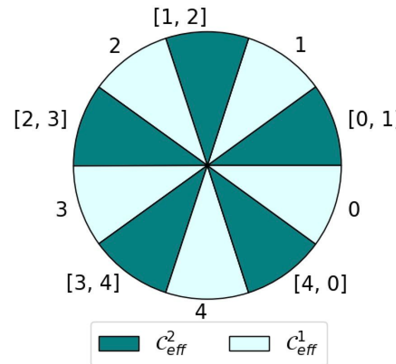

Fig. 2: Example of a Tulip design for $b = 5$

Theorem 1. *Among the set of contiguous scanning beams, a set of scanning beams with* Tulip design *generates the maximal number of possible feedback sequences for the channel with* $p = 1$ *and* 2*, for an arbitrary distribution of channel AoD that is nonzero on any points in the range* $[0, 2\pi)$.

Proof. Please see [13]. The Tulip design is defined as follows.

Definition 2. *Tulip design is given by a set of contiguous SBs* $\mathcal{B} = \{\Phi_i\}, i \in [b]$ *where each beam may only have intersection with its adjacent beams with the exception of* Φ_1 *and* Φ_b *for which the intersection might be nonempty. This means* $\Phi_i \cap \Phi_j = \varnothing$, $1 < |i - j| < b - 1$.

An example of a Tulip design is given in Fig. 2. Any scanning beam in the Tulip design consists of three parts, namely two side lobes (common between two neighboring scanning beams) and one middle lobe. We denote by each component a *component beam* (CB) and define the set $\mathcal{C}$ as the CB set. We denote by the middle lobes the CBs of first type and and by the side lobes the CBs of second type. In the above example we have,

$$\mathcal{C}^1 = \{0, 1, 2, 3, 4\}, \mathcal{C}^2 = \{[0, 1], [1, 2], [2, 3], [3, 4], [4, 0]\} \quad (6)$$

Next, we define the notion of trade-off curve.

B. Trade-off curve

Let $\lambda_{\mathcal{P}}(s)$ denote the beamwidth of TB for the received feedback sequence s and $\eta_i = |\Phi_i|$ denote the beamwidth of the SB i, i.e., Φ_i . Consider a measure $\mu(.)$ which is a function from a set of real numbers to a single real number. In this paper, we particularly are interested in two measures: max and average. We define the max function as $\mu_1(\{\eta_i\}_{i=1}^b) = \max(\{\eta_i\}_{i=1}^b)$ and the average function as $\mu_2(\{\eta_i\}_{i=1}^b) = (1/b)\sum_{i=1}^b \eta_i$. The tradeoff curve for the policy $\mathcal{P}$, using b scanning beams for the channel with p path is defined as the minimum value of $\mu(\{\lambda_{\mathcal{P}}(s)\}_{s \in S})$ for any given value of $\mu(\{\eta_i\}_{i=1}^b)$ where S is the set of all possible feedback sequences.

IV. Main Results

In order to find the trade-off curve, one needs to consider all possible SB designs in general. In this section, we provide an interesting result on the class of SB designs that outperform all other designs in terms of the trade-off between the SB and TB sizes. In particular, we prove that for the channel with single dominant path $p = 1$, the trade-off curve is always achieved by a SB design that follows tulip design for both max and average measures. The results for multi-path channels is out of the scope of this paper. We note that for $p = 1$ both the diversity and beamforming policies collapse into one. As mentioned before, different measures for the SBs and TBs can be taken into account for demonstrating the trade-off. For the max measure, we consider the pair $(\eta_{max}, \lambda_{max})$ where λ_{max} represents the maximum resulting beamwidth of the beams in the TB set, and η_{max} denotes the maximum beamwidth for the set of SBs. The trade-off curve may be interpreted as follows. For a given value of η_{max} there is a minimum achievable λ^{min}_{max}. The *trade-off curve* is defined as the set of all achievable $(\eta_{max}, \lambda^{min}_{max})$. Also, for a given value of λ_{max} there is a minimum achievable η^{min}_{max}. Alternatively, the *trade-off curve* is composed of all achievable $(\eta^{min}_{max}, \lambda_{max})$. For the average measure, we consider the pair $(\bar{\eta}, \bar{\lambda})$ where $\bar{\lambda}$ represents the average resulting beamwidth of the beams in the TB set, and $\bar{\eta}$ denotes the average beamwidth for the set of SBs. We have the following theorem.

Theorem 3. *The optimal (i.e., generating the trade-off curve) BA scheme is achieved by a SB set that follows the Tulip design.*

Proof. Each SB can be represented by its coverage interval which is an arc on the Trigonometric circle. Hence, if we move in a counter clockwise direction each SB has a starting point and an ending point. A *marker* is defined as any starting point or ending point. Please note that in general multiple starting or ending points may be concurrent (i.e., the same) in which case such marker is called *non-singular* marker. If a marker only represents a single starting point or a single ending point it is called a *singular* marker. The proof consist of multiple steps.

Step 1: We exchange each of the non-singular markers with multiple singular markers which are only a differential amount apart in such a way that all starting point are moved in clockwise directions and all ending points are moved in counter clockwise direction. Please note that in this process the ordering between the starting points or between the ending points is not important.

Step 2: There are exactly $2b$ markers and hence $2b$ arcs on the circle. Each SB is comprised of one or more arcs that are adjacent. An arc may be shared between multiple SBs. The *order* of an arc is the number of scanning beams that intersect with that arc. Since each SB is contiguous and each marker is singular, the order of the arcs are alternatively odd and even on the circle. It can be easily verified that the order of the arc in the Tulip design is alternatively 1 and 2. For the original design, we associate a Tulip design where the corresponding arcs in both designs are either both odd or even (See Fig. 2).

On one hand, we have

$$\mu_2(\{\eta_i\}_{i=1}^{b}) = (1/b)\sum_{i=1}^{b}\eta_i = (1/b)\sum_{i=1}^{2b}\mathrm{arc}_i n_i$$

where arc_i is the length of the arc i and n_i is its order. Similarly, for $\mu_2(\{\eta_i\}_{i=1}^{b})$ we note that each SB in the Tulip design is entirely covered by at least one SB in the original design. To see this, we note that each SB in Tulip design consists of three arcs, say, $i-1, i, i+1$ where their orders n_{i-1}, n_i, n_{i+1}, are even, odd, and even, respectively. We note that $|n_{i-1} - n_i| = 1$ and $|n_{i+1} - n_i| = 1$ since each marker is singular and hence the sets of the beams that cover any two adjacent arc exactly differ by one beam. Either, we have $n_{i-1} > n_i > n_{i+1}$ in which case all the beams that cover the arc $i+1$ would cover the arcs $i-1$ and i which means that in the original design there is a beam that covers all three arcs $i-1, i, , i+1$. Hence, in the Tulip design the beam that covers exactly the arcs $i-1, i, i+1$ is covered by a beam in the original design. If $n_{i-1} > n_i$, $n_i < n_{i+1}$, all the beams that cover the arc i have to cover the arc $i-1$ and $i+1$. If $n_{i-1} < n_i < n_{i+1}$, all the beams that cover the arc $i-1$ have to cover the arc i and $i+1$. Finally, if $n_{i-1} < n_i$, $n_i > n_{i+1}$, all the beams that cover the arc $i-1$ are the same as the beams that cover the arc $i+1$ and they cover the arc i as well, but the arc i covers exactly one different beam. Hence, it is immediate that the average (or max) of SB for a tulip design that matches the same marker positions is not more than the average (or max) of SB for the original design. On the other hand, we note that for both the original design and the Tulip design the position of the markers are the same. Moreover, the set of TBs for Tulip design is all possible $2b$ arcs. Since any other beam design including the original design can at most distinguish the same $2b$ arcs, hence, any measure function (e.g., max or mean) on the set of TBs for Tulip design is always less than or equal to that of the original design. This means that the trade-off curve is achieved by the Tulip design.

Step3: We note that after replacing the original SBs with the Tulip design as described in Step 2, it is possible to undo the operation of the step 1 and revert back by collapsing the corresponding group of singular markers into a single marker, for each non-singular marker. This would not affect the correctness of the arguments in step 2. This can also be interpreted as taking the limit over the differential value δ as it goes to zero which obviously collapses the corresponding set of singular points to its original non-singular point. We also note that the introduction of Step 1 is necessary for the way that the Tulip design is compared with the original design. Also, the step 3 will also reveal that the design which achieves the trade-off curve might be a special form of the Tulip design where, e.g., some arcs are diminishing to zero.

We note that the above theorem holds for any channel distribution, i.e., the distribution of the paths. For the case of the multi-path channel, i.e., $p \geq 2$, Theorem 3 does not necessarily hold true for an arbitrarily policy. The reason lies in the fact that the TB is in general composite, i.e., it consists of multiple arcs, and the composition of TBs is a function of the design. Therefore, in the evaluation section, we present the achievable trade-off curve for Tulip design in comparison to that of the other designs without any optimality claim.

A. Analytical Results for Uniform Distribution

The trade-off curve for uniform distribution somewhat exhibits the worst case performance. In this section, we drive the trade-off curve for both max and average measures under uniform distribution. The results can also facilitate the visualization of the trade-off curve for more general cases even though the shape and the end points of the trade-off curve vary for different distribution. First, by using Theorem 3 we derive the closed-form solution for the trade-off curve.

Theorem 4. *The trade-off curve for the average measure* $(\bar{\eta}, \bar{\lambda}) = \big(\mu_2(\{\eta_i\}_{i=1}^b), \mu_2(\{\lambda_{\mathcal{P}}(s)\}_{s\in S})\big)$ *and under uniform distribution follows a polynomial from order* 2.

Proof. Consider a Tulip design with $2b$ CBs where $\{\alpha_i\}_{i=1}^b$ and $\{\beta_i\}_{i=1}^b$ denote the CBs of degree 1 and 2 respectively. For a given $\mu_2(\{\eta_i\}_{i=1}^b)$ it holds that,

$$\mu_2(\{\eta_i\}_{i=1}^b) = 1/b\sum_{i=1}^b \eta_i = (1/b)\sum_{i=1}^b(\alpha_i + 2\beta_i) =$$

$$(1/b)\sum_{i=1}^b(\alpha_i+\beta_i) + (1/b)\sum_{i=1}^b\beta_i = \frac{2\pi}{b} + (1/b)\sum_{i=1}^b\beta_i \quad (7)$$

Hence

$$\sum_{i=1}^b \alpha_i = 4\pi - b\bar{\eta} \quad \text{and} \quad \sum_{i=1}^b \beta_i = b\bar{\eta} - 2\pi \quad (8)$$

For uniform distribution on the users AoD we can write

$$\bar{\lambda} = 1/2\pi\sum_{i=1}^b(\alpha_i^2 + \beta_i^2) \quad (9)$$

For a given $\bar{\eta}$, the summation $\sum_{i=1}^b \alpha_i$ and $\sum_{i=1}^b \beta_i$ is fixed, hence, it is easy to verify that the minimum value of $\bar{\mu}$ in (9) is obtained when $\alpha_i = \alpha$ and $\beta_i = \beta$ for all $i \in [b]$ for proper values of α and β. We have,

$$\bar{\lambda} = b/2\pi\left((\alpha+\beta)^2 - 2\alpha\beta\right) \quad (10)$$

Following equation (8) we get $\alpha = 2\gamma - \bar{\eta}$ and $\beta = \bar{\eta} - \gamma$. Replacing these values in (10) it immediately follows that

$$\bar{\lambda} = (2/\gamma)\bar{\eta}^2 - 6\bar{\eta} + 5\gamma \quad (11)$$

Hence, $\bar{\lambda}$ is a polynomial of degree 2 in terms of η. ☐

We note that for any SB design we have $\gamma \le \bar{\eta} \le 2\gamma$. On the other hand, the analytical curve in (11) has its minimum in $\bar{\eta} = 1.5\gamma$ and it is symmetric with respect to this point. Since a good operating point for $(\bar{\eta}, \bar{\lambda})$ is where both $\bar{\eta}$ and $\bar{\lambda}$ are minimized, it is trivial that the trade-off curve only for $\gamma \le \bar{\eta} \le 1.5\gamma$ is of interest. For example for $\gamma \le \bar{\eta} \le 1.5\gamma$, both $(\bar{\eta}, \bar{\lambda})$ and $(3\gamma - \bar{\eta}, \bar{\lambda})$ are on the trade-off curve and obviously $(\bar{\eta}, \bar{\lambda})$ is much more efficient than $(3\gamma - \bar{\eta}, \bar{\lambda})$. Hence, we focus our attention to the curve between $\gamma \le \bar{\eta} \le 1.5\gamma$.

Algorithm 1 *Greedy-SA*

Input: $\bar{\eta}_0, N, b, f_\Psi(\psi), done = \varnothing$
$G_N \doteq$ a set of N points in $[0, 2\pi]$
$\{z_i\}_{i=1}^{2b} \doteq$ random points from G_N such that $\bar{\eta} \le \bar{\eta}_0$
$\{z_i\}_{i=1}^{2b} \Leftrightarrow \mathcal{C}_{\text{eff}} \doteq \{(z_i, z_{i+1 \mod b}), i \in [2b]\}$
Compute $\bar{\lambda}$ using $f_\Psi(\psi)$ and $\mathcal{C}_{\text{eff}}$
while not $done$ **do**
  $(done, \bar{\lambda}, \mathcal{C}_{\text{eff}})$ = *modify-sol*(G_N, $\bar{\lambda}$, $\mathcal{C}_{\text{eff}}$, $\bar{\eta}_0$)
end while
Return $\bar{\lambda}$, $\mathcal{C}_{\text{eff}}$

Algorithm 2 *Modify-Sol*

Input: G_N, $\bar{\lambda}$, $\mathcal{C}_{\text{eff}}$, $\bar{\eta}_0$, $s = True$, $count = 0$
$dir = \{forward, backward\}$, $\bar{\lambda}_{old} = \bar{\lambda}$
$perm$ = Shuffle $\{(p,q,r) | p,q \in [2b], r \in dir, p \le q\}$
repeat
  Orderly select next tuple (p,q,r) from $perm$
  Slide $\{z_i\}_{i=p}^q$ in $r \in dir$ direction on points in G_N
  Compute $\bar{\eta}$ using $\mathcal{C}_{\text{eff}}$
  if $\bar{\eta} > \bar{\eta}_0$ **then**
    $count++$
  else
    Compute $\bar{\lambda}_{new}$ using $f_\Psi(\psi)$ and $\mathcal{C}_{\text{eff}}$
    if $\bar{\lambda}_{new} \ge \bar{\lambda}_{old}$ **then**
      $count++$
    end if
  end if
until $(count = 2b^2 + b) \vee (\bar{\lambda}_{new} < \bar{\lambda}_{old})$
if $count = 2b^2 + b$ **then**
  Return $(True, \bar{\lambda}_{new}, \mathcal{C}_{\text{eff}})$
else
  Return *modify-sol* (G_N, $\bar{\lambda}_{new}$, $\mathcal{C}_{\text{eff}}$)
end if

V. PERFORMANCE EVALUATION

The trade-off curve for arbitrary distribution in general may not admit an analytical form. Theorem 3 proves useful in this case by restricting the set of scanning beam to Tulip design. We use Algorithm 1 to compute the points on the trade-off curve for a given value of $\bar{\eta} = \bar{\eta}_0$. We start by setting $\bar{\eta}_0 = \gamma$ and increase its value by small value δ at each step. At each step for a given value of $\bar{\eta}_0$, we find the minimum $\bar{\lambda}$ by simply iterating Algorithm 1 for a given number of trials where at least one such starting set of scanning beams defined by $\{z_i\}_{i=1}^{2b}$ at each step is the one that correspond to the minimum value of $\bar{\lambda}$ in previous step. As we discussed earlier, the trade-off curve for the uniform distribution achieves its minimum at $\bar{\eta} = 1.5\gamma$ and any other distribution beside uniform achieves its minimum at $\bar{\eta} < 1.5\gamma$. Hence, it is not necessary to search for the value of $\bar{\eta} > 1.5\gamma$. After the trade-off curve reaches its minima say at $\bar{\eta}^*$ increasing the value of $\bar{\eta} > \bar{\eta}^*$ would not improve the curve. This is verified through our numerical evaluation as well. For empirical measurements, We consider uniform

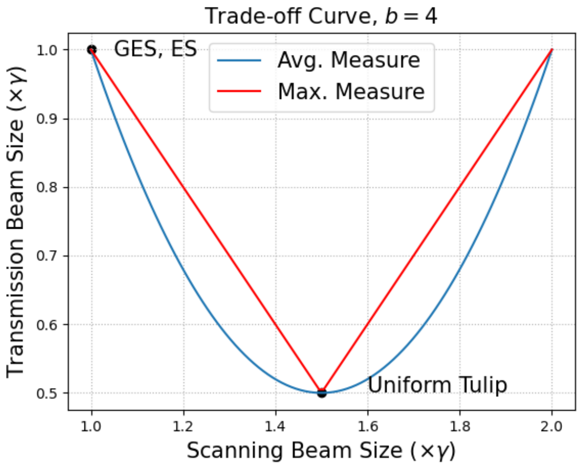

Fig. 3: Theoretical trade-off Curve for Uniform dist.

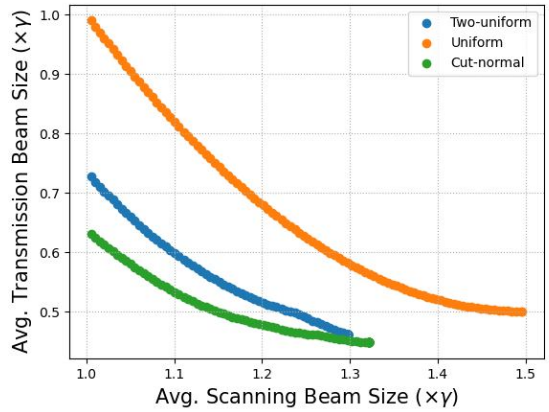

Fig. 4: Empirical trade-off curves for general dist.

distribution, the cut-normal distribution, i.e., $\mathcal{N}(\mu = \pi, \sigma = 1)$ that is truncated beyond the range $(0, 2\pi)$, and the piece-wise uniform distribution with two segments $A = [0, 2\pi/3]$ with $p(A) = 3/4$, and $B = [2\pi/3, 2\pi]$ with $p(B) = 1/4$. Fig. 4 illustrates the trade-off curve for $b = 5$ beams using average measure evaluated based on the proposed algorithm for these three distributions. As derived in Section IV the trade-off curve for uniform distribution is a a convex curve that starts from $\bar{\eta} = \gamma$ and admits its minimum at $\bar{\eta} = 1.5\gamma$. It is easy to verify that the trade-off curve for average measure is convex since a time sharing between every two SB sets can be used to generate points on the line connecting the performance points of these two SB sets. The trade-off curves for the same number of beams b for non-uniform distributions is always strictly below that of the uniform distribution. This can be interpreted as uniform distribution has the worst case in terms of trade-off. We note that other two trade-off curves for both piece-wise uniform distribution, and cut-normal distribution in Fig. 4 start from $\bar{\eta} = \gamma$ and take minima at a point for which $\bar{\eta} < 1.5\gamma$. Fig. 3 depicts the theoretical trade-off for uniform distribution on the user AoD for both measures discussed earlier. The curve corresponding to the Avg. and Max. measures follow parabolic (equation (11)) and linear curves, respectively. We note that the point (γ, γ) on the Avg. trade-off curve corresponds to the GES scheme and on the Max. trade-off curve to the ES scheme.

VI. Conclusions

We reveal a fundamental trade-off between the SBs and TBs gains and define the notion of trade-off curve. We show a fundamental class of SB set design, namely *Tulip Design* achieves the trade-off curve. We present a closed-form solution for the trade-off curve for special channel with uniform distribution on AoA. For general distributions, we provided an algorithm and evaluation results to find the trade-off curve. Our results emphasize that the state-of-the-art beam search algorithms should further optimize their SB sets based on this fundamental trade-off between the SB penetration and TB gain.

References

[1] S. A. Busari, S. Mumtaz, S. Al-Rubaye, and J. Rodriguez, “5g millimeter-wave mobile broadband: Performance and challenges,” *IEEE Communications Magazine*, vol. 56, no. 6, pp. 137–143, 2018.

[2] N. Torkzaban, M. A. Amir Khojastepour, and J. S. Baras, “Codebook design for composite beamforming in next-generation mmwave systems,” in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 1545–1550.

[3] N. Torkzaban and M. A. A. Khojastepour, “Codebook design for hybrid beamforming in 5g systems,” in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 1391–1396.

[4] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[5] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, “Steering with eyes closed: Mm-wave beam steering without in-band measurement,” in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2416–2424.

[6] S. Shahsavari, M. A. Amir Khojastepour, and E. Erkip, “Robust beam tracking and data communication in millimeter wave mobile networks,” in *2019 International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2019, pp. 1–8.

[7] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, “Initial access in millimeter wave cellular systems,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 7926–7940, 2016.

[8] M. Giordani, M. Mezzavilla, C. N. Barati, S. Rangan, and M. Zorzi, “Comparative analysis of initial access techniques in 5g mmwave cellular networks,” in *2016 Annual Conference on Information Science and Systems (CISS)*, 2016, pp. 268–273.

[9] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, “Initial beamforming for mmwave communications,” in *2014 48th Asilomar Conf. on Signals, Systems and Computers*, 2014, pp. 1926–1930.

[10] M. Hussain and N. Michelusi, “Throughput optimal beam alignment in millimeter wave networks,” in *2017 Information Theory and Applications Workshop (ITA)*, 2017, pp. 1–6.

[11] S. Noh, M. D. Zoltowski, and D. J. Love, “Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5689–5701, 2017.

[12] N. Torkzaban, M. A. Khojastepour, and J. S. Baras, “Multi-user beam alignment in presence of multi-path,” in *2022 56th Annual Conference on Information Sciences and Systems (CISS)*, 2022, pp. 224–229.

[13] N. Torkzaban and M. Khojastepour. [Online]. Available: https://www.nec-labs.com/uploads/Documents/Mobile-Communications/5GTech/Multi-user_Beam_Alignment_in_presence_of_Multi-path.pdf