Realizing Uplink MU-MIMO Communication in mmWave WLANs: Bayesian Optimization and Asynchronous Transmission

Shichen Zhang*, Bo Ji†, Kai Zeng‡, and Huacheng Zeng*

*Department of Computer Science and Engineering, Michigan State University
†Department of Computer Science, Virginia Tech
‡Department of Electrical and Computer Engineering, George Mason University

Abstract—**With the rapid proliferation of mobile devices, the marriage of millimeter-wave (mmWave) and MIMO technologies is a natural trend to meet the communication demand of data-hungry applications. Following this trend, mmWave multi-user MIMO (MU-MIMO) has been standardized by the IEEE 802.11ay for its downlink to achieve multi-Gbps data rate. Yet, its uplink counterpart has not been well studied, and its way to wireless local area networks (WLANs) remains unclear. In this paper, we present a practical uplink MU-MIMO mmWave communication (UMMC) scheme for WLANs. UMMC has two key components: i) an efficient Bayesian optimization (BayOpt) framework for joint beam search over multiple directional antennas, and ii) a new MU-MIMO detector that can decode asynchronous data packets from multiple user devices. We have built a prototype of UMMC on a mmWave testbed and evaluated its performance through a blend of over-the-air experiments and extensive simulations. Experimental and simulation results confirm the efficiency of UMMC in practical network settings.**

Index Terms—**mmWave, MU-MIMO, 6G, Bayesian learning**

I. INTRODUCTION

Recently, the marriage of millimeter-wave (mmWave) and multi-user multiple-input-and-multiple-output (MU-MIMO) technologies has attracted much research and development attention in wireless local area networks (WLANs) as it has potential to offer 100s Gbps data rate via simultaneous transmission of multiple independent data streams [1]. As a concrete step towards its real-life applications, downlink mmWave MU-MIMO has been standardized by IEEE 802.11ay [2], and its theoretical data rate can reach 176 Gbps.

However, the advancement of mmWave MU-MIMO is mainly limited to its downlink. Very limited progress has been made so far for its uplink. While both 802.11ac (sub-6GHz) and 802.11ay (60GHz) support downlink MU-MIMO, neither of them supports uplink MU-MIMO. This stagnation underscores the grand challenges in the design of practical yet efficient uplink mmWave MU-MIMO communication schemes. In addition, the demand of uplink data rate is dramatically increasing in emerging applications such as autonomous driving and video streaming. Ericsson predicts that the amount of global uplink traffic will reach 70 EB per month in 2027 [3]. Therefore, there is a critical need to fill this gap.

In this paper, we present a practical yet efficient uplink MU-MIMO mmWave communication scheme (UMMC) for a wireless local area network (WLAN). UMMC allows multiple stations to *simultaneously* send their data packets to an access point (AP) while not requiring fine-grained inter-station synchronization. We address two challenges in the design of UMMC. *The first challenge lies in the analog beamforming for a multi-antenna AP.* While the literature has a wealth of analog beamforming work, existing approaches can be generally classified into two categories: model-based optimization (e.g., [4], [5, Table V]) and model-free beam search (e.g., [6], [7], [8], [9], [10]). While model-based approaches offer the optimal antenna weight vectors (AWVs) for analog beamforming, they require accurate antenna models and channel knowledge, which are hard to obtain. Therefore, these approaches are not amenable to practical use. Model-free approaches do not require the above knowledge as they aim to find the best beam in a predefined beambook. However, most of them focus on maximizing the signal strength for a single-antenna mmWave device while minimizing their beam search overhead. While maximizing signal strength is equivalent to maximizing data rate in single-antenna systems, it is not the case in MU-MIMO systems. This is because the capacity of an MU-MIMO channel is dependent upon not just the signal strength but also the correlation of MIMO channels. When two stations have highly-correlated channels, the AP may not be capable of decoding their packets even if the signals are strong.

To address this challenge, we design a Bayesian optimization (BayOpt) framework for joint beam search at the AP. This framework is inspired by two facts: i) the relation between a selected beam and its achievable data rate in MU-MIMO communications is complex and unknown in real systems; and ii) BayOpt has been proved to be an effective technique for finding an optimal or near-optimal solution to an optimization problem whose objective function and constraints are unknown and costly to evaluate. The key idea of the BayOpt framework is to guide beam search using the posterior probability derived from those beams that have already been evaluated. The more beams we evaluate, the more accurate information we have for the remaining beams. Compared to exhaustive search, BayOpt appears to be surprisingly efficient in finding a near-optimal beam within a given airtime budget.

Another challenge in the design of UMMC is the synchronization among stations. Actually, the signal detection in uplink MU-MIMO transmission has been well studied in sub-6GHz wireless networks, and some signal detection methods such as zero-forcing (ZF) and minimum mean square error (MMSE) have been widely used in practice. However, existing signal detectors are based on an important assumption — the data packets from different stations are synchronized in time when impinging on the AP. Particularly, in OFDM systems, the time misalignment of the packets when arriving at the AP must be less than the time duration of an OFDM symbol's cyclic prefix (CP). While this requirement can be achieved in narrow-

band (20 MHz) sub-6GHz systems (e.g., using timing advance protocols), it is extremely challenging to achieve in ultra wideband mmWave systems. For instance, using conventional MU-MIMO detectors, the time misalignment of packets in 802.11ay must be less than 36ns, which is hard to maintain in practice. Due to this stringent requirement, uplink MU-MIMO has not yet been supported by 802.11ay standard [2].

To address this challenge, we argue that it is more desirable living with the packet misalignment at the AP instead of employing an onerous protocol to synchronize stations. Towards this goal, we observed that existing MU-MIMO detectors work in the *spatial* domain while the packet misalignment is an imperfection in the *temporal* domain. Since these two domains are orthogonal, spatial MU-MIMO detectors should be immune to temporal misalignment of data packets. In fact, the real problem is that the construction of existing MU-MIMO detectors requires the knowledge of channel, which relies on orthogonal pilots (reference signals) in data packets to estimate. However, misaligned packets cannot maintain the orthogonality of their pilots, making it hard to estimate channels. To solve this problem, we design an asynchronous MU-MIMO detector through a transformation of existing MMSE MU-MIMO detector. This new detector is capable of decoding asynchronous packets from multiple stations without the need of explicit channel knowledge. The key idea behind our design is to use the interfered pilots within each packet to train its detection filter. Doing so eliminates the need of channel matrix in the construction of the detection filter while achieving a surprisingly good performance. The new detector fundamentally relaxes the synchronization requirement for stations in uplink MU-MIMO transmission, making UMMC amenable to practical implementation.

We have evaluated UMMC through a blend of over-the-air experiments and extensive simulations. We implemented UMMC on a two-user MIMO mmWave (60GHz) testbed and demonstrated that it enables real-time uplink packet transmission in the absence of inter-user synchronization. Experimental results show that, compared to exhaustive beam search, BayOpt achieves 92% throughput while reducing the overhead by 98%. In addition, simulation results from a 100-user mmWave network show that, compared to exhaustive beam search, BayOpt achieves more than 80% of its throughput while entailing less than 5% of its overhead in all two-user, three-user, and four-user MIMO cases.

The contributions of this paper are summarized as follows.

- We design a practical uplink MU-MIMO mmWave communication scheme for WLANs. We demonstrate that it works in realistic scenarios via over-the-air experiments.
- We introduce the first-of-its-kind BayOpt framework for beam search in mmWave MU-MIMO systems, and show its efficiency through both simulation and experiments.
- We propose a new MU-MIMO detector that can decode the asynchronous data packets from multiple user devices. For the first time, it demonstrates via theory and experiments that fine-grained inter-user synchronization is not needed for uplink MU-MIMO mmWave transmission.

II. PROBLEM DESCRIPTION

We consider the uplink MU-MIMO communication in a WLAN as shown in Fig. 1, where an AP wishes to decode con-
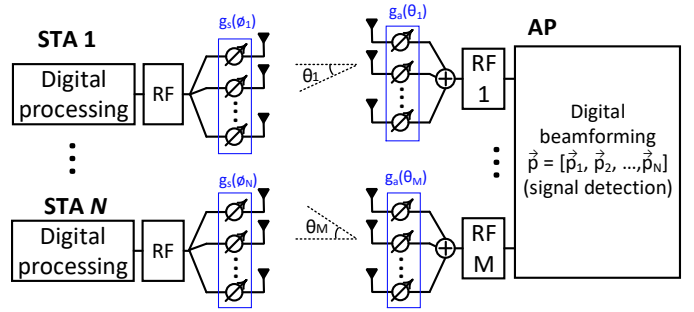

Fig. 1: Uplink MU-MIMO transmission in WLAN.

current data packets from multiple stations. Our objective is to maximize the uplink throughput through the design of analog and digital beamforming for the AP. In the pursuit of this objective, we assume that a beam has already been selected for each station using an existing beam search scheme such as sector-level sweep (SLS) and beam refinement protocol (BRP) [11]. We focus on the analog and digital beamforming at the AP for uplink MU-MIMO transmission.

A. Problem Formulation

Analog Beamforming. Denote M as the number of phased-array antennas (RF chains) on the AP and N as the number of stations involved in the uplink MU-MIMO transmission (assuming $N \leq M$). We assume that all the phased-array antennas on the AP are identical. Suppose that a linear phased-array antenna intends to steer its beam energy to the direction of θ. Then, its antenna weight vector (AWV) can be modeled as: $G_{\mathrm{ap}}(\theta) = [e^{j \frac{d_{\mathrm{ap}}}{\lambda} i \sin(\theta)}]_{0 \leq i \leq N_{\mathrm{ap}}-1}$, where d_{ap} is the patch element spacing, λ is the wavelength, and N_{ap} is the number of patch elements. Similarly, for the phased-array antenna on a station, suppose it intends to steer its beam energy to the direction of ϕ. Then, its AWV can be modeled as: $G_{\mathrm{sta}}(\phi) = [e^{j \frac{d_{\mathrm{sta}}}{\lambda} i \sin(\phi)}]_{0 \leq i \leq N_{\mathrm{sta}}-1}$, where d_{sta} is the patch element spacing and N_{sta} is the number of patch elements. Then, the signal received by the AP's mth RF chain can be written as:

$$y_m = \sum_{n=1}^{N} G_{\mathrm{ap}}(\theta_m) \, \tilde{\mathbf{H}}_{mn} \, G_{\mathrm{sta}}(\phi_n)^{\top} x_n + w_m, \qquad (1)$$

where x_n is the signal transmitted by the nth station, w_m is the received noise, $\tilde{\mathbf{H}}_{mn} \in \mathbb{C}^{N_{\mathrm{ap}} \times N_{\mathrm{sta}}}$ is the over-the-air channel between the AP's mth antenna and the nth station's antenna.

Digital Beamforming. At the AP (receiver), digital beamforming serves for the purpose of MU-MIMO Detection. Denote $\vec{y} = [y_1, y_2, \cdots, y_M]^{\top}$ as the received signals and $\vec{p}_n$ as the AP's spatial filter for decoding the data packets from station n. Then, the decoded version of the signal from station n can be written as: $\hat{x}_n = \vec{p}_n^{\mathsf{H}} \vec{y}$, for $1 \leq n \leq N$, where $(\cdot)^{\mathsf{H}}$ is the conjugate transpose operator.

Design Objective. At the AP, denote $\vec{\theta} = [\theta_1, \theta_2, \cdots, \theta_M]$ as the beam angle vector, which can be directly used to calculate the AWV for analog beamforming. Denote $\vec{p} = [\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_N]$ as the detection vector. Denote EVM_n as the error vector magnitude (EVM) of the decoded signals from station n, i.e., $\mathrm{EVM}_n \equiv \frac{\mathbb{E}[|x_n - \hat{x}_n|^2]}{\mathbb{E}[|x_n|^2]}$. Without loss of

generality, we assume that the transmit power at stations are normalized, i.e., $\mathbb{E}[|x_n|^2] = 1$. Then, we have

$$\text{EVM}_n = \mathbb{E}[|x_n - \hat{x}_n|^2]. \quad (2)$$

The link capacity (spectral efficiency) between station n and the AP can be written as: $c_n = \log_2(1 + \frac{1}{\text{EVM}_n})$.

In uplink MU-MIMO, it is important not only to maximize the data rate but also ensure the fairness among users. Thus, our objective is to pursue the best analog and digital beams so that the bottleneck link data rate can be maximized. Mathematically, it can be formulated as:

$$[\vec{\theta}^*, \ \vec{p}^*] = \underset{\vec{\theta} \in \mathcal{B}, \ \vec{p}}{\arg \max} \left(\min_n \log_2(1 + \frac{1}{\text{EVM}_n}) \right), \quad (3)$$

where $\mathcal{B}$ is the predefined beambook that includes all possible beam angle vectors.

The optimization problem in (3) can be divided into two subproblems: i) analog beam selection (determining $\vec{\theta}$), and ii) MU-MIMO detector construction (determining $\vec{p}$). These two subproblems are tightly coupled with each other. Given the complex nature of this problem, it is intractable to pursue a global optimal solution in real systems. Therefore, we develop a practical yet efficient scheme to solve the two subproblems.

B. Key Challenges

Inaccurate Models. Solving the above optimization is nontrivial as the gradients of the objective function are unknown, so first-order methods like gradient descent cannot be applied. In addition, we used $G_{\text{ap}}(\theta)$ and $G_{\text{sta}}(\phi)$ to model the response of ideal linear phased-array antennas. In practice, phased-array antennas have many imperfections in their radiation patterns. Their actual mathematical models are unknown. The discrepancy between the ideal and real antenna model significantly affects the beamforming design.

Channel Correlation. The capacity of MU-MIMO transmission is determined by not only *signal strength* but also *MIMO channel correlation*. Existing approaches based on signal strength only are not suitable for beam search in MU-MIMO. Therefore, it calls for a beam search scheme that can *jointly* identify the best beams for *all* antennas. One straightforward approach is exhaustive search. However, it will entail a large airtime overhead and thus compromise the throughput gain of MU-MIMO. *Therefore, an efficient joint beam search scheme is needed.*

Inter-Station Timing Synchronization. Uplink MU-MIMO detection has been well studied. However, existing schemes require fine-grained inter-user timing synchronization for signal detection. That is, the time misalignment of data packets from different stations must be less than OFDM CP length. In 802.11ay [2], the normal guard interval duration (CP) is 36.36ns. Maintaining the inter-user synchronization within 36.36ns not only entails a large overhead but also complicates the network design and operation. For this reason, neither 802.11ac (sub-6GHz) nor 802.11ay (60GHz mmWave) supports uplink MU-MIMO.

III. OVERVIEW OF UMMC

In this section, we first highlight our approaches to overcoming the above challenges and then present the overall system diagram of UMMC. In what follows, we denote
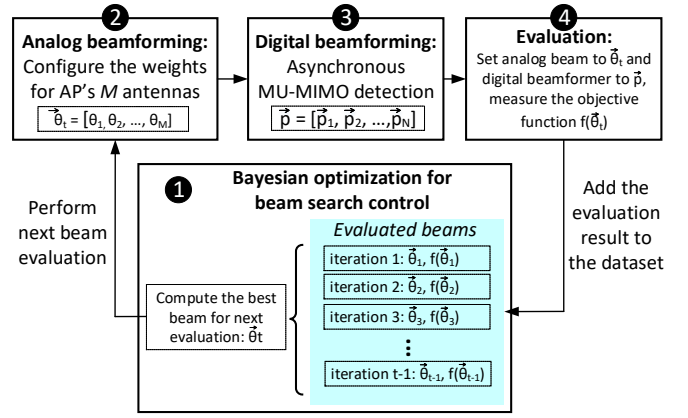

Fig. 2: The high-level system diagram of UMMC.

$f(\vec{\theta}) = \max_n\{\text{EVM}_n\}$. When $\vec{p}$ is given, the optimization in (3) is equivalent to minimizing $f(\vec{\theta})$.

A. Our Approaches

Analog Beam Search. To address the beam search challenge, we design a BayOpt scheme for joint beam search. BayOpt has been proved to be an effective technique for solving sequential optimization problems where the objective function is complex (treated as a black-box), the (sub-)gradient is unknown, and the evaluation is expensive [12]. To illustrate the idea behind BayOpt, let us consider the beams in a beambook $[\vec{\theta}_1, \vec{\theta}_2, \cdots, \vec{\theta}_{3600}]$. Suppose that we have measured two beams, say $\vec{\theta}_{10}$ and $\vec{\theta}_{1000}$, and found that $f(\vec{\theta}_{10}) = 5$ and $f(\vec{\theta}_{1000}) = 0.1$. Then, in the next iteration we should select a beam in the neighborhood of $\vec{\theta}_{1000}$ to evaluate, because the global minimum is more likely sitting in the neighborhood of $\vec{\theta}_{1000}$ compared to $\vec{\theta}_{10}$. BayOpt is a principled strategy to guide the process of joint beam search based on posterior probability.

MU-MIMO Detection. Inter-station synchronization is a fundamental problem for uplink MU-MIMO. Achieving the required timing alignment for packet transmission among distributed stations is extremely hard. In light of this, we live with the timing misalignment among the stations and focus on enabling asynchronous MU-MIMO detection. To this end, we revisit the conventional (synchronous) MMSE detector and find that a transformation can make it applicable to decoding asynchronous data packets from independent stations.

B. System Diagram

Fig. 2 shows the system diagram of UMMC. The AP measures the performance of a sequence of analog beams $[\vec{\theta}_1, \vec{\theta}_2, \cdots, \vec{\theta}_t, \cdots, \vec{\theta}_T]$, where t is the evaluation/iteration index and T is the predefined maximum number of evaluations/iterations allowed (e.g., $T = 30$). In the end of T iterations, UMMC chooses the beam that yields the best performance. In each iteration t, the operations of UMMC include the following four steps:

- Step ❶: The AP selects a beam $\vec{\theta}_t$ for evaluation in the current iteration based on the posterior probability derived from the past evaluations, i.e., $(\vec{\theta}_{t'}, f(\vec{\theta}_{t'}))$ for $1 \le t' < t$. Details are presented in Section IV.
- Step ❷: The AP reconfigures its phased-array antennas by setting their beam patterns to $\vec{\theta}_t$.

- Step ❸: The AP first calculates its digital beamformers (a.k.a. MU-MIMO detector) $\vec{p} = [\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_N]$, and then uses them to decode asynchronous signal frames from the N stations. Details are presented in Section V.
- Step ❹: The AP measures the EVM of the decoded signals from each station. By doing so, it obtains $f(\vec{\theta}_t)$. Then, $(\vec{\theta}_t, f(\vec{\theta}_t))$ is added to the dataset and will be used to guide the future beam search.

IV. BAYESIAN OPTIMIZATION FOR BEAM SEARCH

In this section, we assume that the algorithm for determining $\vec{p}$ is given and focus on the BayOpt design to find a near-optimal beam $\vec{\theta}$ for the AP. The design of $\vec{p}$ will be presented in the next section.

A. Why Bayesian Optimization?

Recall that the objective function is $f(\vec{\theta}) = \max_n \{\text{EVM}_n\}$. It has the following salient features.

- $f(\vec{\theta})$ *has a complex structure:* Fig. 3 shows an example of $f(\vec{\theta})$ obtained through exhaustive beam search on our two-user MIMO 60GHz mmWave testbed[1]. It is evident that $f(\vec{\theta})$ is hard to optimize due to its non-convexity.
- $f(\vec{\theta})$ *is unknown:* Practical mmWave communication systems typically suffer from hardware imperfections such as phase noise and clock jitters [13], which are hard to characterize and model. As such, the beam pattern may largely deviate from its ideal model $G_{\text{ap}}(\theta)$. The accurate objective function $f(\vec{\theta})$ is unknown and can only be obtained via exhaustive experimental measurements.
- *Evaluating $f(\vec{\theta})$ is costly:* To evaluate $f(\vec{\theta})$ for a given $\vec{\theta}$, the AP needs to physically set up the beam pattern and measure the resultant signal quality. This process incurs a fixed airtime overhead. For example, in 802.11ay, measuring the value of $f(\vec{\theta})$ for a given $\vec{\theta}$ may take the time of one Control PHY Preamble (about $3.7\mu s$), let alone other airtime overhead incurred in this process. Therefore, there is a tradeoff between the quality of $\vec{\theta}$ and the number of evaluations of $f(\vec{\theta})$.

Fortunately, BayOpt is an effective technique to optimize such a function that is unknown yet expensive to evaluate [12]. It makes use of the laws of probability to combine prior belief with observed data to compute posterior distribution of the objective function. Therefore, we will design a BayOpt framework for analog beam search.

B. A Bayesian Optimization Framework

To perform BayOpt, one needs to address two problems: i) finding a statistical process to model the function being optimized, and ii) selecting an acquisition function as a surrogate approximation to guide the search in each iteration. In what follows, we address these two problems in order.

Gaussian Process Regression. We model the iterative beam search problem as a Gaussian process. In the tth iteration, the AP has observed $t-1$ beams. Denote $\mathbf{\Theta} = \{\vec{\theta}_i\}_{i=1}^{t-1}$ as the set of beams that the AP has already observed. Denote $f(\mathbf{\Theta}) = \{f(\vec{\theta}_i)\}_{i=1}^{t-1}$ as the objective function values of those observed beams. We treat $f(\mathbf{\Theta})$ as a multi-variate Gaussian

[1]The detailed experimental setup is presented in Section VI-A.
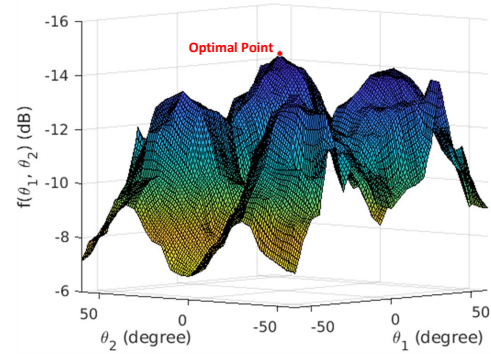

Fig. 3: An instance of $f(\vec{\theta})$ obtained from experimental measurements on a two-user MIMO 60GHz testbed, where $\vec{\theta} = [\theta_1, \theta_2]$ and $f(\vec{\theta}) = \max(\text{EVM}_1, \text{EVM}_2)$ in dB.

distribution, with $\mu(\mathbf{\Theta})$ as its mean and $k(\mathbf{\Theta}, \mathbf{\Theta})$ as its covariance kernel. Here, $\mu(\mathbf{\Theta})$ is a $(t-1) \times 1$ vector, while $k(\mathbf{\Theta}, \mathbf{\Theta})$ is a $(t-1) \times (t-1)$ matrix. Let $\vec{\theta}$ be an arbitrary beam in the beambook. Then, per the definition of Gaussian process, the joint distribution of the function values corresponding to $\vec{\theta}$ and $\mathbf{\Theta}$ should satisfy:

$$\begin{bmatrix} f(\mathbf{\Theta}) \\ f(\vec{\theta}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu(\mathbf{\Theta}) \\ \mu(\vec{\theta}) \end{bmatrix}, \begin{bmatrix} k(\mathbf{\Theta}, \mathbf{\Theta}) & k(\mathbf{\Theta}, \vec{\theta}) \\ k(\vec{\theta}, \mathbf{\Theta}) & k(\vec{\theta}, \vec{\theta}) \end{bmatrix} \right), \quad (4)$$

where $\mu(\cdot)$ and $k(\cdot, \cdot)$ should be understood as an element-wise operational function. There are various definitions for Gaussian kernel, such as *Matérn* kernel, *exponentiated quadratic* kernel, and *radial basis function* kernel [14]. In our experiments, we choose radial basis function kernel, $k(\vec{\theta}_i, \vec{\theta}_j) = \exp(-\frac{1}{2\sigma^2}\|\vec{\theta}_i - \vec{\theta}_j\|^2)$, where σ is a hyper-parameter that governs the kernel width. In our experiments, we let $\sigma = 1$.

The posterior distribution on the arbitrary beam $\vec{\theta}$ can be calculated through standard Bayesian rules. Specifically, the distribution of $f(\vec{\theta})$ can be modeled as:

$$f(\vec{\theta}) \sim p\big(f(\vec{\theta})|\vec{\theta}, \mathbf{\Theta}, f(\mathbf{\Theta})\big) = \mathcal{N}\big(\mu(\vec{\theta}), \Sigma(\vec{\theta})\big), \quad (5)$$

where

$$\mu(\vec{\theta}) = k(\vec{\theta}, \mathbf{\Theta})k(\mathbf{\Theta}, \mathbf{\Theta})^{-1}f(\mathbf{\Theta}), \quad (6)$$

$$\Sigma(\vec{\theta}) = k(\vec{\theta}, \vec{\theta}) - k(\vec{\theta}, \mathbf{\Theta})k(\mathbf{\Theta}, \mathbf{\Theta})^{-1}k(\mathbf{\Theta}, \vec{\theta}). \quad (7)$$

Acquisition Function. There are different acquisition functions available for BayOpt problems such as Probability of Improvement (PoI), Expected Improvement (EI), and Gaussian process Upper Confidence Bound (GP-UCB) [14]. We choose EI for two reasons: i) compared to PoI, it has been shown to be better-behaved; and ii) unlike GP-UCB, it does not involve tuning parameters [15]. The acquisition function can be written as:

$$\text{EI}(\vec{\theta}) = \mathbb{E}\big[\max(f(\vec{\theta}) - f(\vec{\theta}^+), 0)\big], \quad (8)$$

where $\vec{\theta}^+$ is the best beam found so far. Under the Gaussian process model, it can be analytically written as follows:

$$\text{EI}(\vec{\theta}) = \big(\mu(\vec{\theta}) - f(\vec{\theta}^+) - \xi\big)\text{CDF}(Z) + \Sigma(\vec{\theta})\text{pdf}(Z), \quad (9)$$

where $Z = \frac{\mu(\vec{\theta}) - f(\vec{\theta}^+) - \xi}{\Sigma(\vec{\theta})}$, $\text{CDF}(\cdot)$ and $\text{pdf}(\cdot)$ are the cu-

mulative distribution function and the probability density function of standard normal distribution, respectively, and ξ is a parameter that determines the amount of exploration during the optimization. A large value of ξ leads to more exploration, while a small value leads to more exploitation. In our experiments, we empirically set ξ to 0.1.

Beam Selection. In the tth iteration, the beam selected for evaluation is obtained by solving the following problem:

$$\vec{\theta}_t = \arg\max_{\vec{\theta} \in \mathcal{B} \setminus \mathbf{\Theta}} \text{EI}(\vec{\theta}), \qquad (10)$$

where $\mathcal{B}$ is the set of all predefined beams and $\mathbf{\Theta}$ is the set of beams that has been evaluated so far. We note that (10) is easy to solve because (9) is a simple, disciplined function.

C. Practical Considerations

There are two challenges associated with the above BayOpt framework when it is applied to beam search. In the following, we first point out the challenges and then present our solutions.

Limited number of evaluations. MmWave systems have a fixed airtime budget for beam search/training, which determines the maximum number of evaluations/iterations that can be performed before data transmission. In practice, given the limited airtime budget for beam search, it is unlikely to find the optimal beam for data transmission. Therefore, the beam search problem is further constrained by the number of evaluations. To address this challenge, we propose a *recenter-and-shrink* (RaS) scheme for the Gaussian process regression. This scheme was inspired by [16]. The basic idea is that, when approaching the evaluation budget, we *recenter* the search space to the current optimal beam and *shrink* the search space. Doing so increases the probability of finding a better beam when we reach the evaluation budget. Following this idea, we modify the acquisition function in (10) to:

$$\vec{\theta}_t = \arg\max_{\vec{\theta} \in \mathcal{B} \setminus \mathbf{\Theta}} \text{EI}(\vec{\theta}) \qquad (11)$$

$$\text{s.t. } \theta_m \in \begin{cases} [-\frac{\pi}{2},\ \frac{\pi}{2}] & \text{if } 1 \le t < T/2 \\ [\theta_m^+ - \frac{\phi_t}{2},\ \theta_m^+ + \frac{\phi_t}{2}] & \text{if } T/2 \le t \le T. \end{cases}$$

where t is the iteration/evaluation index, T is the maximum number of evaluations, $\vec{\theta}^+ = [\theta_m^+]_{m=1}^M$ is the best beam found so far, and ϕ_t is the reduced search range. Empirically, we set $\phi_t = (\frac{3}{2} - \frac{t}{T})\pi$ in our experiments.

Cubic Computational Complexity. The computational complexity of Gaussian process regression is cubic to the number of data samples, i.e., $\mathcal{O}(t^3)$, where t is the number of evaluations that have been performed [14]. Clearly, the computation rapidly increases as the evaluation procedure evolves. To overcome the computation challenge of Gaussian process, a wealth of sparse approximations have been recently suggested, such as the subset of data (SoD) approximation, the subset of regressors (SoR) approximation, the deterministic training conditional (DTC) approximation, and partially independent training conditional (PITC) approximation [17]. *In these methods, a subset of the latent variables are treated exactly while the remaining variables are treated approximately to reduce the computation.* Here, we employ the SoR approximation for the beam search as it demonstrates a good tradeoff between performance and computation (see Tables 8.1 & 8.2 in [17]).

Algorithm 1 Bayesian optimization for analog beam search

1: **Required:** T: the budgeted number of evaluations.
2: **Output:** A beam $\vec{\theta}^*$ in the predefined beambook $\mathcal{B}$ for data packet reception at the AP
3: Initialization $\mathbf{\Theta} = [\vec{0}]$.
4: **for** $t = 1, 2, \cdots, T$ **do**
5:     Calculate $\mathbf{\Phi}$ using (14)
6:     Calculate $\mu(\vec{\theta})$ using (12) and $\Sigma(\vec{\theta})$ using (13)
7:     Construct the surrogate function $\text{EI}(\vec{\theta})$ using (9)
8:     Find the next beam direction $\vec{\theta}_t$ by solving (11)
9:     Add $\vec{\theta}_t$ to $\mathbf{\Theta}$
10: **end for**
11: **return** $\vec{\theta}^* = \arg\min_{\vec{\theta} \in \mathbf{\Theta}} f(\vec{\theta})$.

Denote $\mathbf{\Phi}$ as the subset of training data samples that are selected for exact regression, where $\mathbf{\Phi} \subset \mathbf{\Theta}$. Per [17], the Gaussian process regression can be characterized by the approximate mean and covariance as follows:

$$\mu(\vec{\theta}) = \sigma^{-2} k(\vec{\theta}, \mathbf{\Phi}) \mathbf{Q}^{-1} k(\mathbf{\Phi}, \mathbf{\Theta}) f(\mathbf{\Theta}), \qquad (12)$$

$$\Sigma(\vec{\theta}) = k(\vec{\theta}, \mathbf{\Phi}) \mathbf{Q}^{-1} k(\mathbf{\Phi}, \vec{\theta}), \qquad (13)$$

where $\mathbf{Q} = \sigma^{-2} k(\mathbf{\Phi}, \mathbf{\Theta}) k(\mathbf{\Theta}, \mathbf{\Phi}) + k(\mathbf{\Phi}, \mathbf{\Phi})$.

A question to ask is how to select the active data samples for $\mathbf{\Phi}$. Empirically, we define an integer number $\tau \in \mathbb{Z}$ which is smaller than t. We choose the τ beams in $\mathbf{\Theta}$ that are closest to $\vec{\theta}^+$ as the active samples for $\mathbf{\Phi}$. Denote $g(\vec{\theta}) \triangleq \|\vec{\theta}^+ - \vec{\theta}\|^2$ as the metric for $\vec{\theta}$. Based on this metric, we sort the elements in $\mathbf{\Theta}$ in a non-decreasing order and denote the resulting vector as $\mathbf{\Theta}_{srt} = [\vec{\theta}_{s_1}, \vec{\theta}_{s_2}, \cdots, \vec{\theta}_{s_t}]$. Then, we let:

$$\mathbf{\Phi} = [\vec{\theta}_{s_1}, \vec{\theta}_{s_2}, \cdots, \vec{\theta}_{s_\tau}]. \qquad (14)$$

With the approximation in (12)-(14), the computational complexity of Gaussian process regression in the tth iteration decreases to $\mathcal{O}(\tau^2 t)$ [14]. More importantly, the complexity scales linearly (rather than cubically) with the number of iterations.

We present the proposed BayOpt algorithm in Alg. 1. In a nutshell, it is a non-parametric online learning algorithm that guides the beam search using the posterior probability of those data samples that have been evaluated so far.

V. ASYNCHRONOUS MU-MIMO DETECTION

In this section, we first review the MMSE MU-MIMO detector, and then present a transformation for MMSE MU-MIMO detector so that it can decode asynchronous data packets. The resulting detector fundamentally relaxes the inter-user synchronization for uplink MU-MIMO, and thus is particularly suited for mmWave communications. Finally, we conduct performance analysis of the proposed detector in mmWave networks.

A. Conventional (Synchronous) MMSE MU-MIMO Detector

Consider the uplink MU-MIMO transmission from N stations to an M-antenna AP as shown in Fig. 1. Suppose that data packets from the N stations are perfectly aligned in time
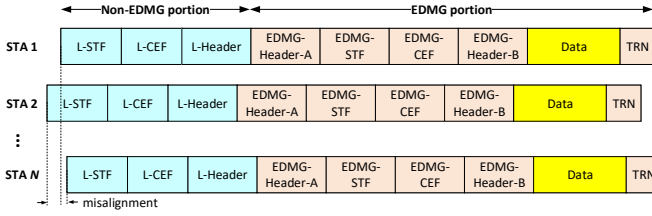
Fig. 4: An illustration of the received asynchronous packets from multiple stations at the AP in an 802.11ay WLAN.

when impinging on the AP. Then, the signal transfer model in the digital domain can be written as:

$$\vec{y} = \mathbf{H}\vec{x} + \vec{w}, \tag{15}$$

where $\vec{y} \in \mathbb{C}^{M \times 1}$ is the received digital baseband signal vector at the AP, $\vec{x} = [x_1, x_2, \cdots, x_N]^\top$ is the transmit signal vector, where x_n is the signal at the nth station, $\vec{w} \in \mathcal{C}^{M \times 1}$ is the noise vector, and $\mathbf{H} = [H_{mn}]_{1 \le m \le M, 1 \le n \le N} \in \mathbb{C}^{M \times N}$ is the compound channel between the N stations and the AP.

To decode the N data packets, the AP can first estimate the compound channel using the orthogonal pilots (a.k.a. reference signals) in the N data packets and then construct the MMSE MIMO detector as follows:

$$\mathbf{P} = \mathbf{H}^{\mathsf{H}}(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{\sigma_w^2}{\sigma_x^2}\mathbf{I})^{-1}, \tag{16}$$

where $\mathbf{I}$ is an identity matrix of proper dimension, σ_x^2 is signal power, and σ_w^2 is noise power. After constructing the MMSE detector, the AP can perform MU-MIMO detection as follows: $\hat{\vec{x}} = \mathbf{P}\vec{y}$, where $\hat{\vec{x}}$ is an estimated copy of $\vec{x}$.

Conventional MU-MIMO detectors can work only when the data packets are well aligned in time. Roughly speaking, the time misalignment of data packets must be less than the time duration of an OFDM symbol's cyclic prefix [18]. For example, in 802.11ay, the time misalignment must be less than 36.36ns [2]. In real systems, this requirement is extremely hard to satisfy as many factors (e.g., propagation delays, digital processing delays, and clock jitters) contribute to the time misalignment. For this reason, uplink MU-MIMO is not standardized in IEEE 802.11ac (sub-6GHz) [19] and 802.11ay (60GHz) [2].[2]

B. A Transformation of MMSE MU-MIMO Detector

Since it is hard to maintain the time alignment of the data packets for the AP, we wish to design an MIMO detector for the AP so that it can decode the misaligned data packets as shown in Fig. 4. In this case, if the AP knows the MMSE MIMO detector $\mathbf{P}$ in (16), then it should still be able to decode those asynchronous data packets. This is because $\mathbf{P}$ is a *spatial* filter and its effectiveness is not affected by the *temporal* imperfections (i.e., time misalignment of data packets). In other words, the spatial and temporal properties of data packets are orthogonal to each other. The key question here is how to obtain $\mathbf{P}$ when the AP receives asynchronous data packets. In synchronous MU-MIMO, the data packets from different stations carry orthogonal pilots for the AP to estimate the

---

channel matrix $\mathbf{H}$, based on which the AP can calculate $\mathbf{P}$ using (16). In asynchronous MU-MIMO, the data packets from different stations cannot maintain the orthogonality of their pilots. As a result, the AP cannot estimate the channel $\mathbf{H}$ and thus (16) does not work for this case.

To overcome this challenge, we show that a transformation of the MMSE detector in (16) can eliminate the need of channel knowledge $\mathbf{H}$ and obtain an approximation of $\mathbf{P}$, which allows the AP to decode those asynchronous data packets separately. Denote $\mathcal{R}_n\{\cdot\}$ as the nth row of a matrix or a vector. Per the conventional MMSE detection, we have

$$\hat{x}_n = \mathcal{R}_n\{\hat{\vec{x}}\} = \mathcal{R}\{\mathbf{P}\vec{y}\} = \mathcal{R}_n\{\mathbf{P}\}\vec{y}. \tag{17}$$

Denote $\mathbf{R}_x$ as the correlation matrix of $\vec{x}$, i.e., $\mathbf{R}_x = \mathbb{E}[\vec{x}\vec{x}^{\mathsf{H}}]$. Denote $\mathbf{R}_w$ as the correlation matrix of $\vec{w}$, i.e., $\mathbf{R}_w = \mathbb{E}[\vec{w}\vec{w}^{\mathsf{H}}]$. In practice, signal and noise are always independent. Then, we have $\mathbf{R}_x = \sigma_x^2\mathbf{I}$ and $\mathbf{R}_w = \sigma_w^2\mathbf{I}$. Per (16), we have

$$\mathcal{R}_n\{\mathbf{P}\} = \mathcal{R}_n\{\mathbf{H}^{\mathsf{H}}(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \frac{\sigma_w^2}{\sigma_x^2}\mathbf{I})^{-1}\}$$

$$\overset{(a)}{=} \mathcal{R}_n\{\mathbf{R}_x\mathbf{H}^{\mathsf{H}}(\mathbf{H}\mathbf{R}_x\mathbf{H}^{\mathsf{H}} + \mathbf{R}_w)^{-1}\}$$

$$= \mathcal{R}_n\{\mathbb{E}[\vec{x}\vec{x}^{\mathsf{H}}]\mathbf{H}^{\mathsf{H}}(\mathbf{H}\mathbb{E}[\vec{x}\vec{x}^{\mathsf{H}}]\mathbf{H}^{\mathsf{H}} + \mathbb{E}[\vec{w}\vec{w}^{\mathsf{H}}])^{-1}\}$$

$$= \mathbb{E}[\mathcal{R}_n\{\vec{x}\}\vec{x}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}]\mathbb{E}[\mathbf{H}\vec{x}\vec{x}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}} + \vec{w}\vec{w}^{\mathsf{H}}]^{-1}$$

$$= \mathbb{E}[x_n(\mathbf{H}\vec{x})^{\mathsf{H}}]\mathbb{E}[(\mathbf{H}\vec{x} + \vec{w})(\mathbf{H}\vec{x} + \vec{w})^{\mathsf{H}}]^{-1}$$

$$\overset{(b)}{=} \mathbb{E}[x_n(\mathbf{H}\vec{x} + \vec{w})^{\mathsf{H}}]\mathbb{E}[(\mathbf{H}\vec{x} + \vec{w})(\mathbf{H}\vec{x} + \vec{w})^{\mathsf{H}}]^{-1}$$

$$= \mathbb{E}[x_n\vec{y}^{\mathsf{H}}]\mathbb{E}[\vec{y}\vec{y}^{\mathsf{H}}]^{-1}, \tag{18}$$

where (a) and (b) follow from the assumptions that $\mathbf{R}_x$ is of full rank and $\mathbb{E}[x_n\vec{w}] = 0$, respectively. Both assumptions are always valid in practice.

Eq. (18) shows that the MMSE detector can be computed without channel knowledge, but using $\mathbb{E}[x_n\vec{y}^{\mathsf{H}}]$ and $\mathbb{E}[\vec{y}\vec{y}^{\mathsf{H}}]$. Now a question to ask is how to compute these two terms. In UMMC, we use the sample averaging operation to approach statistic expectation based on the fact that every packet in practical systems carries reference signals (a.k.a., pilots or preamble) for signal detection. Consider the 802.11ay frame shown in Fig. 4 for example. The reference signals include L-STF, L-CEF, EDMG-STF, and EDMG-CEF, which are pre-defined and known to all stations and APs. These reference signals will be used to compute $\mathbb{E}[x_n\vec{y}^{\mathsf{H}}]$ and $\mathbb{E}[\vec{y}\vec{y}^{\mathsf{H}}]$ in (18).

In the following, we slightly abuse the notation by introducing l as the index of OFDM symbol and k as the index of OFDM subcarrier. Denote $\mathcal{A}_n(k)$ as the set of reference symbols (pilots) in the data packet transmitted by station n on OFDM subcarrier k. Then, we have

$$\vec{p}_n[k] \triangleq \mathcal{R}_n\{\mathbf{P}(k)\} \overset{(18)}{=} \mathbb{E}[x_n\vec{y}^{\mathsf{H}}]\mathbb{E}[\vec{y}\vec{y}^{\mathsf{H}}]^{-1} \tag{19}$$

$$\triangleq \left[\sum_{(l,k')\in\mathcal{A}_n(k)} x_n(l,k')\vec{y}(l,k')^{\mathsf{H}}\right]\left[\sum_{(l,k')\in\mathcal{A}_n(k)} \vec{y}(l,k')\vec{y}(l,k')^{\mathsf{H}}\right]^{\dagger},$$

where $(\cdot)^{\dagger}$ is the pseudo-inverse operator, and $x_n(l,k')$ and $\vec{y}(l,k')$ represent the transmitted and received reference signal on OFDM symbol l and subcarrier k', respectively.

With the MU-MIMO detector in (19), the AP decodes the data packet from station n as follows: $\hat{x}_n(l,k) = \vec{p}_n(k)^\top\vec{y}(l,k)$, where $\vec{y}(l,k)$ is the received payload signal

---

[2]Note that 802.11ax is the only WLAN standard that supports uplink MU-MIMO. Yet, there is still no 802.11ax product that supports this feature. In addition, 802.11ax is more of a centralized rather than distributed network.

vector at the AP and $\hat{x}_n(l,k)$ is its decoded payload signal from station n, $1 \leq n \leq N$.

C. Performance Analysis and Discussions

Performance Analysis. Since analyzing the performance of the proposed detector in general settings is extremely hard, we focus on an ideal case. Suppose that the number of reference signals (e.g., pilots in L-STF, L-CEF, EDMG-STF, and EDMG-CEF in Fig. 4) is greater than or equal to the number of stations, i.e., $|\mathcal{A}_n(k)| \geq N$. Then, we have the following lemma.

Lemma 1: If $M \geq N$ and $\sigma_w = 0$, the MU-MIMO detector in (19) *can perfectly recover the misaligned signals from the asynchronous stations, i.e., $\hat{x}_n(l,k) = x_n(l,k)$ for $1 \leq n \leq N$, $1 \leq k \leq K$, and $1 \leq l \leq L$.*

Proof Sketch. We omit the subcarrier index k to simplify the notation. Given that $M \geq N$, $\mathbf{H}$ is a square or tall/thin matrix. Then, based on (19), we have:

$$\vec{p}_n \overset{(a)}{=} \Big[\sum_{l \in \mathcal{A}_n} x_n(l)\vec{y}(l)^{\mathsf{H}} \Big] \Big[\sum_{l \in \mathcal{A}_n} \vec{y}(l)\vec{y}(l)^{\mathsf{H}} \Big]^{\dagger}$$

$$\overset{(b)}{=} \Big[\sum_{l \in \mathcal{A}_n} x_n(l)\vec{x}(l)^{\mathsf{H}}\mathbf{H}^{\mathsf{H}} \Big] \Big[\sum_{l \in \mathcal{A}_n} \mathbf{H}\vec{x}(l)\vec{x}(l)^{\mathsf{H}}\mathbf{H}^{\mathsf{H}} \Big]^{\dagger}$$

$$\overset{(c)}{=} \Big[\mathcal{R}_n\{\hat{\mathbf{R}}_{\mathbf{x}}\}\mathbf{H}^{\mathsf{H}} \Big] \Big[\mathbf{H}\hat{\mathbf{R}}_{\mathbf{x}}\mathbf{H}^{\mathsf{H}} \Big]^{\dagger} \overset{(d)}{=} \mathcal{R}_n\{\mathbf{H}^{\dagger}\}, \qquad (20)$$

where (a) follows from (19) by omitting the subcarrier index k; (b) follows from the fact that $\vec{y} = \mathbf{H}\vec{x}$ when $\sigma_w = 0$; (c) follows from our definition that $\hat{\mathbf{R}}_x = \sum_{l \in \mathcal{A}_n} \vec{x}(l)\vec{x}(l)^{\mathsf{H}}$; (d) follows from the fact that $\mathbf{H}$ is a square or tall matrix since $M \geq N$ and that $\hat{\mathbf{R}}_x$ is a square matrix of full rank since $|\mathcal{A}_n(k)| \geq N$. Based on (20), we have $\hat{x}_n(l,k) = \vec{p}_n(k)\vec{y}(l,k) = \mathcal{R}_n\{\mathbf{H}(k)^{\dagger}\}\vec{y}(l,k) = x_n(l,k)$. ∎

In practice, the assumption of $|\mathcal{A}_n(k)| \geq N$ and $M \geq N$ are typically valid, but $\sigma_w \neq 0$. For the realistic case, we will evaluate this detector through experiments in Section VI-A.

Explicit Channel Knowledge is Not Needed. It is evident that the MU-MIMO detector in (19) does not require explicit channel knowledge $\mathbf{H}$ for packet detection. Instead, it uses the reference signals in data packets to compute the detectors for each individual data stream. As such, this MU-MIMO detector is particularly suitable for an AP to decode asynchronous data packets, while ZF/MMSE detector is incapable of doing so.

Unique Features of mmWave MU-MIMO. MmWave communication systems are typically equipped with directional antennas (e.g., phased-array antenna), which significantly reduce the multipath effect of channels. As a result, the mmWave channels are more frequency-flat compared to sub-6GHz systems. In addition, compared to SISO mmWave WLANs (e.g., 802.11ad), MU-MIMO mmWave networks (e.g., 802.11ay) have pilots in both legacy preamble (L-STF and L-CEF) and enhanced preamble (EDMG-STF and EDMG-CEF); see Fig. 4. Lemma 1 shows that these two properties make the proposed asynchronous MMSE detector particularly suitable for 802.11ay networks.

VI. PERFORMANCE EVALUATION

A. Experimental Results (Two-User MIMO Case)

Implementation: We built a 60 GHz MU-MIMO testbed that comprises an AP and two stations as shown in Fig. 5. The AP
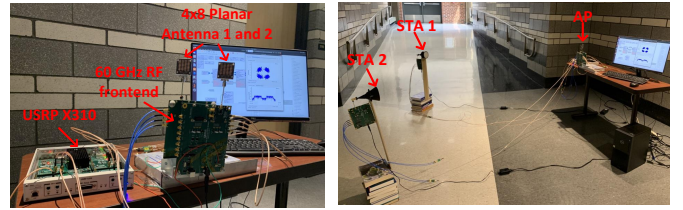
(a) Two-antenna AP.　　　　(b) Experimental setup.

Fig. 5: Illustration of our prototype and experimental setup.
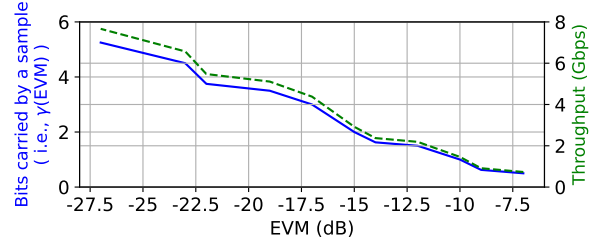

Fig. 6: EVM specified in IEEE 802.11ay standard [2].

was built using two HMC6300 Boards (60 GHz RF Frontend) and one USRP X310. We modified the clock circuits of the two HMC6300 boards to synchronize their clock for MU-MIMO applications. The AP was equipped with two planar antennas, each of which has 4×8 patch elements. Each station was built using one HMC6300 Board and one USRP X310, and connected with a horn antenna. The two stations worked independently, and there is no external clock to synchronize their packet transmissions. The instantaneous bandwidth of this MU-MIMO testbed is 100 MHz. We used GNURadio OTT (in C++) to implement the signal processing modules of a simplified 802.11ay PHY layer (512 FFT for OFDM modulation, QPSK, without LDPC codes) for the uplink MU-MIMO transmission. A demo video can be found in [20].

Experimental Setting. We consider three indoor scenarios for our experiments. Scenario 1: short distance (2.5m) for both stations. Scenario 2: long distance (5m) for both stations. Scenario 3: short distance (2.5m) for station 1 and long distance (5m) for station 2.

Performance Metrics. We use EVM and throughput as the performance metrics. EVM is widely used for the performance measurement of wireless receivers in industry. It was defined in (2). Based on the measured EVM, we calculate the throughput of 802.11ay networks as follows: $r_n = B \cdot \frac{\tau_{ofdm}}{\tau_{gi}+\tau_{ofdm}} \cdot \frac{N_{data}}{N_{fft}} \cdot \gamma(\text{EVM}_n)$, where $B = 2.64$GHz is the sampling rate, $\tau_{gi} = 36.36$ns is the normal guard interval, $\tau_{ofdm} = 194.56$ns is the OFDM symbol duration, $N_{data} = 336$ is the number of subcarriers for data, $N_{fft} = 512$ is the FFT size, and $\gamma(\text{EVM}_n)$ is the adaptive rate specified by [2] and shown in Fig. 6. Recall that our objective is to maximize the minimum of user's throughput. Therefore, we denote $\text{EVM} = \max(\text{EVM}_1, \text{EVM}_2)$ and $\text{Throughput} = \min(r_1, r_2)$.

Asynchronous MU-MIMO Detection. We first validate the feasibility of the proposed asynchronous MU-MIMO detector on the testbed, where the two stations are continuously transmitting data packets but have no synchronization mechanism. For both AP and stations, we perform exhaustive search to find their best analog beams. Fig. 7(a-b) shows the two constellation diagrams observed at the AP. It is clear that the
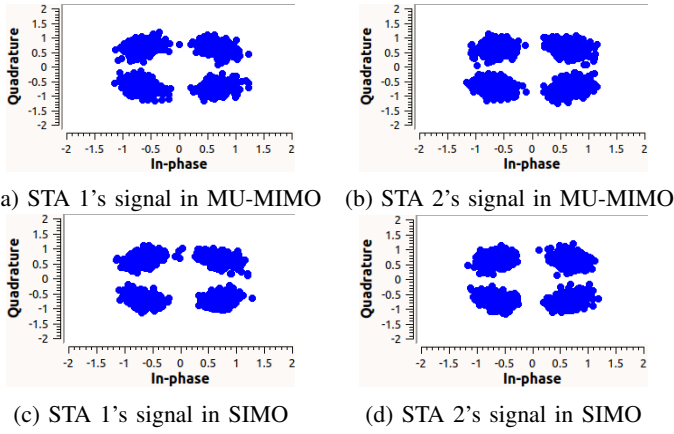
(a) STA 1's signal in MU-MIMO

(b) STA 2's signal in MU-MIMO

(c) STA 1's signal in SIMO

(d) STA 2's signal in SIMO

Fig. 7: Constellation diagram of the decoded signals at the AP.
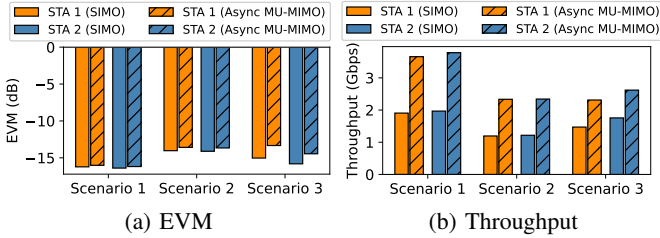

(a) EVM

(b) Throughput

Fig. 8: Comparison of our proposed asynchronous MU-MIMO technique and conventional SIMO technique.

proposed detector is able to decode the data packets in the absence of inter-station synchronization.

As a comparison baseline, we also implemented the single-input and multiple-output (SIMO) transmission scheme on the testbed in the same settings. In this case, each station uses a half of the airtime for packet transmission in turn (i.e., TDMA mode). When serving each station, the AP selects its best antenna to decode its data packets. Fig. 7(c-d) shows the two constellation diagrams observed at the AP. It can be seen that the AP observes similar constellation diagrams in the two cases. This reveals the effectiveness of our proposed MU-MIMO detector in decoding asynchronous data packets.

We repeated the above tests in all three scenarios to quantify the EVM and throughput of the two techniques (Async MU-MIMO and SIMO). Fig. 8(a) shows the EVM comparison. It shows that the two techniques have a similar EVM. This is a bit surprising, because in theory SIMO should offer a better (e.g., 3 dB) EVM performance than Async MU-MIMO. We conjecture that it was caused by the non-negligible phase noise of 60GHz mmWave RF devices. Phase noise increases linearly with carrier frequency in communication systems. When phase noise is strong, it dictates the communication performance and marginalizes the difference caused by other factors.

Fig. 8(b) shows the throughput comparison. It can be seen that Async MU-MIMO almost doubles the throughput of SIMO. This is because the AP can only serve the stations in turn in SIMO, while Async MU-MIMO allows the AP to serve both stations simultaneously.

Impact of MU-MIMO Channel Correlation. For the two antennas at the AP, we consider two approaches for their beam search: i) exhaustive *separate* search, and ii) exhaustive *joint* search. In the separate search, each individual antenna finds

TABLE I: Comparison of exhaustive *separate* beam search and exhaustive *joint* beam search.

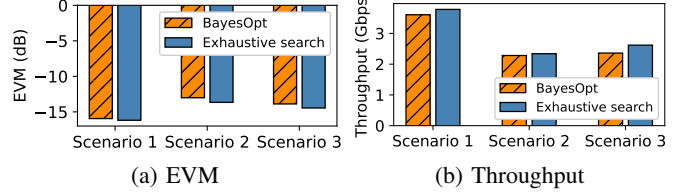| Search approach	Scenario 1		Scenario 2		Scenario 3	
	Joint	Separate	Joint	Separate	Joint	Separate
Best angle for ant 1 (θ_1^*)	-30°	-45°	-30°	-13°	30°	-47°
Best angle for ant 2 (θ_2^*)	15°	-23°	-30°	-24°	30°	25°
EVM (dB)	-16.2	-13.0	-13.7	-10.0	-14.5	-10.1
Throughput (Gbps)	3.65	2.28	2.28	0.91	2.37	1.46

(a) EVM

(b) Throughput

Fig. 9: Comparison of BayOpt and exhaustive search.

the beam angle that maximizes its signal strength. In the joint search, the two antennas try all possible beam combinations to find the one that maximizes the bottleneck of user data rates.

Table I shows our experimental results. It is clear that joint and separate search approaches lead to different beam results. Consider scenario 1 for example. When using separate beam search, the optimal angle is -45° for antenna 1 and -23° for antenna 2. This combination is optimal in terms of the signal strength at each individual antenna, but it is not optimal in terms of user throughput. For the joint search approach, the combination (-30°, 15°) yields the best EVM and thus the best user throughput. Similar phenomena can also be observed in scenarios 2 and 3. This confirms that signal strength is not a good criterion for beam search in MU-MIMO systems.

BayOpt Search versus Exhaustive Search. Using the proposed MU-MIMO detector, we compare two joint beam search approaches: *BayOpt search* and *exhaustive search*. For exhaustive search, we search the beams for each antenna every 5 degrees, and the search range is from -60° to 60°. So the total number of beam combinations for search is $(120/5 + 1)^2 = 625$. Fig. 3 shows an instance of exhaustive search results. For BayOpt search, we fix the number of search iterations (evaluations) to 20. Therefore, the overhead of BayOpt search is only 3.2% of the exhaustive search.

Fig. 9 shows the comparison of these two joint beam search approaches in three scenarios. It can be seen that BayOpt can achieve a similar EVM and throughput performance of exhaustive search. More accurately, BayOpt achieves 94.3% throughput of exhaustive search. It is important to point out that the throughput in Fig. 9(b) does not take into account the airtime overhead of beam training. If the beam training overhead is taken into consideration, BayOpt would easily outperform exhaustive search.

B. Simulation Results (More-User MIMO Case)

Due to the hardware limitation, we resort to simulations for the evaluation of BayOpt in more-user MIMO cases. We consider a 400ft^2 conference room where the AP is deployed on a wall and 100 users are uniformly and randomly distributed over the whole room. We use the model in [21] to calculate the path loss based on the distance between a user and the AP, and use the model in [22] to generate the gain of phased-array antennas for a given direction. In each time slot, the AP randomly selects N users for uplink MU-MIMO
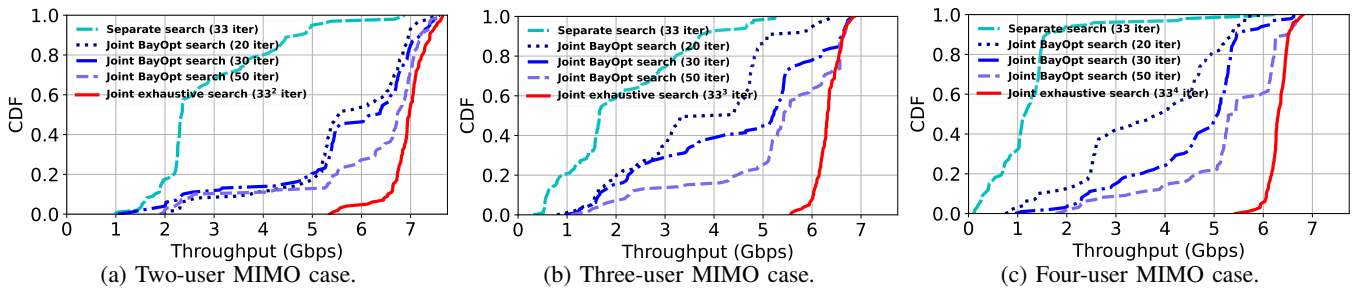
Fig. 10: Throughput comparison of Separate beam search, BayOpt beam search, and exhaustive search.

TABLE II: Representative work on beam search in literature.

Beam search technique	Approach
Learning-based search: [26], [27], [28], [29]	Train the machine learning model to predict the current best beam direction.
Out-of-Band assistance: [6], [30], [7], [31], [9], [32]	Utilize the out-of-band information (sub-6GHz, light, camera) to align the beam.
Compressive sensing: [33], [10]	Finding the best alignment beam direction with sparse measurement.
Hierarchical search: [34], [35], [36], [37], [38], [39]	Design a beamforming codebooks and training in a hierarchical way.

transmission, where $N \in \{2, 3, 4\}$ as defined in 802.11ay. In the simulations, we focus on the comparison of three different beam search approaches (separate exhaustive search, joint exhaustive search, and BayOpt) without considering packet misalignment issue. An ideal MMSE detector is used to decode concurrent packets and calculate their EVM and throughput.

We present the simulation results in Fig. 10. Compared to separate exhaustive search, BayOpt-30 (BayOpt with 30 iterations) has a similar airtime overhead (30 vs. 33 iterations), but it improves the throughput by 95.8%, 109.8%, and 267.2% in the two-user, three-user, and four-user cases, respectively. Compared to joint exhaustive search, BayOpt-50 achieves 88.6% throughput with 4.6% overhead in two-user MIMO case, 82.1% throughput with 0.1% overhead in three-user MIMO case, and 83.5% throughput with 0.004% overhead in four-user MIMO case. Note that the throughput presented in Fig. 10 does not take into account beam search overhead.

VII. RELATED WORK

This work is relevant to mmWave MIMO communications, beam search, MU-MIMO detection, and system prototyping.

802.11ad/ay and Cellular Networks. In 2012, the IEEE 802.11ad amendment standardized communication in 60 GHz unlicensed band to offer up to 6.75 Gbps data rate in a short range [11]. While 802.11ad devices may have multiple antennas, they do not support MU-MIMO transmission. As a follow-up, 802.11ay was standardized in 2020 [2], which supports new features including channel bonding, higher-order modulation, and downlink MU-MIMO. However, it does not support uplink MU-MIMO yet. The 3GPP specification for 5G cellular networks has already supported MU-MIMO, hybrid beamforming, and mmWave communications in the 24–53 GHz band [23]. While abundant literature has studied the beam design and MU-MIMO for mmWave, most of them are limited to signal processing and numerical analysis [24], [25].

Beam Search. There is a large body of work on beam training for mmWave communications. Table II lists some of the representative works and their basic ideas. Of existing

work, most focuses on finding the best beam in a predefined beambook to maximize *signal strength* while minimizing the associated cost. As we explained before, maximizing the signal strength is not a good strategy for MU-MIMO.

Uplink MU-MIMO in Sub-6GHz Networks. Uplink MU-MIMO has been supported in 4G cellular networks and will be supported by 5G and beyond [40]. In contrast, the way of uplink MU-MIMO to 802.11 standards was rocky. Thus far, no on-market WiFi devices support uplink MU-MIMO. Similar to 802.11ay, 802.11ac supports downlink MU-MIMO but does not support uplink MU-MIMO [19]. This can be attributed to the fact that WLANs are distributed, contention-based systems and lack inter-user coordination. Although 802.11ax will support uplink MU-MIMO, the symbol-level synchronization remains an outstanding challenge [41].

Inter-User Synchronization for Uplink MU-MIMO. Timing advance (TA) is the main mechanism used in wireless networks to compensate inter-user time misalignment and offset the signal propagation delays for uplink MU-MIMO and other multi-access technologies. Per [42] and [43], the timing error achieved by TA in cellular networks cannot meet the requirement of mmWave MU-MIMO based on the 802.11ay numerology. [44] validated the throughput gain of MU-MIMO via offline experiments but did not address the timing problem.

VIII. CONCLUSION

In this paper, we presented a practical yet efficient uplink MU-MIMO communication (UMMC) scheme for mmWave networks. This scheme has two key components: BayOpt for beam search and asynchronous MU-MIMO detection. UMMC provides the first BayOpt framework for beam search in mmWave MU-MIMO systems, and introduced a new MU-MIMO detector that can decode asynchronous data packets from multiple users. It has demonstrated through both theory and experiments that fine-grained inter-user synchronization is not needed for uplink MU-MIMO transmission. We have evaluated the performance of UMMC through a blend of experiments and simulations. Experimental and simulation results confirm the practicality and efficiency of UMMC.

REFERENCES

[1] Y. Ghasempour, C. R. Da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11 ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 186–192, 2017.

[2] "IEEE 802.11ay-2021." https://standards.ieee.org/ieee/802.11ay/6142/, Accessed: 09-July-2022.

[3] "Mobile data traffic outlook." https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast, Accessed: 08-July-2022.

[4] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Transactions on Signal processing*, vol. 63, no. 14, pp. 3628–3640, 2015.

[5] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 949–973, 2015.

[6] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-Wave beam steering without in-band measurement," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, pp. 2416–2424, IEEE, 2015.

[7] S. Sur, I. Pefkianakis, X. Zhang, and K.-H. Kim, "WiFi-assisted 60 GHz wireless networks," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, pp. 28–41, 2017.

[8] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra, "Beamspy: Enabling robust 60 GHz links under blockage," in *Proceedings of 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pp. 193–206, 2016.

[9] M. K. Haider, Y. Ghasempour, D. Koutsonikolas, and E. W. Knightly, "Listeer: mmwave beam acquisition and steering by tracking indicator LEDs on wireless APs," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pp. 273–288, 2018.

[10] H. Hassanieh, O. Abari, M. Rodriguez, M. Abdelghany, D. Katabi, and P. Indyk, "Fast millimeter wave beam alignment," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, pp. 432–445, 2018.

[11] "IEEE 802.11ad-2012." https://standards.ieee.org/ieee/802.11ad/4527/, Accessed: 08-July-2022.

[12] J. Mockus, *Bayesian approach to global optimization: theory and applications*, vol. 37. Springer Science & Business Media, 2012.

[13] X. Yang, M. Matthaiou, J. Yang, C.-K. Wen, F. Gao, and S. Jin, "Hardware-constrained millimeter-wave systems for 5G: challenges, opportunities, and solutions," *IEEE Communications Magazine*, vol. 57, no. 1, pp. 44–50, 2019.

[14] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, vol. 2. MIT press Cambridge, MA, 2006.

[15] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," *Advances in neural information processing systems*, vol. 25, 2012.

[16] N. Stander and K. Craig, "On the robustness of a simple domain reduction scheme for simulation-based optimization," *Engineering Computations*, vol. 19, no. 4, pp. 431–450, 2002.

[17] J. Quinonero-Candela, C. E. Rasmussen, and C. K. Williams, "Approximation methods for gaussian process regression," in *Large-scale kernel machines*, pp. 203–223, MIT Press, 2007.

[18] M. A. Alvarez and U. Spagnolini, "Distributed time and carrier frequency synchronization for dense wireless networks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 4, pp. 683–696, 2018.

[19] "IEEE 802.11ac-2013." https://standards.ieee.org/ieee/802.11ac/4473/, Accessed: 09-July-2022.

[20] Anonymous, "Demo video of real-time uplink MU-MIMO mmWave communication." https://youtu.be/Q2Bk7i6O5mg , Accessed:30-July-2022.

[21] G. R. MacCartney, S. Deng, and T. S. Rappaport, "Indoor office plan environment and layout-based mmWave path loss models for 28 GHz and 73 GHz," in *Proceedings of the IEEE 83rd vehicular technology conference (VTC Spring)*, pp. 1–6, IEEE, 2016.

[22] M. Rebato, J. Park, P. Popovski, E. De Carvalho, and M. Zorzi, "Stochastic geometric coverage analysis in mmWave cellular networks with realistic channel and antenna radiation models," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3736–3752, 2019.

[23] 3GPP, "NR and NG-RAN overall description," 2018.

[24] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[25] D. H. Nguyen, L. B. Le, T. Le-Ngoc, and R. W. Heath, "Hybrid MMSE precoding and combining designs for mmWave multiuser systems," *IEEE Access*, vol. 5, pp. 19167–19181, 2017.

[26] Y. Heng and J. G. Andrews, "Machine learning-assisted beam alignment for mmwave systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1142–1155, 2021.

[27] I. Aykin, B. Akgun, M. Feng, and M. Krunz, "MAMBA: A multi-armed bandit framework for beam tracking in millimeter-wave systems," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*, pp. 1469–1478, IEEE, 2020.

[28] S. Khosravi, H. Shokri-Ghadikolaei, and M. Petrova, "Learning-based handover in mobile millimeter-wave networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 663–674, 2020.

[29] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37328–37348, 2018.

[30] M. Hashemi, C. E. Koksal, and N. B. Shroff, "Out-of-band millimeter wave beamforming and communications to achieve low latency and high energy efficiency in 5G systems," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 875–888, 2017.

[31] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beamselection using out-of-band spatial information," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038–1052, 2017.

[32] B. Salehi, M. Belgiovine, S. G. Sanchez, J. Dy, S. Ioannidis, and K. Chowdhury, "Machine learning on camera images for fast mmWave beamforming," in *Proceedings of IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 338–346, IEEE, 2020.

[33] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive tracking with 1000-element arrays: A framework for multi-Gbps mm wave cellular downlinks," in *Proceedings of 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 690–697, IEEE, 2012.

[34] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du, "On the efficient beamforming training for 60GHz wireless personal area networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 504–515, 2012.

[35] 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.

[36] J. Song, J. Choi, and D. J. Love, "Common codebook millimeter wave beam design: Designing beams for both sounding and communication with uniform planar arrays," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1859–1872, 2017.

[37] W. Yuan, S. M. Armour, and A. Doufexi, "An efficient and low-complexity beam training technique for mmWave communication," in *Proceedings of IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 303–308, IEEE, 2015.

[38] Y. M. Tsang, A. S. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmWave communication system," in *Proceedings of IEEE Global Telecommunications Conference-GLOBECOM*, pp. 1–6, IEEE, 2011.

[39] N. Jha, A. Mishra, J. Bapat, and D. Das, "Fast beam search with two-level phased array in millimeter-wave massive MIMO: A hierarchical approach," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1371–1376, IEEE, 2022.

[40] "5G massive MIMO." https://res-www.zte.com.cn/mediares/zte/Files/PDF/white_book/202009101153.pdf, Accessed: 09-July-2022.

[41] Y. Son, S. Kim, S. Byeon, and S. Choi, "Symbol timing synchronization for uplink multi-user transmission in IEEE 802.11ax WLAN," *IEEE access*, vol. 6, pp. 72962–72977, 2018.

[42] 3GPP Tdoc R1-1901252, "Evaluation on TSN requirements," Jan. 2019.

[43] A. Mahmood, M. I. Ashraf, M. Gidlund, J. Torsner, and J. Sachs, "Time synchronization in 5G wireless edge: Requirements and solutions for critical-MTC," *IEEE Communications Magazine*, vol. 57, no. 12, pp. 45–51, 2019.

[44] R. Zhao, T. Woodford, T. Wei, K. Qian, and X. Zhang, "M-cube: A millimeter-wave massive MIMO software radio," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pp. 1–14, 2020.