

Mobile Millimeter Wave Channel Tracking: A Bayesian Beamforming Framework Against DOA Uncertainty

Yan Yang*, Shuping Dang[†], Miaowen Wen[‡], Shahid Mumtaz[§], and Mohsen Guizani[¶]

*State Key Lab. of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

[†]Computer, Electrical and Mathematical Science and Engineering Division,

King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

[‡]School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China

[§]Instituto de Telecomunicações, Campus Universitário de Santiago, Aveiro, Portugal

[¶]Department of Computer Science, University of Idaho, Moscow ID, USA

*yyang@bjtu.edu.cn; [†]shuping.dang@kaust.edu.sa; [‡]emwwen@scut.edu.cn; [§]smumtaz@av.it.pt; [¶]mguizani@ieee.org

Abstract—A Bayesian approach for joint beamforming and tracking is presented, which is robust to uncertain direction-of-arrival (DOA) estimation in millimeter wave (mmWave) multiple input multiple output (MIMO) systems. The uncertain or completely unknown DOA is modeled as a discrete random variable with a priori distribution defined over a set of candidate DOAs, which describes the level of uncertainty. The estimation problem of DOA is formulated as a weighted sum of previously observed DOA values, where the weights are chosen according to a posteriori probability density function (pdf) of the DOA. In particular, we present a motion trajectory-based a priori probability approximation method, which implies a high probability to perform a directional estimate within a specific spatial region. We demonstrate that the proposed approach is robust to DOA uncertainty, and the beam tracking problem can be addressed by incorporating the Bayesian approach with an expectationmaximization (EM) algorithm. Simulation results validate the theoretical analysis and demonstrate the effectiveness of the proposed solution.

Keywords—Bayesian beamforming; direction of arrival (DOA) estimation; millimeter wave, channel tracking; expectationmaximization (EM) algorithm

I. INTRODUCTION

Millimeter wave (mmWave) systems offer a large available bandwidth for Gigabit per second (Gbit/s) data rates in future cellular and wireless LAN systems [1]–[5]. Due to the spatial sparsity and quasi-optical property of mmWave channels, a typical mmWave channel is dominated by a few paths. The large-scale antenna arrays with increasing antennas in limited physical space will lead to a much better resolvability of multiple paths. Accordingly, the problem of mmWave channel estimation can be naturally formulated by the angle-based representation of the channel and thus can be simplified as a straightforward approach to estimate the dominant direction of arrival (DOA) and the corresponding path gains [6].

Mobile mmWave communication systems will likely operate over wideband channels with frequency selectivity. These antenna beams must be precisely aligned to afford a feasible link margin, and an efficient approach

for beam training and tracking is crucial to determine suitable directions of transmission and reception [7]. For a downlink mobile mmWave system, a key consideration is that the received wideband signal is quite sensitive to the error of steering direction, i.e., a slight beam misalignment may lead to a significant drop of reception quality. Therefore, accurate DOA information is essential for both base station (BS) and Mobile stations (MS's). However, precise knowledge of the steering vector is usually difficult to be determined in a fast-varying channel environment. For example, the DOA information is usually inaccurate and even lost due to the effects of multiple positionally indeterminate scatterers, uncalibrated arrays, node mobility, as well as DOA estimation error. In a mobile mmWave network, destructive effect caused by mobility on the DOA accuracy shall be dealt with carefully. To address this issue, many Bayesian frameworks for adaptive and/or blind beamforming algorithms have been developed to combat DOA uncertainty [8]–[10].

In this paper, we investigate the robust beamforming for mobile mmWave channel tracking in a downlink multiple input, multiple output-orthogonal frequency division multiplexing (MIMO-OFDM) system with low-resolution analog-to-digital converters (ADCs). We propose a Bayesian-based parameter estimation method to mitigate the performance degradation subject to the DOA uncertainty. With our proposed Bayesian framework, a statistically inference approach is formulated to address the uncertainty of DOA estimation. By exploiting the characteristics of the motion trajectory of a reference receiver, we propose an approximation method of a priori probability. The advantage of this method is that the calculation of a posteriori probability only needs to be performed within a range where the actual DOA occurrence probability is high.

The organization of the paper is as follows. In Section II, we present the system model and the problem description. In Section III, we mathematically formulate the Bayesian beamformer. A beam tracking method is introduced, discusses key technical considerations in Section IV. Section V present simulation results and Section VI

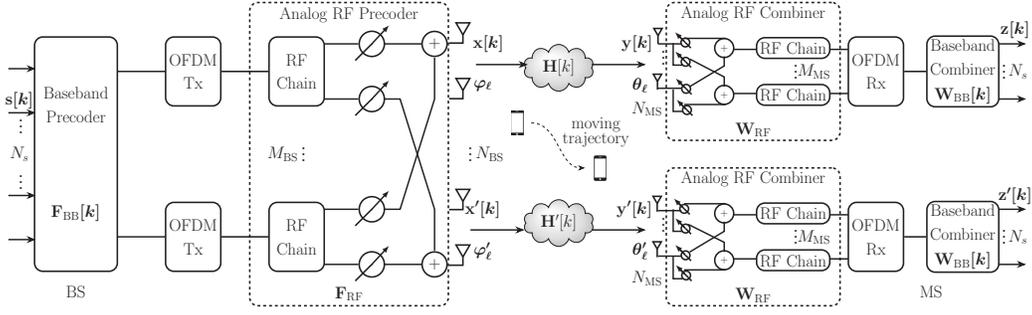


Fig. 1: A downlink mmWave MIMO-OFDM system with hybrid beamforming structure: received signal at subcarrier k .

conclusions of this work.

II. SYSTEM MODEL AND PROBLEM PRELIMINARY

Consider a wideband downlink mmWave MIMO-OFDM system with N_{MS} receive antennas and N_{BS} transmit antennas, which is constructed by a two-stage hybrid digital/analog beamforming architecture as shown in Fig. 1. From the view of implementation in the frequency domain, the transmitted data symbols are organized into frequency vectors $\mathbf{s}[k] \in \mathbb{C}^{N_s \times 1}$, $k = 1, \dots, K$, and the individual data stream is transmitted via N_s length- K data symbol blocks with the property $\mathcal{E}[\mathbf{s}[k]\mathbf{s}[k]^T] = \mathbf{I}$, where $\mathcal{E}(\cdot)$ denotes the expected value of the enclosed argument.

Within the hybrid architecture, a BS with M_{BS} radio frequency (RF) chains serves a MS with M_{MS} RF chains. In this paper, we use \mathbb{C} to denote the field of complex numbers and $\mathbb{C}^{m \times n}$ denotes an m by n dimensional complex space. The BS first applies a digital baseband precoder $\mathbf{F}_{BB}[k] \in \mathbb{C}^{M_{BS} \times N_{BS}}$ to modify the N_s data symbols $\mathbf{s}[k]$ over the k th subcarrier, and is transformed to the time domain using M_{BS} K -point IFFT. After that, an RF (analog) precoder $\mathbf{F}_{RF} \in \mathbb{C}^{N_{BS} \times 1}$, is implemented using analog phase shifters to support multiple directional beamers in downlink multi-user communication. Similarly, the hybrid beamforming structure at the receiver is constructed by concatenation of a RF combiner $\mathbf{W}_{RF}[k] \in \mathbb{C}^{M_{MS} \times N_{MS}}$ and a low-dimensional digital baseband combiner $\mathbf{W}_{BB}[k] \in \mathbb{C}^{M_{BS} \times N_{BS}}$. Since the transceiver can only form a beam in one direction, we mainly concern the path directions of the source and/or reflectors, i.e., (ϕ, θ) , each pair of physical path's angle of arrival and departure (AoA/AoD).

To simplify our exposition, we first adopt an extended Saleh-Valenzuela geometric model, where the matrix channel \mathbf{H} is supposed to be determined by the sum of L multipath clusters, assuming each cluster to be a complex Gaussian variable [11]. With this multipath-cluster channel model, the channel matrix of the k th subcarrier $\mathbf{H}[k]$ can be written as

$$\mathbf{H}[k] = \mathbf{N}_L \sum_{\ell=0}^{L-1} \alpha_{k,\ell} \mathbf{a}_{\theta_\ell} \mathbf{a}_{\varphi_\ell}^H e^{-j \frac{2\pi d_\ell k}{K}}, \quad (1)$$

where $\mathbf{N}_L = \sqrt{\frac{N_{BS} N_{MS}}{L}}$ is a normalized factor; $\alpha_{k,\ell} \sim \mathcal{CN}(0, 1)$ is the complex gain of ℓ th path over the k

subcarrier; $\mathbf{a}_{\theta_\ell} \in \mathbb{C}^{N_{MS} \times 1}$ denotes the receive array steering vector and $\mathbf{a}_{\varphi_\ell} \in \mathbb{C}^{N_{BS} \times 1}$ denotes the transmit array steering vector; $(\cdot)^H$ denote Hermitian transpose. Correspondingly, θ and φ are the AoA and the AoD at the receiver and transmitter, respectively. The array steering vectors \mathbf{a}_ℓ and \mathbf{a}_ℓ corresponding to θ and φ are given as

$$\begin{aligned} \mathbf{a}_{\theta_\ell} &= \frac{1}{\sqrt{N_{MS}}} \left[1, e^{j \cos(\theta_\ell)}, \dots, e^{j(N_{MS}-1)\cos(\theta_\ell)} \right]^T \\ \mathbf{a}_{\varphi_\ell} &= \frac{1}{\sqrt{N_{BS}}} \left[1, e^{j \cos(\varphi_\ell)}, \dots, e^{j(N_{BS}-1)\cos(\varphi_\ell)} \right]^T. \end{aligned} \quad (2)$$

Even though the sparse scattering nature of the mmWave channel, the measurement results reveal that there are $\{R_\ell\}_{\ell=0}^{L-1}$ rays within each of L multipath clusters. Given an acyclic-prefix length D , the discrete-time wideband MIMO channel matrix \mathbf{H} associated with the delay- d can be obtained as [11]

$$\mathbf{H}[d] = \mathbf{P}_{Loss} \sum_{\ell=0}^{L-1} \sum_{r=0}^{R_\ell-1} \xi(dT_s - \tau_{\ell,r}) \mathbf{a}_{\theta_{\ell,r}} \mathbf{a}_{\varphi_{\ell,r}}^H, \quad (3)$$

where $\mathbf{P}_{Loss} = \sqrt{\frac{N_{BS} N_{MS}}{E_{PL}}}$ is a normalized factor dependent on path loss E_{PL} and $\tau_{\ell,r}$ denoting the time delay of the r th ray within cluster ℓ . For such a scattering environment, the channel matrix over subcarrier k can be expressed by

$$\mathbf{H}[k] = \sum_{d=0}^{D-1} \mathbf{H}[d] e^{-j \frac{2\pi d k}{K}}. \quad (4)$$

In the absence of the frequency offset and phase noise, the final transmitted signal associated with the k th subcarrier can be represented as

$$\mathbf{x}[k] = \mathbf{F}_{RF} \mathbf{F}_{BB}[k] \mathbf{s}[k], \quad (5)$$

where $\mathbf{F}_{RF}[k]$ is assumed to be frequency-flat, while the baseband precoders can be different for each subcarrier. To simplify our exposition, we adopt a common approach that each narrowband channel is modeled through block fading in an ideal setting, where a line-of-sight (LOS) path is considered so that the channel can be approximated by the single-tap model. Consequently, the received signal at the antenna array of MS can be written as

$$\mathbf{y}[k] = \mathbf{H}[k] \mathbf{x}[k] + \mathbf{n}[k], \quad (6)$$

where $\mathbf{H}[k] \in \mathbb{C}^{N_{MS} \times N_{BS}}$ is the complex channel matrix corresponding to the k th subcarrier and $\mathbf{n}[k] \sim$

$\mathcal{CN}(0, \sigma^2 \mathbf{I})$ denotes the vector of independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN); $\mathbf{I} \in \mathbb{C}^{N_{\text{MS}} \times 1}$.

At the receiving end, the overall hybrid combiner can be written as $\mathbf{W}[\mathbf{k}] = \mathbf{W}_{\text{RF}}[\mathbf{k}] \mathbf{W}_{\text{BB}}[\mathbf{k}]$, $\mathbf{W}[\mathbf{k}] \in \mathbb{C}^{M \times N_s}$, where $\mathbf{W}_{\text{RF}}[\mathbf{k}]$ is the RF combiner implemented by phase shifters, assuming that $\|\mathbf{W}_{\text{RF}}[\mathbf{k}]\|_2 = 1$, and $\mathbf{W}_{\text{BB}}[\mathbf{k}]$ is the digital combiner in baseband. It means that the estimate of the desired signal vector $\bar{\mathbf{z}}[\mathbf{k}] \in \mathbb{C}^{N_s \times 1}$ can be given as a linear combination of the receive signal $\mathbf{y}[\mathbf{k}]$ over the current subcarrier. We mathematically have

$$\mathbf{z}[\mathbf{k}] = \mathbf{W}^H[\mathbf{k}] \mathbf{H}[\mathbf{k}] \mathbf{F}[\mathbf{k}] \mathbf{s}[\mathbf{k}] + \mathbf{W}^H[\mathbf{k}] \mathbf{n}[\mathbf{k}], \quad (7)$$

where $\mathbf{F}[\mathbf{k}] = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}[\mathbf{k}]$.

The optimal hybrid precoders at the BS can be thereby determined by the following criterion:

$$\begin{aligned} \mathbf{F}^*[\mathbf{k}] &= \arg \max_{\mathbf{F}[\mathbf{k}]} \mathbf{F}^H[\mathbf{k}] \mathbf{s}^H[\mathbf{k}] \mathbf{s}[\mathbf{k}] \mathbf{F}[\mathbf{k}] \\ &\text{s.t.} \sum_{k=1}^K \text{Tr}(\mathbf{V}^H[\mathbf{k}] \mathbf{V}[\mathbf{k}]) < \Omega_{\text{Tx}}, \end{aligned} \quad (8)$$

where $\text{Tr}(\cdot)$ denotes the trace of the matrix. Similarly, by defining the average signal-to-noise ratio (SNR) at each receive antenna as $\rho_k / N_{\text{BS}} \sigma_k$, the optimal combiner at the receiver can be expressed as

$$\begin{aligned} \mathbf{W}^*[\mathbf{k}] &= \arg \max_{\mathbf{W}[\mathbf{k}]} \frac{|\mathbf{W}^H[\mathbf{k}] \mathbf{H}[\mathbf{k}] \mathbf{F}[\mathbf{k}]|^2 |\mathbf{s}[\mathbf{k}]|^2}{|\mathbf{W}^H[\mathbf{k}] \mathbf{n}[\mathbf{k}]|^2} \\ &\text{s.t.} \sum_{k=1}^K |\mathbf{W}[\mathbf{k}]|^2 = 1, \end{aligned} \quad (9)$$

Assume that the total available power is uniformly allocated over all space-frequency grids and the data streams transmitted from different antennas are statistically independent. In this manner, the effects of fading are greatly reduced. From Shannon's information theory, the bound on the capacity of a single-user system is [12]

$$\begin{aligned} C_{\text{SU}} &= \sum_{k=1}^K \beta_k \log_2 \left(1 + \frac{\rho_k}{N_{\text{BS}} \sigma_k^2} \right) \\ &= \sum_{k=1}^K \beta_k \log_2 \left(1 + \frac{|\mathbf{W}^H[\mathbf{k}] \mathbf{H}[\mathbf{k}] \mathbf{F}[\mathbf{k}]|^2}{N_{\text{BS}} \sigma_k^2} \right), \end{aligned} \quad (10)$$

where β_k represents the normalized weight of the rate over the k th subcarrier.

Because, our work mainly focuses on urban LOS and non-LOS (NLOS) channels, a closed-form expression for the desired signal can be expressed as

$$\begin{aligned} \mathbf{z}[\mathbf{k}] &= \underbrace{\mathbf{G}_0 \mathbf{W}^H[\mathbf{k}] \mathbf{a}_{\text{R}_0} \mathbf{a}_{\text{T}_0}^H \mathbf{F}[\mathbf{k}] \mathbf{s}[\mathbf{k}]}_{\text{LOS signal}} \\ &+ \underbrace{\mathbf{W}^H[\mathbf{k}] \tilde{\mathbf{H}}[\mathbf{k}] \mathbf{F}[\mathbf{k}] \mathbf{s}[\mathbf{k}] + \mathbf{W}^H[\mathbf{k}] \mathbf{n}[\mathbf{k}]}_{\mathbf{e}[\mathbf{k}]: \text{effective NLOS signals plus noise}}, \end{aligned} \quad (11)$$

where \mathbf{G}_0 is the antenna gain of LOS. Consider that the global optimization problem is decoupled and the optimal beamformer and combiner can be systematically solved, accordingly, we do not take the global optimization into

account because of its overlarge computational complexity. After some manipulations, the final received signal can be simplified as

$$\mathbf{z}[\mathbf{k}] = \mathbf{W}^H[\mathbf{k}] \mathbf{a}[\theta] \mathbf{a}^H[\phi_0] \mathbf{F}_{\text{opt}}[\mathbf{k}] \mathbf{s}[\mathbf{k}] + \mathbf{e}[\mathbf{k}], \quad (12)$$

where $\mathbf{e}[\mathbf{k}]$ represents the perturbation term and ϕ_0 is a temporary steering angle in the selected beam direction. For the convenience of derivation, we use θ to represent the DOA information set in the following sections which can be viewed as the combination of any azimuth and elevation angle.

III. BAYESIAN BEAMFORMING APPROACH

A. Bayesian Estimation to Uncertain DOA

Let $\mathbf{Z} = \{\mathbf{z}[n, k], \dots, \mathbf{z}[n + m - 1, k]\}$ denotes a collection of m snapshots of the received data vectors. From (12), the k th subcarrier of the n th received signal at the subcarrier level, can be rewritten as

$$\mathbf{z}[n, k] = \mathbf{W}^H[n, k] \mathbf{a}[\theta] \mathbf{a}^H[\phi_0] \mathbf{F}_{\text{opt}}[k] \mathbf{s}[n, k] + \mathbf{e}[n, k], \quad (13)$$

where θ could be viewed as the incidence DOA of the received signal at time slot n .

Using a Bayesian approach, unknown DOA parameter at each sampling moment is generally assumed to be a discrete random variable with a priori $p(\theta)$, and there exists a discrete set of J points $\theta = \{\theta_j\} \forall j \in \{1, \dots, J\}$ over the whole parameter space. From the viewpoint of statistical inference, given observation \mathbf{Z} with a hidden random variable θ , the detection probability for the desired \mathbf{Z} is proportional to the likelihood function $L(\theta|\mathbf{Z})$, which can be written as

$$L(\theta|\mathbf{Z}) = f(\mathbf{Z}|\theta) = \prod_{i=0}^{m-1} f(\mathbf{z}[n + i, k]|\theta), \quad (14)$$

where $f(\mathbf{Z}|\theta)$ is the mass probability function, i.e., the conditional pdf of observations.

By adopting a parametric representation of the Bayesian approach, the estimation problem with uncertainty is to infer θ in terms of $f(\mathbf{Z}|\theta)$ from the noisy observation \mathbf{Z} . It is clear that the azimuth and elevation angles associated with DOA are orthogonal, and θ could be therefore considered as a single variable to reduce the computational complexity. Again, denote the joint pdf of \mathbf{Z} and θ by $g(\mathbf{Z}, \theta)$, we have

$$g(\mathbf{Z}, \theta) = f(\mathbf{Z}|\theta) f(\theta) = f(\theta|\mathbf{Z}) \int_{\theta} f(\mathbf{Z}|\theta) f(\theta) d\theta, \quad (15)$$

Considering the estimation problem, only a posteriori probabilities can be used for inference, which can be calculated by

$$f(\theta|\mathbf{Z}) = \frac{f(\mathbf{Z}|\theta) f(\theta)}{\int_{\theta} f(\mathbf{Z}|\theta) f(\theta) d\theta}. \quad (16)$$

According to Bayes' theorem, our task is to realize statistical inferences about the unknown parameter θ . Suppose that a meaningful prior pdf is a normal distribution with given parameter μ_n and δ_n that can be chosen for the parameter θ , then the posteriori pdf of θ is a normal

distribution and provides increased estimation accuracy.

In the presence of DOA uncertainty, given the noisy observation data set \mathbf{Z} , the Bayesian approach enables us to reconstruct DOA parameter value by using the MAP estimate. With the logarithmic representation, the MAP estimate can be obtained by

$$\hat{\boldsymbol{\theta}}_{MAP} = \arg \max_{\boldsymbol{\theta}} f(\mathbf{Z}|\boldsymbol{\theta})f(\boldsymbol{\theta}). \quad (17)$$

A simple approach, KullbackLeibler (KL) divergence, can be used to define the level of uncertainty, which can be calculated by

$$D(\boldsymbol{\theta}_j; \boldsymbol{\theta}) = \mathcal{E}_j \left\{ \ln \frac{f(\mathbf{Z}|\boldsymbol{\theta}_j)}{f(\mathbf{Z}|\boldsymbol{\theta})} \right\}. \quad (18)$$

Recall from (9) that the weight vectors of combined matrix \mathbf{W} can be determined by solving the optimization problem by certain criteria. In this paper, we follow the approach proposed in [4]. Consider $\hat{\mathbf{Z}}$ with $f(\mathbf{Z}|\boldsymbol{\theta}_j)$, the optimization problem of the Bayesian beamformer $\mathbf{W}_B[\boldsymbol{\theta}; n, k]$ turns out to minimize the conditional minimum mean square error (MMSE) in space Θ , namely:

$$\mathbf{W}_B[\boldsymbol{\theta}; n, k] = \arg \min_{\mathbf{W}_B, \boldsymbol{\theta} \in \Theta} \text{Tr} \left[(\hat{\mathbf{Z}} - \mathbf{Z})(\hat{\mathbf{Z}} - \mathbf{Z})^H \right]. \quad (19)$$

Note that, based directly on Bayes' theorem, the use of Bayesian estimator is equivalent to yielding an estimate of $\hat{\boldsymbol{\theta}}$ as close as possible to the actual value of $\boldsymbol{\theta}$, with respect to the posterior information. The resulting beamformer is a weighted sum of MMSE beamformers pointed at a set of candidate DOA's [8]. To obtain a good estimate of $\hat{\boldsymbol{\theta}}$, it can be seen that a priori pdf needs to be determined, and the precision of the a posteriori pdf is also required.

B. Approximation of Priori and Posteriori pdf

Motivated by the results presented in [6], the angle-delay of the ℓ th path is modeled by the angle domain channel coefficient matrix $\boldsymbol{\Gamma}[\ell] \in \mathbb{C}^{N_{MS} \times N_{BS}}$, in which the (i, j) th entry of the matrix $\boldsymbol{\Gamma}[\ell]$ is the channel gain between the j th discrete transmit angle and the i th discrete receive angle, equivalent to the angle domain channel representation proposed in [6]. Accordingly, the coefficient Γ_i of $\boldsymbol{\Gamma}$ is generated by the following expression:

$$f(\Gamma_i; \boldsymbol{\theta}) = \lambda_0 \delta(\Gamma_i) + \sum_i \lambda_i \mathcal{CN}(\Gamma_i; \mu_i, \Xi_i), \quad (20)$$

where $\lambda_0 = \Pr\{\Gamma_i = 0\}$ and $\lambda_i (0 \leq \lambda_i \leq 1)$, μ_i , Ξ_i are the weights, means, and variances of the GM over $\boldsymbol{\theta}$, respectively; $\delta(\cdot)$ is the Dirac delta function. The parameter λ provides an additional degree of freedom that controls the GM distribution. As the details of the modeling process are out of the scope of this paper, we do not present them here. In our work, we consider that the velocity is not too great, and the angle-delay is assumed to be quasi-static, i.e., $\boldsymbol{\Gamma}, \boldsymbol{\theta}$ are supposed to be invariant within M time slots

Generally, the estimates of a priori pdf in (20) can be determined by a series of iterative quantization algorithms, e.g., MAP and Bayesian linear regression. However, the estimates of (μ_i, Ξ_i) of a priori pdf have to consider all

possible combinations of discrete DOA values in the whole parameter space as well as all possible combinations of a posterior probability estimation. Therefore, once the optimization procedure falls into a wide search range, it would take a great amount of time in order to converge to the optimal solution. Such a pitfall will also lead to the inherent difficulty of computational complexity.

To solve the pitfall above, we propose a moving trajectory-based constraint relying on the motion features that are not completely random. Therefore, we can to some extent utilize the regularity to restrict the search space of DOA estimates. We assume that each moving trajectory corresponds to a specific a priori pdf subset, and the channel information (spatial directions) is collected from multiple specific moving trajectories. With this limitation, the parameter range of the expected a posteriori pdf can be limited to a smaller region, and a Gaussian approximation is often reasonable in reducing the computational complexity, resulting in a faster convergence rate.

The motivation behind this constraint is that the moving trajectory contains valuable a priori information, since the BS is geographically stationary and the moving trajectory of target object is usually regular. Particularly, this regular feature might be more obvious in high-mobility scenarios. Suppose that there exists a \mathcal{K} -element overlapping subset $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_{\mathcal{K}}\}$ in the whole DOA parameter search space. Each element in the overlapping subset could be defined as a subspace and is comprised of a priori pdf corresponding to a specific moving trajectory. This feature implies that $\boldsymbol{\theta}$ can be reconstructed by

$$\boldsymbol{\theta}_1 \cup \boldsymbol{\theta}_2 \dots \cup \boldsymbol{\theta}_{\mathcal{K}} = \boldsymbol{\theta}. \quad (21)$$

Accordingly, the a posteriori pdf of $\boldsymbol{\theta}$, a discrete setting of (16), can be expressed as

$$f(\Gamma_i; \boldsymbol{\theta} | \mathbf{z}) = \frac{f(\mathbf{Z}|\Gamma_i; \boldsymbol{\theta})f(\Gamma_i; \boldsymbol{\theta})}{\sum_{j=1}^J f(\mathbf{Z}|\Gamma_i; \boldsymbol{\theta}_j)f(\Gamma_i; \boldsymbol{\theta})}. \quad (22)$$

The DOA-dependent hyperparameter estimates $\hat{\mu}_i$ and $\hat{\Xi}_i$ are obtained by maximizing the marginal log-likelihood distribution [6]. Mathematically, it can be written as

$$(\hat{\mu}_i, \hat{\Xi}_i) = \arg \max_{\mu_i, \Xi_i} \sum_{j=1}^J (\log f(\mathbf{Z}|\Gamma_i; \mu_i, \Xi_i) + \log f(\Gamma_i; \mu_i, \Xi_i)). \quad (23)$$

Up to this point, the optimization objective becomes

$$\hat{\boldsymbol{\theta}}_{MAP} = \arg \max_{\boldsymbol{\theta} \in \boldsymbol{\theta}_{\mathcal{K}}} (\log f(\mathbf{Z}|\Gamma_i; \boldsymbol{\theta}) + \log f(\Gamma_i; \boldsymbol{\theta})). \quad (24)$$

IV. ROBUST CHANNEL TRACKING

According to a time-variant state-space model, the dynamics of DOA can be described as a first-order Gauss-Markov process by

$$\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)} + \Delta \boldsymbol{\theta} + \xi, \quad (25)$$

where $\xi \sim \mathcal{CN}(0, \sigma_s^2)$ is zero-mean Gaussian random noise term that allows $\boldsymbol{\theta}$ to change over time.

To find the unknown parameters of the a posterior

pdf, we adopt a data-aided EM algorithm, where the transmitted symbol \mathbf{x} is supposed to be known. Typically, the EM algorithm is an iterative method and each iteration cycle consists two steps: the expectation step (E-step) and the maximization step (M-step) [13].

E-step : Define $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})$ as the expected value of the logarithmic a posteriori function of $\boldsymbol{\theta}$. The new parameters are estimated from the recent snapshot \mathbf{Z} by

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i-1)}) = \mathcal{E} \left[\log f(\mathbf{z}|\boldsymbol{\theta}) | \mathbf{z} = \mathbf{Z}, \boldsymbol{\theta}^{(i-1)} \right]. \quad (26)$$

M-step : Iteratively find the parameters according to the criterion given as follows:

$$\boldsymbol{\theta}^{(i)} = \arg \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i-1)}). \quad (27)$$

Note that the EM algorithm starts with an arbitrary initial guess, due to the GM assumption, the parameters is characterized in Section III. In the M-step, the latest parameters are updated. The M-step ensures that each iteration cycle increases the a posteriori until a local maximum is reached, i.e., the algorithm converges to a stationary point of the MAP function. For hyperparameter estimates of GMs, $\hat{\mu}_i$ and $\hat{\Sigma}_i$ can be obtained by (23) in a straightforwardly, and there is only a slightly different in the form of the Q function, as shown in (22) and (23).

We start with traditional beam searching and training procedure for initial access beam training [7], and obtain the a priori pdf. At the transmitter, the BS uses the training precoding matrix of the first level of the codebook. At the receiver, the MS uses the measurement vectors of the first level to combine the received signal. In the iteration step, we assume the MS is moving along a fixed trajectory or a mobility pattern. According to Bayes' theorem, the estimation of a latest set $\hat{\boldsymbol{\theta}}$ of suitable directions and alleviate the DOA uncertainty under mobility. During the iterative EM procedure, we do not need to feed back explicit DOA estimation to the BS, and the training overhead is therefore reduced.

V. NUMERICAL RESULTS

We consider a 64×16 hybrid beamforming architecture with a uniform planar array (UPA), where the arrays at the BS and MS are equipped with $N_{\text{BS}} = 8 \times 8$, $N_{\text{MS}} = 4 \times 4$ antenna elements and 4 RF chains, respectively. To investigate the performance of the proposed approach in close to the real scenarios, our simulation is conducted at 28 GHz frequency bands in a typical urban macro-cellular (UMa) scenario with LOS and NLOS components, which has a bandwidth of 100 MHz.

Fig. 2 compares the proposed Bayesian method and existing data-independent methods in high SNR regime. It demonstrates the significant gains can be achieved with the proposed Bayesian beamforming approach. As can be seen from Fig. 3, the method proposed in [7] may not take into account the effect of array calibration error and does not form the expected gain in a particular direction. In this case, we notice that [14] performs better than [7] but have a direction error of about 5° .

In Fig. 3, the MSE performance of DOA estimation is evaluated at low ADC quantization levels. As described

in Section III, the statistical properties of DOA are the product of the GM distribution, where a sufficiently large number of snapshots could be generally used to obtain the correct statistical properties. Correspondingly, the mean square error of the DOA estimate can be expressed as $\mathcal{E}|\theta - \hat{\theta}|^2$. Fig. 3 shows the MSE performance used for the estimated parameters θ using 1 and 2-bit quantization in the high SNR regime.

Fig. 4 shows the posteriori PDF estimation performance for the unknown parameters θ and ϕ , considering an approximation of a GM distribution with two-component mixed model. The estimates of Gaussian parameters are obtained by using an iterative EM algorithm. Compared with the posteriori pdf estimation without priori knowledge on the right part, it is also noticeable that the proposed approach has better estimation performance.

Finally, we validate the performance of beam tracking algorithms in more realistic situations. For a fair comparison, an identical feedback cycle between the transmitter and receiver is devised, which is sufficient to achieve the optimal system capacity with limited feedback hybrid precoding. As shown in Fig. 5, the proposed method evidently performs better tracking performance than the angle tracking strategies developed in the [14] and [7], which applied with the extended Kalman filter (EKF) to track the channel with the aid of beamswitching at the transmitter and the receiver. It can be noted that beam tracking algorithm provides more DOA error suppression. An intuitive explanation for this phenomenon is that, only when the nonlinearity of the solution problem is mild, the Kalman filter will have good performance. However, the actual dynamics of DOA and the property of ADC quantization can be nonlinear. It should further observed that, although the accumulated DOA estimation error will severely degrade system performance. However, with abrupt channel changes detection [15], once the abrupt changing detected, it suggests that the performance penalty can be negligible via a limited feedback link.

VI. CONCLUSION

In this paper, a Bayesian beamforming approach is proposed to deal with the DOA uncertainty in a mobile mmWave system. By exploiting a priori information of the motion trajectory dependent DOA, we have demonstrated that the proposed Bayesian beamformer is robust to uncertain DOA estimation and is suitable when the DOA is completely unknown. To alleviate the computational complexity, we present a motion trajectory-based priori probability approximation method. By means of simulations, the numerical results showed that the proposed solution is effective to achieve a high estimation accuracy within a specific spatial region. The solution framework proposed in this study effectively integrates the soft decision method based on approximate posterior probability and the iterative-based EM parameter estimation algorithm, and the angle variations can be successfully tracked.

REFERENCES

- [1] Z. Pi and F. Khan. An introduction to millimeter-wave mobile broadband systems. *IEEE Commun. Mag.*, 49(6):101–107, 2011.

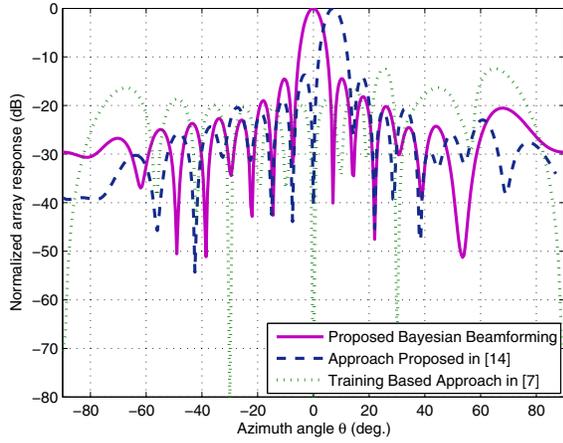


Fig. 2: The synthetic array response patterns of beamformer with Gaussian random perturbation. ϕ is set to be $\pi/2$; Ξ_c^2 is set to be 0.01; SNR=20 dB.

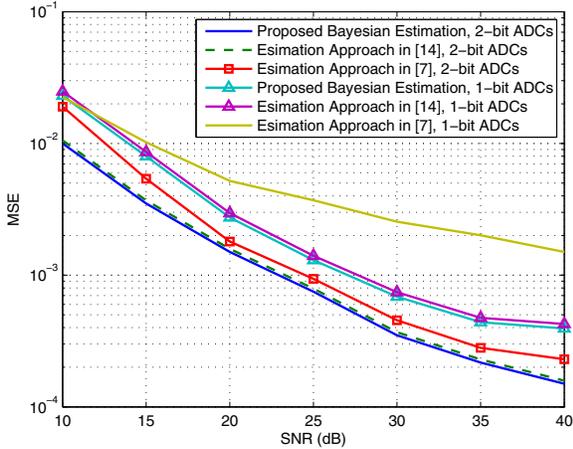


Fig. 3: MSE of DOA estimation vs. SNR

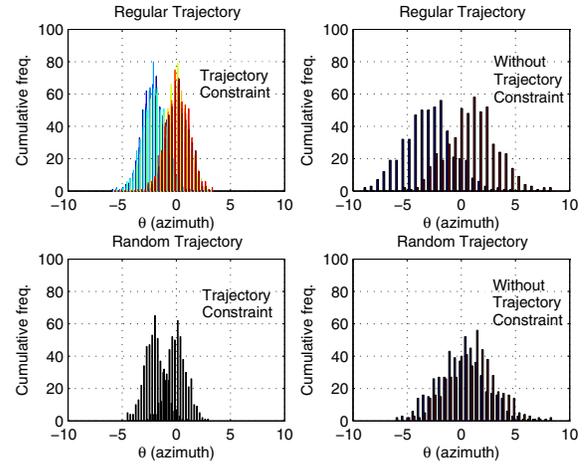


Fig. 4: Posteriori pdf with Bayesian estimation approximately calculated by 1000 snapshots. GM model is adopted by $p(\theta) = (1 - \lambda)g_1(\theta) + \lambda g_2(\theta)$ ($\mu_1 = 0$, $\mu_2 = 3$, $\sigma = 0.25$). SNR = 20 dB.

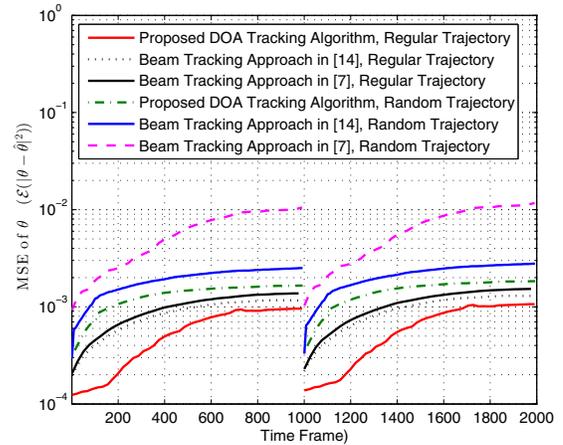


Fig. 5: Average tracking error using the proposed method.

[2] T. S. Rappaport et al. Millimeter wave mobile communications for 5G cellular: It will work! volume 1, pages 335–349, May 2013.

[3] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and Jr. R. W. Heath. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Commun. Mag.*, 52(12):122–131, December 2014.

[4] F. Sahrabi and W. Yu. Hybrid digital and analog beamforming design for large-scale antenna arrays. *IEEE J. Sel. Topics Signal Process.*, 10(3):501–513, April 2016.

[5] M. Xiao et al. Millimeter wave communications for future mobile networks. *IEEE J. Sel. Areas Commun.*, 35:1909–1935, June 2017.

[6] J. Mo, P. Schniter, and Jr. Robert W. Heath. Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs. *IEEE Trans. Signal Process.*, 66(5):1141–1154, March 2018.

[7] D. D. Donno, J. Palacios, and J. Widmer. Millimeter-wave beam training acceleration through low-complexity hybrid transceivers. *IEEE Trans. Wireless Commun.*, 16:3646–3660, June 2017.

[8] K. Bell, Y. Ephraim, and H. Van Trees. A Bayesian approach to robust adaptive beamforming. *IEEE Trans. Signal Process.*, 48(2):386–398, February 2000.

[9] C. Lam and A. Singer. Bayesian beamforming for DOA uncer-

tainty: Theory and implementation. *IEEE Trans. Signal Process.*, 54(11):4435–4445, November 2006.

[10] S. Malik, J. Benesty, and J. Chen. A Bayesian framework for blind adaptive beamforming. *IEEE Trans. Signal Process.*, 62(9):2370–2384, May 2014.

[11] Study on channel model for frequency spectrum above 6 GHz (release 14), June 2018.

[12] A. Goldsmith et al. Capacity limits of MIMO channels. *IEEE J. Sel. Areas Commun.*, 21(5):684–702, June 2003.

[13] J. A. Bilmes. *A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models*. Berkeley, CA: Comput. Sci. Div. Dep. Elect. Eng. Comput. Sci. U.C. Berkeley, 1998.

[14] D. Zhu, J. Choi, Q. Cheng, W. Xiao, and R. W. Heath Jr. High-resolution angle tracking for mobile wideband millimeter-wave systems with antenna array calibration. *IEEE Trans. Wireless Commun.*, 17(11):7173–7189, November 2018.

[15] C. Zhang, D. Guo, and P. Fan. Tracking angles of departure and arrival in a mobile millimeter wave channel. In *Proc. IEEE ICC '16*, pages 1–6, Kuala Lumpur, Malaysia, May 2016.