3D On and Off-Grid Dynamic Channel Tracking for Multiple UAVs and Satellite Communications

Jiadong Yu, Xiaolan Liu, Yue Gao, Senior Member, IEEE
Xuemin (Sherman) Shen, Fellow, IEEE

Abstract—The space-air-ground integrated network (SAGIN) has drawn increasing attention for its benefits, such as wide coverage, high throughput for 5G and 6G communications. As one of the links, space-air communications between multiple unmanned aerial vehicles (UAVs) and Ka-band orbiting low earth orbit (LEO) satellites face a crucial challenge in tracking the 3D dynamic channel information. This paper exploits a statistical dynamic channel model called the multi-dimensional Markov model (MD-MM), which investigates the more realistic spatial and temporal correlation in the sparse UAVs-satellite channel. Specifically, the spatial and temporal probabilistic relationships of multi-user (MU) hidden support vector, single-user (SU) joint hidden support vector, and SU hidden value vector are investigated. The specific transition probabilities that connect the SU and MU hidden support vector for both azimuth and elevation directions are defined. Moreover, based on the proposed MD-MM, we derive a novel multi-dimensional dynamic turbo approximate message passing (MD-DTAMP) algorithm for tracking the 3D dynamic channel in multiple UAVs systems. Furthermore, we also develop a gradient update scheme to recursively find the azimuth and elevation offset for 3D off-grid estimation. Numerical results verify that the proposed algorithm shows superior 3D channel tracking performance with smaller pilot overhead and comparable complexity.

Index Terms—Channel tracking, message passing, space-air communications.

I. INTRODUCTION

With increasing communication traffic demands for emerging applications, such as the internet of things (IoT), cloud computing, autonomous cars and so on, future networks have been expected to provide more robust communication services. Hence, the space-air-ground integrated network (SAGIN) has been widely studied for its advantages such as large coverage, high throughput and strong resilience for 5G and beyond systems [1]. SAGIN has mainly three segments: space segment consists of satellite network, air segment consists of aerial network, and ground segment consists of terrestrial network [2]. Acting as the middle segment, aerial network with unmanned aerial vehicles (UAVs) is cost–effective. It can provide a shorter line of sight (LoS) range for transmission compared with satellite-ground links [3]. UAVs are deployed as base stations in both crowded areas and non-accessible emergency scenarios for UAV-aided air-ground communications, which can provide seamless coverage. For space-air links, UAVs play a role as a relay to transmit the information received from satellite (e.g. low earth orbit (LEO) and medium earth orbit (MEO) satellites [4]) to terrestrial network with mmWave Ka-band [5], [6]. In the above two UAV-related communication links, due to the velocity of UAVs, it is undeniable that the estimation of the highly dynamic communication channels with lower overhead and good accuracy faces some challenges: the spatial and temporal variations caused by the mobility of UAVs in three dimensional (3D) space [7].

A. Related Work

Conventionally, channel state information (CSI) is estimated using techniques, such as least square (LS) and minimum mean square error (MMSE). However, for these methods, training overhead is proportional to the number of transmit antennas which is infeasible to mmWave communication with massive MIMO system [8]. To reduce the training overhead, compressive sensing (CS) based techniques, such as orthogonal matching pursuit (OMP) and compressive sampling matching pursuit (CoSaMP), have been widely studied to estimate the approximate sparse wireless channel in different domain [9].

Based on fundamental CS concepts and algorithms, some studies explored the more detailed sparse structure of the channel through spatial correlation [10]–[13] and temporal correlation [14]–[17]. Because of the physical scattering of the environment, [10] has concluded that the angular domain shows a burst sparsity structure. Hence, they proposed the burst least absolute shrinkage and selection operator (burst-LASSO) to extract the burst structure properties in the massive MIMO channels with enhanced performance. [11] also exploited the clustered sparse structure of the channel support in the spatial domain through Markov prior. [12] proposed a hidden Markov model to capture the angular domain structured sparsity of the channel. Moreover, [13] exploited the double sparsity of the multipath components in both angular domain and delay domain. However, these studies only considered the angular domain structure ([13] also considered the delay domain) based on static channel environment. Effectively utilizing the prior information can reduce the training overhead and maintain a good current channel estimation performance for dynamic
channel tracking. [14] proposed a prior aided channel tracking scheme by taking the practical user motion model into account through a temporal variation law of the physical direction. [15] modeled the channel by a Markov process and the Kalman filter to track the dynamic model. A structured matching pursuit [16] and differential orthogonal matching pursuit (D-OMP) [17] are proposed both by considering the temporal correlation that path delays change slowly over time but path gains evolve faster. 

Jointly considering both spatial correlation and temporal correlation can improve the channel tracking performance [18]–[21]. [18] proposed the channel tracking scheme based on the prior-aided iterative angle and Doppler shift estimation algorithm and with the help of the previously estimated channel parameters. [19] decoupled the channel information into angles of arrival/departure (AoAs/AoDs) information and gain knowledge. A modified unscented Kalman filter tracks the AoAs/ AoDs, and the gain is estimated from the proposed beam training. [20] and [21] formulated the more realistic dynamic models of the AoDs support over time through hidden Markov model in flat and 3D space, respectively.

B. Motivation and Contribution

The aforementioned channel estimation and tracking work in [10]–[12], [14]–[20] can be directly adapted to the ground links and air-ground links with mobile scenario. However, for space-air links, which consist of multiple UAVs and one satellite, there are some non-neglected features: the known orbiting satellite, the 3D UAVs moving trajectories, and the dominant LoS ray between each UAV and satellite [2].

Although [10]–[12] considered the angular domain sparsity of the channel, it’s non-trivial to include the temporal information for channel tracking, especially in highly mobile space-air links with orbiting satellite and UAVs navigation. Comparatively, [14]–[17] all considered the prior channel information for time-varying channel tracking. However, they ignored the spatial angular structured sparsity, [18] studied the prior information of both angle and Doppler shifts for channel tracking, the probabilistic relationship of the angle movement is not explored. Moreover, although [19] and [20] both jointly considered spatial and temporal correlation of the dynamic channel, the premise of the system were both communicated through a flat surface. Furthermore, although our previous work [21] considered a more realistic 3D space-air communications with only one UAV, however, the multiple-UAVs system can provide wider coverage in practice. Compared with one UAV in the system, for multiple UAVs, there will have more than one structured sparse points on the quantized grids to represent the multi-user (MU) spatial correlation with sharing azimuth or elevation supports. Hence, a more complicated scenario with multiple UAVs should be further explored.

The above algorithms all assumed that the true angles lie on the quantized grids. However, in the real world, the true angle is continuously distributed [22]. To solve the off-grid problem, [22] and [23] all proposed algorithms based on a two-dimensional (2D) flat environment. Hence, none of the mentioned work considered the 3D off-grid case of the channel model.

Motivated by the above literature review, we consider the multiple UAVs space-air links to recursively track the 3D off-grid dynamic channel with smaller training overhead, comparable complexity, and improved estimation performance. The main contributions are summarized as follows:

- We develop a statistical 3D on-grid and off-grid channel model called multi-dimensional Markov model (MD-MM) for space-air communication links with 3D navigation multiple UAVs and orbit LEO satellite in the system. The MD-MM captures the structured sparsity of the MU azimuth and elevation spatial domains and the probabilistic relationship between MU and SU support.
- Based on the proposed MD-MM channel model, we propose a novel multi-dimensional dynamic turbo approximate message passing (MD-DTAMP) algorithm which can recursively track the 3D on-grid dynamic channel over time.
- We further develop a gradient azimuth and elevation 3D off-grid estimation scheme for a more realistic channel model, which iteratively update the optimum 3D offset through gradient calculation.
- The proposed MD-DTAMP algorithm with and without an off-grid scheme all show superior estimation performance compared with benchmark algorithms. Moreover, the proposed algorithm can achieve solid performance with smaller pilot overhead and comparable complexity.

This paper is organized as follows. The downlink LEO satellite to multiple UAVs system model is introduced in Section II. In Section III, the 3D on-grid and off-grid MD-MM channel model are elaborated. The problem of channel tracking is formulated in Section IV. The proposed algorithm MD-DTAMP and the 3D gradient off-grid estimation scheme are presented in Section V. Numerical results and conclusions are then provided in Section VI and Section VII, respectively.

**Notation:** Throughout the paper, we use the following common notation. The complex numbers are denoted by $\mathbb{C}$. The transpose and the conjugate transpose are denoted by $(\cdot)^{T}$, $(\cdot)^{H}$ respectively. $I$ is the identity matrix. $CN$ (mean, covariance) indicates a complex Gaussian random vector with defined mean and covariance. $\| \cdot \|$ is the $l_2$ norm.

II. SYSTEM MODEL

In this paper, we consider the narrow-band flat fading transmission model through space-air downlink communication from LEO satellite to multiple UAVs. Since Ka-band is weather dependent, to mitigate the high path loss, the LEO satellite is installed with uniform planar array (UPA) with $M = N_x \times N_y$ antenna elements to compensate the transmission loss [6], [24]. The $k^{th}$ $\{1,\ldots, K\}$ UAV is equipped with a high gain omnidirectional single antenna$^1$ [25], [26]. The illustration of the

$^1$For simplicity of only exploring the temporal and spatial correlation of the azimuth and elevation AoD of the satellite, the high gain omni-directional single antenna is assumed to be carried by the $k^{th}$ UAV. However, the proposed system model can be further extended with UAVs also installed with UPA. The main difference is that the channel model should further take the receiver array response vectors with azimuth and elevation directions of AoAs into consideration.
The important notations of the system and channel model are shown in Table I. In training procedure, as symbols $S_k \in \mathbb{C}^{M \times N_a}$ are transmitted as UAV, the received signal at $k^{th}$ UAV at time slot $t$ can be expressed as

$$y_{k,t} = h_{k,t} S_k + n_{k,t}^{\text{RX}},$$

with $h_{k,t} \in \mathbb{C}^{1 \times M}$ as the frequency domain channel vector, and $n_{k,t}^{\text{RX}} \sim \mathcal{CN}(0, \sigma_n^2 I)$ is the additive white Gaussian noise.

![Diagram](image-url)  

Fig. 1. The illustration of space-air communications system model at time slot $t$.

### TABLE I

**TABLE OF IMPORTANT NOTATIONS**

| Notations | Explanations |
|-----------|--------------|
| $n_x, N_x$ | Index and total number of x-axis antenna/sparsity size |
| $n_y, N_y$ | Index and total number of y-axis antenna/sparsity size |
| $N_a$ | Total pilot number |
| $m, M$ | Index and total number of the antenna and the sparse grid |
| $t, T$ | Index and total number of the time slot |
| $\theta_{t,k}, \phi_{t,k}$ | $k^{th}$ UAV azimuth and elevation directions of AoD at $t$ |
| $\Delta \theta_{t,k}, \Delta \phi_{t,k}$ | $k^{th}$ UAV on-grid azimuth and elevation directions of AoD at $t$ |
| $\Delta \theta_{t,k}, \Delta \phi_{t,k}$ | $k^{th}$ UAV azimuth and elevation offset directions of AoD at $t$ |
| $\theta_{range}$ | The range of azimuth directions of AoD |
| $\phi_{range}$ | The range of elevation directions of AoD |
| $a(\theta_{t,k}, \phi_{t,k})$ | Transmit array response vector at $t$ |
| $h_{k,t}$ | $k^{th}$ UAV frequency domain channel model at $t$ |
| $S_k$ | $k^{th}$ UAV transmit pilot symbols |
| $y_{k,t}^{\text{RX}}$, $y_{k,t}^{(i)}$ | $k^{th}$ UAV received signal at $t$ |
| $y_{k,t}^{(j)},...,y_{k,t}^{(e)}$ | $k^{th}$ UAV received total signal for total period $t$ |
| $n_{k,t,1}, n_{k,t,2}$ | $k^{th}$ UAV additive white Gaussian noise at $t$ |
| $\bar{x} y_{k,t}^{(i)}$ | $k^{th}$ UAV complex channel gain for $i^{th}$ element of the $m^{th}$ AoD at $t$ |
| $\bar{x} y_{k,t}^{(i)}$ | $k^{th}$ UAV sparse matrix complex channel gain at $t$ |
| $\bar{x} y_{k,t}^{(i)}$ | $k^{th}$ UAV total sparse matrix complex channel gain for period $t$ or $T$ |
| $h_{k,t,1}, h_{k,t,2}$ | $k^{th}$ UAV hidden support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV hidden support vector for total period $t$ or $T$ |
| $a_{k,t,1}, a_{k,t,2}$ | MU Hidden support vector at $t$ |
| $\bar{x} a_{k,t}^{(i)}$ | $k^{th}$ UAV azimuth support element for the $m^{th}$ AoD at $t$ |
| $\bar{x} a_{k,t}^{(i)}$ | $k^{th}$ UAV azimuth support element for the $m^{th}$ AoD at $t$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support element for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |
| $\bar{x} h_{k,t}^{(i)}$ | $k^{th}$ UAV elevation support vector for total period $t$ or $T$ |

**A. Channel Model**

The LoS ray is the dominant path in space-air communication links [27]. Hence, the $h_{k,t}$ denotes the frequency domain channel vector with one dominant LoS path, which is written as,

$$h_{k,t} = \tilde{g}_{k,t} a^H(\theta_{k,t}, \phi_{k,t}),$$

where $\tilde{g}_{k,t}$ is the complex channel gain, $a(\theta_{k,t}, \phi_{k,t}) \in \mathbb{C}^{(N_x \times N_y) \times 1}$ represents the transmit array response vectors, $\theta_{k,t}$ and $\phi_{k,t}$ as azimuth and elevation directions of AoD for $k^{th}$ UAV at time slot $t$. As for UPA installed LEO satellite, the steering vector can be written as

$$a(\theta_{k,t}, \phi_{k,t}) = \frac{1}{\sqrt{N_x \times N_y}} \begin{bmatrix} 1, ..., e^{\frac{2\pi i}{N_y}} \sin \theta_{k,t} x sin \phi_{k,t} + \frac{n_x}{y} \cos \phi_{k,t} \end{bmatrix}_T,$$

where $\lambda$ is the signal wavelength, $0 \leq n_x < N_x$ and $0 \leq n_y < N_y$ are the indices of an antenna element with $N_x$ and $N_y$ as antenna array sizes.

**B. On-grid and Off-grid Virtual Channel Model**

Conventionally, to explore the sparse property of the channel, AoDs are assumed to lie on the quantized grids for simplicity [28]. With discretized AoDs, the channel model is called on-grid channel. However, in practice, the true angle is continuously distributed. The channel with this kind of AoDs is named as off-grid channel [23]. Hence in this subsection, we present both on-grid and off-grid cases.

For on-grid case, the channel model can be further written into a sparse representation as

$$h_{k,t} = g_{k,t} \tilde{A}^H(\tilde{\theta}_{k,t}, \tilde{\phi}_{k,t}),$$

where $g_{k,t} \in \mathbb{C}^{1 \times M}$ is a sparse vector with one non-zero elements (which means only one dominant path for the $k^{th}$ UAV), $M$ is the total number of spatial grids, the quantized AoD pairs are considered, i.e., $\theta_{k,t} = [\theta_{k,t}^{\text{range}}/N_x, ..., \theta_{k,t}^{\text{range}}/(N_x-1)/N_x]$ and $\phi_{k,t} = [\phi_{k,t}^{\text{range}}/N_y, ..., \phi_{k,t}^{\text{range}}/(N_y-1)/N_y]$, which means the virtual $k^{th}$ UAV’s AoD pair can consist of a $M$ uniform grid, and

$$\tilde{A}(\tilde{\theta}_{k,t}, \tilde{\phi}_{k,t}) \in \mathbb{C}^{M \times M}$$

is the dictionary matrix. For off-grid case, we define $\Delta \theta_{k,t,n_x}$ and $\Delta \phi_{k,t,n_y}$ to represent the offset angles for both azimuth and elevation directions, respectively. Hence, the true AoD azimuth and elevation angle should be the combination of the nearest on-grid angle and the relative offset.
angle. As an example, we demonstrate the $k^{th}$ UAV’s AoD azimuth direction. The illustration can be found in Fig. 2. We assume that the true azimuth angle $\theta_{k,t}$ has the nearest on-grid angle $\tilde{\theta}_{k,t}$ and the offset angle is $\Delta \theta_{k,t,n_z} = \theta_{k,t} - \tilde{\theta}_{k,t,n_z}$.

Similarly, the AoD elevation direction offset angle can be written as $\Delta \phi_{k,t,n_y} = \phi_{k,t} - \tilde{\phi}_{k,t,n_y}$. Hence, the off-grid channel can be written as

$$h_{k,t} = g_{k,t} \tilde{A}^H \left( \theta_{k,t} + \Delta \theta_{k,t}, \phi_{k,t} + \Delta \phi_{k,t} \right),$$

where

$$\tilde{A} \left( \theta_{k,t} + \Delta \theta_{k,t}, \phi_{k,t} + \Delta \phi_{k,t} \right) = 
\begin{bmatrix}
\alpha \left( \tilde{\theta}_{k,t,1} + \Delta \tilde{\theta}_{k,t,1}, \tilde{\phi}_{k,t,1} + \Delta \tilde{\phi}_{k,t,1} \right), ... , \\
\alpha \left( \tilde{\theta}_{k,t,n_x} + \Delta \tilde{\theta}_{k,t,n_x}, \tilde{\phi}_{k,t,n_x} + \Delta \tilde{\phi}_{k,t,n_x} \right), ... ,
\end{bmatrix}$$

(6)

is the off-grid array response vector dictionary matrix. To be noticed, the AoD azimuth offset vector $\Delta \theta_{k,t} = [\Delta \theta_{k,t,1}, ..., \Delta \theta_{k,t,n_x}]^T$ with

$$\Delta \theta_{k,t,n_x} = \begin{cases} 
\tilde{\theta}_{k,t} - \tilde{\theta}_{k,t,n_z}, & \text{Active path,} \\
0, & \text{Otherwise.}
\end{cases}$$

(7)

The otherwise case represents the non-active virtual on-grid offsets which can be observed in Fig. 2 with the notation from $\Delta \theta_{k,t,1}$ to $\Delta \theta_{k,t,n_x}$ without the active path offset $\Delta \theta_{k,t,n_z}$. Similarly, the AoD elevation offset vector $\Delta \phi_{k,t} = [\Delta \phi_{k,t,1}, ..., \Delta \phi_{k,t,n_y}]^T$ with

$$\Delta \phi_{k,t,n_y} = \begin{cases} 
\tilde{\phi}_{k,t} - \tilde{\phi}_{k,t,n_y}, & \text{Active path,} \\
0, & \text{Otherwise.}
\end{cases}$$

(8)

When $\Delta \theta_{k,t} = 0$ and $\Delta \phi_{k,t} = 0$, the off-grid channel can be treated as the on-grid channel in (4).

III. MULTI-DIMENSIONAL MARKOV MULTIUSER CHANNEL MODEL

In our model, communication links connected between LEO satellite and multiple UAVs are considered. Because of the LEO satellite height, a certain portion of the earth’s surface is covered at any given time slot. This means a constellation of LEO satellites is necessary for full coverage of the earth [29]. We consider a condition that one LEO satellite communicates with the same group of multiple UAVs simultaneously before any communication link is handed over to the adjacent LEO satellite. This means the LEO satellite would orbit in a certain distance range, limiting the azimuth and elevation angle in fixed ranges. With assumptions above, we transfer the complex continuous relative displacement relationships between LEO satellite and UAVs into time-varying angular domain variance with transition probability similar to [19]–[21] (this is further elaborated in Section III-C).

In the MU scenario with multiple UAVs, we specially explore the relationship between the MU and SU hidden support vector. The MU azimuth $\tilde{\alpha}^{(T)}$ and elevation $\tilde{\epsilon}^{(T)}$ support vector are both modeled as two dimensional Markov model (2D-MM) similarly in [21]. $\tilde{\alpha}^{(T)}$ and $\tilde{\epsilon}^{(T)}$ represent the MU azimuth and elevation support vector over total time period $T$. During time slot $t$, $\tilde{\alpha}_t = [\tilde{\alpha}_{t,1}, ..., \tilde{\alpha}_{t,N_y}] \in \{0,1\}^{N_x}$ with $\tilde{\alpha}_{t,n_x} = \tilde{\alpha}_{t,m} \lor \tilde{\alpha}_{t,n_x} \lor ... \lor \tilde{\alpha}_{t,M}$, which means MU azimuth support is the union of the SU azimuth support of total $K$ UAVs at time slot $t$. Similarly, for MU elevation support $\tilde{\epsilon}_t = [\tilde{\epsilon}_{t,1}, ..., \tilde{\epsilon}_{t,N_y}] \in \{0,1\}^{N_y}$ with $\tilde{\epsilon}_{t,n_y} = \tilde{\epsilon}_{t,m} \lor \tilde{\epsilon}_{t,n_y} \lor ... \lor \tilde{\epsilon}_{t,M}$. Furthermore, the MU joint hidden support vector $b_{k,t}$ is modeled as three dimensional hidden Markov model (3D-HMM) which jointly combines the single-user (SU) azimuth and elevation support.

The details of the complete multi-dimensional Markov model (MD-MM) channel model are introduced in four parts: probability model for SU channel, 2D Markov model (2D-MM) for azimuth and elevation MU hidden support vector, 3D-HMM for azimuth and elevation SU joint hidden support vector, and Gaussian Markov model for SU hidden value vector.

A. Probability Model for SU Channel

The dynamic angular domain channel for the $k^{th}$ UAV $g_k^{(T)} = \{g_{k,t,1}, ..., g_{k,T}\}$ with $g_{k,t} = \{g_{k,t,1}, ..., g_{k,t,M}\}$ can be modeled as a probabilistic signal model with two hidden random processes, SU joint hidden support vector $b_k^{(T)} = \{b_{k,1}, ..., b_{k,T}\}$ and SU hidden value vector $\vartheta_k^{(T)} = \{\vartheta_{k,1}, ..., \vartheta_{k,T}\}$. Here, $b_{k,t} = [b_{k,t,1}, ..., b_{k,t,N_y,N_z}] = [b_{k,t,1}, ..., b_{k,t,M}] \in \{0,1\}^M$ represents the SU joint hidden support vector at time $t$, which indicates the $k^{th}$ UAV channel sparsity with $b_{k,t,m} \in \{1,0\}$. The SU hidden value vector $\vartheta_{k,t} = [\vartheta_{k,t,1}, ..., \vartheta_{k,t,M}] \in \mathbb{C}^M$ represents the temporal correlation of the $k^{th}$ UAV channel coefficients. The dynamic channel element for the $m^{th}$ AoD direction in the time $t$ of the $k^{th}$ UAV can be written as

$$g_{k,t,m} = b_{k,t,m} \vartheta_{k,t,m},$$

(9)

where $b_{k,t,m}$ denotes whether there is an active path or not (i.e., $b_{k,t,m} = 1$ means the $m^{th}$ AoD path is activated) and $\vartheta_{k,t,m}$ represents the path gain.

To formulate the more realistic dynamic channel (i.e., the inactive AoD path at the current time $t$ may be activated at $2$ For paring the azimuth and elevation directions, $m = (n_x - 1) \times N_x + n_y$.}
time \( t+1 \), a probabilistic channel model of \( k^{th} \) UAV with channel prior distribution can be formed as

\[
p \left( g_k^{(T)}, b_k^{(T)}, \vartheta_k^{(T)} \right) = p \left( g_k^{(T)} \mid b_k^{(T)}, \vartheta_k^{(T)} \right) p \left( b_k^{(T)} \right) p \left( \vartheta_k^{(T)} \right),
\]

(10)

where the \( k^{th} \) UAV channel vector conditional prior is

\[
p \left( g_k^{(T)} \mid b_k^{(T)}, \vartheta_k^{(T)} \right) = \prod_{t=1}^{T} \prod_{m=1}^{M} p \left( g_{k,t,m} \mid b_{k,t,m}, \vartheta_{k,t,m} \right)
= \prod_{t=1}^{T} \prod_{m=1}^{M} \delta \left( g_{k,t,m} - b_{k,t,m}, \vartheta_{k,t,m} \right),
\]

(11)

with \( \delta(\cdot) \) is the Dirac delta function. The factor graph of the combined \( k^{th} \) UAV channel can be found in Fig. 3, with \( p_{k,t,m} \) denotes the conditional prior \( p \left( g_{k,t,m} \mid b_{k,t,m}, \vartheta_{k,t,m} \right) \). This factor graph shows that the final \( k^{th} \) UAV channel gain \( g_{k,t} \) is the combination of the SU joint hidden support vector \( b_{k,t} \) (on the left hand side of Fig. 3) and the SU hidden value vector \( \vartheta_{k,t} \) (on the right hand side of Fig. 3).

B. Gaussian Markov Model for SU Hidden Value Vector

The illustration of the factor graph for the \( k^{th} \) UAV SU hidden value vector \( \vartheta_{k,t} \) and the Gaussian-Markov model can be found in Fig. 3 (on the right hand side). Due to the path gains change smoothly over time [17], the hidden value vector can be formulated as the Gauss-Markov processes [30]

\[
\vartheta_{k,t,m} = (1 - \beta_k) (\vartheta_{k,t-1,m} - \mu_k) + \beta_k \omega_{k,t,m} + \mu_k,
\]

(12)

where \( \beta_k \in [0,1] \), \( \omega_{k,t,m} \sim \mathcal{CN}(0, \varsigma_k) \), \( \mu_k \in \mathbb{C} \) is the mean of the process. When \( \beta_k = 0 \), then \( \vartheta_{k,t,m} = \vartheta_{k,t-1,m} \), which indicates that \( \vartheta_{k,t} \) remain unchanged over time. When \( \beta_k = 1 \), then \( \vartheta_{k,t,m} = \beta_k \omega_{k,t,m} + \mu_k \), which means that the SU hidden value vector shows the i.i.d Gaussian distribution with mean \( \mu_k \) over time. If \( 0 < \beta_k < 1 \), the conditional probability can be written as

\[
p \left( \vartheta_{k,t,m} \mid \vartheta_{k,t-1,m} \right) \sim \mathcal{CN} \left( \vartheta_{k,t,m}; (1 - \beta_k) \vartheta_{k,t-1,m} + \beta_k \mu_k, \varsigma_k^2 \right).
\]

(13)

The distribution of the steady state of the process is \( \vartheta_{k,t,m} \sim \mathcal{CN}(\mu_k, \varsigma_k^2) \). And the joint distribution of \( \vartheta_k^{(T)} \) can be formulated as

\[
p \left( \vartheta_k^{(T)} \right) = \prod_{m=1}^{M} \prod_{t=2}^{T} p \left( \vartheta_{k,t,m} \mid \vartheta_{k,t-1,m} \right),
\]

(14)

where \( p \left( \vartheta_{k,1,m} \right) \) and \( p \left( \vartheta_{k,t,m} \mid \vartheta_{k,t-1,m} \right) \) are the transition probabilities that equal to the factor nodes \( f_{k,1,m} \) and \( f_{k,t,m} \) in Fig. 3.

C. 2D-MM for Azimuth and Elevation MU Hidden Support Vector

The details of the azimuth and elevation MU hidden support vector are introduced as follows: To explain the 2D-MM for MU hidden support vector of azimuth and elevation directions, we take three UAV users as an example in Fig. 4(a). The relationship of the MU azimuth support \( \bar{\alpha}_t \) and the MU elevation support \( \bar{\epsilon}_t \) with three activated paths are shown. To be noticed, in this example, there are two users share the same MU azimuth support. Hence, the MU azimuth support \( \bar{\alpha}_t \) has only two activated points at time \( t \). For MU elevation support \( \bar{\epsilon}_t \) has three activated points at time \( t \).

Because of the navigation of the UAVs [31], together with the known orbital voyage of the LEO satellite, the relative displacement of the UAV and satellite is continuous. Hence, the spatial sparsity pattern will act highly related with the prior adjacent pattern, i.e., \( \bar{\alpha}_{t+1,n} \) has higher probability be influenced by the adjacent \( \bar{\alpha}_{t,n-1} \) in previous time slot [10]. [17] has also verified that the channel support change slowly over time, which indicates \( \bar{\alpha}_{t+1,n} \) has higher probability to be influenced by itself in previous time slot \( \bar{\alpha}_{t,n} \). This relationship which considers both spatial correlation and temporal correlation is named as dynamic 2D sparsity. The illustration of this 2D-MM can be found in the factor graph in Fig. 4(b). The model over total time period \( T \) can be expressed as:

\[
p \left( \bar{\alpha}_1^{(T)}, \bar{\epsilon}_1^{(T)} \right) = p \left( \bar{\alpha}_1, \bar{\epsilon}_1 \right) \prod_{t=2}^{T} p \left( \bar{\alpha}_{t,n} \mid \bar{\alpha}_{t-1,n} \right) \prod_{t=2}^{T} p \left( \bar{\epsilon}_{t,n} \mid \bar{\alpha}_{t,n} \right)
\]

(15)

where \( \bar{\alpha}_{1,n}, \bar{\alpha}_{1,n}^{(A)}, \bar{\alpha}_{1,n-1}, \bar{\alpha}_{1,n-1}^{(A)} \) and \( \bar{\epsilon}_{1,n}, \bar{\epsilon}_{1,n}^{(A)}, \bar{\epsilon}_{1,n-1}, \bar{\epsilon}_{1,n-1}^{(A)} \) are the transition probabilities that equal to the factor nodes \( \bar{\alpha}_{1,n}, \bar{\alpha}_{1,n}^{(A)}, \bar{\alpha}_{1,n-1}, \bar{\alpha}_{1,n-1}^{(A)} \). The support of the MU elevation support of AoD \( p \left( \bar{\epsilon}_1^{(T)} \right) \) can be modeled similarly.
D. 3D-HMM for azimuth and elevation SU joint hidden support vector

To explore the relationship between MU and SU, we modeled the joint prior distribution of MU and SU azimuth or elevation support as HMM with conditional probability. Take the $k^{th}$ UAV SU azimuth support as an example:

$$p\left(\alpha^{(T)}_{k,t},\alpha^{(T)}_{k,t,n_z}\right) = \prod_{t=1}^{T} \prod_{k=1}^{K} p\left(\alpha_{k,t}\mid \alpha_t\right)$$

(16)

where $q_{k,t,n_z}^{(A)} \alpha_{k,t,n_z}$ is the transition probability equal to the factor node $q_{k,t,n_z}$ in Fig.5(b). Furthermore, $\tilde{\alpha}^{(T)}$ and $\alpha^{(T)}$ are the MU azimuth support vector and $k^{th}$ UAV SU azimuth support vector over time period $T$, respectively. Although in practice, multiple UAVs may share partially MU azimuth or elevation support, i.e., in Fig.5, user 2 and user 3 both have the $4^{th}$ grid activated which means the same azimuth support is shared, our HMM model can still apply. To be noticed, each element in $\alpha_{k,t,n_z}$ generated by the MU azimuth support vector in the HMM with certain conditional probability, i.e.,

$$p\left(\alpha_{k,t,n_z} = 1\mid \bar{\alpha}_{t,n_z} = 0\right) = q_{k,t,0,1}^{(A)}$$

and $p\left(\alpha_{k,t,n_z} = 1\mid \bar{\alpha}_{t,n_z} = 0\right) = q_{k,t,1,1}^{(A)}$. In another word, when the MU azimuth support element $\bar{\alpha}_{t,n_z} = 0$, there is no way that the $k^{th}$ UAV SU azimuth support element $\alpha_{k,t,n_z} = 0$. However, when the MU azimuth support element $\bar{\alpha}_{t,n_z} = 1$, there is a certain probability $q_{k,t,1,1,1}^{(A)}$ that the $n_z^{th}$ element $\alpha_{k,t,n_z} = 1$ in the $k^{th}$ UAV SU azimuth support vector is 1. The $k^{th}$ UAV SU elevation support can be modeled similarly.

Hence, the 3D-HMM of the SU joint hidden support $b^{(T)}_{k,t}$ over total period time $T$ which combines both azimuth and elevation direction supports can be found in Fig. 5(c). The joint conditional prior of the channel support probability of the $k^{th}$ UAV is given by

$$p\left(b^{(T)}_{k,t}\mid \alpha^{(T)}_{k,t}, e^{(T)}_{k,t}\right) = \prod_{t=1}^{T} \prod_{n_z=1}^{N_z} \prod_{n_y=1}^{N_y} p\left(b_{k,t,n_z,n_y}\mid \alpha_{k,t,n_z}, e_{k,t,n_y}\right).$$

(17)

IV. 3D Off-grid Dynamic Sparse Channel Tracking Problem Formulation

By taking the off-grid channel (5) into consideration, the received signal at $k^{th}$ UAV can be written as

$$y^{RX}_{k,t} = g_{k,t}^T A^{H} \left(\tilde{\theta}_{k,t} + \Delta \theta_{k,t}, \tilde{\phi}_{k,t} + \Delta \phi_{k,t}\right) S_{k} + n^{RX}_{k,t},$$

(18)

with $y^{RX}_{k,t} \in \mathbb{C}^{1 \times N_y}$. To convert the received signal into the standard CS model, we reshape $g_{k,t}$ from $1 \times M$ to $M \times 1$. Hence the received signal can be rewritten as

$$y_{k,t} = \Phi_{k,t} \left(\Delta \theta_{k,t}, \Delta \phi_{k,t}\right) g_{k,t} + n_{k,t},$$

(19)

with the sensing matrix $\Phi_{k,t} \left(\Delta \theta_{k,t}, \Delta \phi_{k,t}\right) = S_{k}^T \left(A^{H} \left(\tilde{\theta}_{k,t} + \Delta \theta_{k,t}, \tilde{\phi}_{k,t} + \Delta \phi_{k,t}\right)\right)^T \in \mathbb{C}^{N_z \times M}$. $y_{k,t} \in \mathbb{C}^{N_z \times 1}$. For the training phase, the transmitted pilot $S_{k}$ is designed based on the partial discrete Fourier transform (DFT) random permutation measurement matrix $\Phi_{k,t} \left(\Delta \theta_{k,t}, \Delta \phi_{k,t}\right)$ [11], [32]. Details are shown in Appendix A. Given the received signals of total $t$ time period $y^{(t)}_{k,t}$, the purpose is to track the dynamic channel vector $g_{k,t}$ at time $t$. The channel vector can be estimated as $\hat{g}_{k,t,m} = E\left(g_{k,t,m}\mid y^{(t)}_{k,t}, \Delta \theta_{k,t}, \Delta \phi_{k,t}\right)$, when the expectation is over the marginal posterior

$$p\left(g_{k,t,m}\mid y^{(t)}_{k,t}, \Delta \theta_{k,t}, \Delta \phi_{k,t}\right) \propto \sum_{b^{(t)}_{k,t,m}} \int_{\tilde{\alpha}^{(t)}_{k,t,m}} \phi^{(t)}_{k} p\left(g^{(t)}_{k}, \tilde{\alpha}^{(t)}_{k,t,m}, b^{(t)}_{k,t,m}\right)$$

(20)

\[\tilde{\alpha}^{(t)}_{k,t,m}, b^{(t)}_{k,t,m}, \tilde{\phi}^{(t)}_{k}, \tilde{\theta}^{(t)}_{k,t,m}, \tilde{\phi}^{(t)}_{k,t,m}\mid \Delta \theta_{k,t}, \Delta \phi_{k,t}\] with $g^{(t)}_{k,t,m}$ denotes the vector $g^{(t)}_{k,t}$ without the element $g^{(t)}_{k,t,m}$. Our goal is to obtain the MMSE of the estimated $g^{(t)}_{k,t}$. At the same time, the optimal off-grid azimuth and elevation offset $\Delta \theta^{(opt)}_{k,t}$ and $\Delta \phi^{(opt)}_{k,t}$ are obtained by maximum likelihood.
a contains both SU azimuth and elevation vector support vector
b a T 6) all through time SU hidden value vector ˜ vector
graph contains three sub-graphs: the MU azimuth and elevation
algorithm is elaborated in this section.
the reconstruction performance. The details of the proposed
further employ a turbo iterative framework [37] to improve
challenging to gain the exact marginal posterior in (20). We
factor graph Fig. 6, the complex structure makes it more
structured dynamic channel model. As can be seen from the
ever, the prior distribution knowledge is time-varying in our
sparse channel estimation [30], [35], [36] because of its good
reconstruction accuracy and low complexity. Traditionally,
sUb = argmax
g = argmax
(ML) as follows,

\[ \Delta \theta_{k,t}^{(opt)}, \Delta \phi_{k,t}^{(opt)} = \arg \max_{\Delta \theta_{k,t}, \Delta \phi_{k,t}} \ln p(y_k^{(t)}, \Delta \theta_{k,t}, \Delta \phi_{k,t}) \]  
(21)

\[ = \arg \max_{\Delta \theta_{k,t}, \Delta \phi_{k,t}} \int p(g_k^{(t)}, y_k^{(t)}, \Delta \theta_{k,t}, \Delta \phi_{k,t}) dg_k^{(t)} \]

V. MULTI-DIMENSIONAL DYNAMIC TURBO APPROXIMATE
MESSAGE PASSING ALGORITHM WITH 3D OFF-GRID
ESTIMATION

Approximate Message Passing (AMP) based algorithms
have been widely studied for user detection [33], [34] and
sparse channel estimation [30], [35], [36] because of its good
reconstruction accuracy and low complexity. Traditionally,
AMP is applied based on a static prior information. How-
ever, the prior distribution knowledge is time-varying in our
structured dynamic channel model. As can be seen from the
factor graph Fig. 6, the complex structure makes it more
challenging to gain the exact marginal posterior in (20). We
further employ a turbo iterative framework [37] to improve
the reconstruction performance. The details of the proposed
algorithm is elaborated in this section.

The factor graph of the total MD-MM distribution
\[ p(g_{k,t}^{pri}, \hat{\theta}^{(t)}, \hat{\phi}^{(t)}, b_k^{(t)}, \theta_k^{(t)}, y_k^{(t)}) \] is shown in Fig. 6. This factor
graph contains three sub-graphs: the MU azimuth and elevation
vector \( \hat{a}_t \) and \( \hat{e}_t \) (on the left-hand side of Fig. 6), the SU hidden
support vector \( \hat{b}_{k,t} \) (in the middle of Fig. 6), and the \( k^{th} \) UAV
SU hidden value vector \( \hat{\theta}_{k,t} \) (on the right-hand side of Fig.
6) all through time T. The MU azimuth and elevation vector
\( \hat{a}_t \) and \( \hat{e}_t \) are both structured as 2D-MM. The \( k^{th} \) UAV
SU hidden support vector \( \hat{b}_{k,t} \) is structured as 3D-HMM which
contains both SU azimuth and elevation vector \( \hat{a}_{k,t} \) and \( \hat{e}_{k,t} \).

The details of all the necessary nodes in the factor graph are
already explained in Section III. The expression of the factor
nodes, distributions and functional forms are shown in Table II.

Due to the complex structure of the total factor graph of
the MD-MM, it is challenging to achieve the exact marginal pos-
terior in (20). In this section, to explore the dynamic channel
structure, we explain the structural details of the proposed al-
gorithm based on the turbo framework with the proposed MD-
MM MMSE denoiser. In [37], conventional turbo compressed
sensing is achieved based on a two-step iteration framework:
linear minimum mean-square error (LMMSE) estimator and
i.i.d. prior-based MMSE estimator. Our proposed algorithm
will be introduced in four parts: Module A LMMSE Estimator,
Module B MD-MM MMSE denoiser, message passing through
time, and 3D off-grid estimation.

A. Module A: LMMSE Estimator

In Module A, the angular channel vector \( g_{k,t} \) at time \( t \) for
\( k^{th} \) UAV is estimated based on the the received signal \( y_{k,t} \) with
the prior distribution \( CN \left( g_{k,t}; g_{\text{prior},A,k,t}, v_{\text{prior},A,k,t} \right) \). The extrinsic
mean and variance are defined as [37]

\[ g_{\text{post}}^{A,k,t} = g_{\text{prior}}^{A,k,t} + \frac{v_{\text{prior}}^{A,k,t}}{\sigma^2_v} \Phi_H \left( y_{k,t} - \Phi_y \hat{g}_{A,k,t}^{\text{prior}} \right) \]
and

\[ v_{\text{post}}^{A,k,t} = \frac{N_S}{M} \frac{1}{v_{\text{prior}}^{A,k,t} + \sigma^2_v}. \]

\[ \text{Fig. 6. Factor graph representation of the MD-MM.} \]
respectively. The extrinsic distribution of $g_{k,t}$ is given by

\[
CN\left(\mu_{\text{extr}k,t}^{\text{ext}}, \Omega_{\text{extr}k,t}^{\text{ext}}\right)
\]

where the extrinsic mean and variance are respectively given by

\[
g_{k,t}^{\text{extr}} = \nu_{k,t}^{\text{ext}} (g_{k,t}^{\text{post}} - g_{k,t}^{\text{pri}}) \tag{25}
\]

and

\[
\nu_{k,t}^{\text{ext}} = \left(\frac{1}{\nu_{k,t}^{\text{post}}} - \frac{1}{\nu_{k,t}^{\text{pri}}}\right)^{-1} \tag{26}
\]

B. Module B: Message Passing MMSE Denoiser

Following LMMSE estimator in module A, the main task for module B is to exploit the $k^{th}$ UAV spatial sparsity details of the proposed MD-MM in the multi-UAVs system. Hence, the extrinsic mean and variance pass through the graph, which means $g_{k,t}^{\text{extr}}_{\text{B},k,t} = g_{k,t}^{\text{extr}}_{k,t}$ and $\nu_{k,t}^{\text{ext}}_{\text{B},k,t} = \nu_{k,t}^{\text{ext}}_{k,t}$ [37]. In module B, the purpose is to calculate the approximate posterior distributions of $p(g_{k,t,m} | g_{k,t}^{\text{extr}}_{k,t})$ instead of (20) using the sum-product message passing structure shown in the factor graph. For message passing algorithm, the basic assumption is that $g_{k,t}^{\text{pri}}_{\text{B},k,t} = g_{k,t} + z_{k,t}$ where $z_{k,t} \sim CN(0, \sigma_{\text{B},k,t}^{\text{pri}} I)$ is independent of $g_{k,t}$ [37].

The formulation details are elaborated in Appendix B and the message passing procedure can be summarized as follows:

- Firstly, in module A, based on the prior information passed from the last time slot $t$, the message passing over the $k^{th}$ UAV Gaussian hidden value vector $\theta_{k,t}$ are given by (52) with $\mu_{\text{dyn}}^{\text{post}}$ and $(\Omega_{\text{dyn}}^{\text{post}})^{-1}$ as the input to update the message $\nu_{k,t,m}^{\text{post}} \rightarrow \pi_{k,t,m}$.
- Secondly, in module B, $g_{k,t}^{\text{extr}}_{\text{B},k,t}$ and $\nu_{k,t}^{\text{ext}}_{\text{B},k,t}$ from module A are passed over the path $g_{k,t,m} \rightarrow \pi_{k,t,m} \rightarrow b_{k,t,m} \rightarrow u_{k,t,m} \rightarrow a_{k,t,n} \rightarrow \nu_{k,t,n}^{\text{ext}}$ (using (53) to (56)).
- Thirdly, the forward-backward message passing is executed over the proposed MD-MM MU azimuth branch with three paths: $a_{k,t,n} \rightarrow g_{k,t,n} \rightarrow \tilde{a}_{t,n}$, the 2D-MM of the MU azimuth support $\tilde{a}_{t}$, and $\tilde{a}_{t,n} \rightarrow g_{k,t,n} \rightarrow a_{k,t,n}$ using (59) to (68).
- Then the message pass through path $a_{k,t,n} \rightarrow \nu_{t,m} \rightarrow \pi_{t,n} \rightarrow b_{t,m} \rightarrow u_{k,t,m} \rightarrow a_{k,t,n} \rightarrow \nu_{k,t,n}^{\text{ext}}$ (using (69) to (71)) to another MD-MM elevation branch with similar three paths: $\epsilon_{k,t,n} \rightarrow \nu_{k,t,n} \rightarrow \tilde{e}_{t,n}$, the 2D-MM of the MU elevation support $\tilde{e}_{t}$, and $\tilde{e}_{t,n} \rightarrow \nu_{k,t,n} \rightarrow \epsilon_{k,t,n}$ (using (71) to (82)).
- The overall details of the MD-MM azimuth and elevation SU joint hidden support estimation can be found in Algorithm 1.
Based on the calculated updated messages \( v_{\pi_{k,t,m} \rightarrow g_{k,t,m}} \), the posterior distributions can be written as
\[
p\left(g_{k,t,m} \| g_{B,k,t}^{\text{pri}}\right) \propto v_{\pi_{k,t,m} \rightarrow g_{k,t,m}}(g_{k,t,m}) \cdot v_{\rho_{k,t,m} \rightarrow g_{k,t,m}}(g_{k,t,m}),
\]
where \( v_{\rho_{k,t,m} \rightarrow g_{k,t,m}}(g_{k,t,m}) = \mathcal{CN}\left(\hat{g}_{B,k,t,m}^{\text{pri}},\sigma_{B,k,t,m}^{\text{pri}}\right) \).
Then the posterior mean and variance of each element of \( g_{k,t} \) can be updated as
\[
g_{B,k,t,m}^{\text{post}} = \int_{g_{k,t,m}} g_{k,t,m} p\left(g_{k,t,m} \| g_{B,k,t}^{\text{pri}}\right)
\]
and
\[
v_{B,k,t} = \frac{1}{M} \sum_{m=1}^{M} \text{Var}\left(g_{k,t,m} \| g_{B,k,t}^{\text{pri}}\right),
\]
where \( \text{Var}\left(g_{k,t,m} \| g_{B,k,t}^{\text{pri}}\right) \) denotes the conditional variance of \( g_{k,t,m} \) given \( g_{B,k,t}^{\text{pri}} \). Then similar to (25) and (26), the extrinsic update mean and covariance can be written as
\[
g_{A,k,t}^{\text{pri}} = g_{B,k,t}^{\text{ext}} = v_{A,k,t}^{\text{pri}} \frac{g_{B,k,t}^{\text{post}} - g_{B,k,t}^{\text{pri}}}{v_{B,k,t}^{\text{pri}}}
\]
and
\[
v_{A,k,t}^{\text{pri}} = v_{B,k,t}^{\text{ext}} = \left(1 - \frac{1}{v_{B,k,t}^{\text{pri}}} \right) - 1.
\]

C. Message Passing Through Time

For dynamic scenario, prior information which is also treated as the temporal correlation can be considered for better reconstruction of the time-varying channel model. Hence, we further consider the message passing through time for channel tracking. There are mainly three branches passing through time in Fig. 6 the factor graph of the MD-MM: the path of the \( k^{th} \) UAV Gaussian hidden value vector \( \theta_{k,t} \), the path of MU azimuth support \( \hat{a}_k \) and the path of MU elevation support \( \hat{e}_k \).

1) The \( k^{th} \) UAV Gaussian Hidden Value Vector \( \theta_{k,t} \) Passing Through Time: The message passing from factor node \( \pi_{k,t,m} \) to variable node \( \theta_{k,t,m} \) should be
\[
v_{\pi_{k,t,m} \rightarrow \theta_{k,t,m}}(\theta_{k,t,m}) = p_{\theta_{k,t,m}} \mathcal{CN}\left(\hat{\theta}_{B,k,t,m},\sigma_{B,k,t,m}^{\text{true}}\right)
+ (1 - p_{\theta_{k,t,m}}) \mathcal{CN}\left(\hat{\theta}_{B,k,t,m},\sigma_{B,k,t,m}^{\text{true}}\right).
\]

To be noticed, if \( b_{k,t,m} = 0 \), \( \theta_{k,t,m} \) would be 0 in (9), which makes \( \theta_{k,t,m} \) unobservable. Hence, similarly in (30), a threshold which is slightly less than 1 is introduced. The modified message passing is given by
\[
v_{\pi_{k,t,m} \rightarrow \theta_{k,t,m}}(\theta_{k,t,m}) = \mathcal{CN}\left(\hat{\theta}_{B,k,t,m},\mu_{k,t,m}^{\text{true}},\sigma_{k,t,m}^{\text{true}} \right).
\]

Algorithm 1 MD-MM k-th SU Azimuth and Elevation Support Estimation

| Input: \( p_{\theta_{k,n}}, \hat{A}_{k,t,n,x}, \hat{E}_{k,t,n,y} \) |
| Output: \( \rho_{k,t,n,z}, \hat{e}_{k,t,n,y} \) |
| 1: Calculate \( v_{uk,t,m \rightarrow ak_{t,n,x}} \) as in (59) and \( v_{ek,t,m \rightarrow \epsilon_{k,t,n,y}} \) as in (68) |
| # HMM k-th SU Azimuth Support Estimation Based on 2D-MM MU Azimuth Support: |
| 2: if \( t = 1 \) then |
| 3: Initialization: \( \tilde{z}_{t,1} = \lambda, \tilde{z}_{t,N_e} = \frac{1}{2} \) |
| 4: for \( n_x = 2, \ldots, N_x \) do |
| 5: Use \( \rho_{t,1} \) to calculate \( \tilde{z}_{t,n_x} \) as in (64) |
| 6: end for |
| 7: for \( n_x = 1, \ldots, N_x - 1 \) do |
| 8: Use \( \rho_{t,2} \) to calculate \( \tilde{z}_{t,n_x} \) as in (65) |
| 9: end for |
| else if \( t > 1 \) then |
| 11: Initialization: \( \tilde{z}_{t,1} = \rho_{t,1}^{(T)} \left(1 - \rho_{t,1}^{\text{dyna}}\right) + \left(1 - \rho_{t,1}^{(T)}\right) \rho_{t,1}^{\text{dyna}},\tilde{z}_{t,N_e} = \frac{1}{2} \) |
| 12: for \( n_x = 2, \ldots, N_x \) do |
| 13: Calculate the \( \tilde{z}_{t,n_x} \) as in (64) |
| 14: end for |
| 15: for \( n_x = 1, \ldots, N_x - 1 \) do |
| 16: Calculate \( \tilde{z}_{t,n_x} \) as in (65) |
| 17: end for |
| 18: end if |
| 19: Calculate the \( k-th \) SU AoD azimuth support vector \( \rho_{k,t,n_x}^{\text{out}} = \rho_{k,t,n_x}^{\text{out}}(\theta_{k,t,n_x}^{(A)} \hat{a}_{k,n_x}^{(A)},\epsilon_{k,t,n_x}^{(A)} \hat{e}_{k,t,n_x}^{(A)} \hat{e}_{k,t,n_x}) \) |
| # Message Passing Over k-th SU Azimuth to Elevation Support: |
| 20: Calculate individual message \( \rho_{k,t,n_x}^{\text{out}} \) using (70) |
| 21: Calculate message \( v_{uk,t,n_x \rightarrow uk_{t,n_x}}(\theta_{k,t,n_x}) \) and \( v_{ek,t,n_x \rightarrow ek_{t,n_x}}(\epsilon_{k,t,n_x}) \) as in (69) and (71), respectively |
| 22: Calculate \( v_{ek,t,n_x \rightarrow ek_{t,n_x}}(\epsilon_{k,t,n_x}) \) as in (73) and \( v_{\pi_{k,t,n_x} \rightarrow \epsilon_{k,t,n_x}} \propto \sum_{\epsilon_{k,t,n_x}} \) as in (82) |
| # HMM k-th SU Elevation Support Estimation Based on 2D-MM MU Elevation: |
| 23: Similar to the procedures from step 2 to step 18 by using elevation related parameters, Details can be found from (76) to (79) |
| 24: Calculate the \( k-th \) SU AoD elevation support vector \( \rho_{k,t,n_y}^{\text{out}} = \rho_{k,t,n_y}^{\text{out}}(E) \) |

with
\[
(\mu_{k,t,m}^{(in)},\sigma_{k,t,m}^{(in)})^2 = \left\{ \begin{array}{ll}
\tilde{g}_{B,k,t,m}^{\text{out}} & \rho_{k,t,m} \leq T_h, \\
\tilde{g}_{B,k,t,m}^{\text{out}} + \epsilon_{k,t,m} > T_h,
\end{array} \right.
\]

where \( \epsilon \) is set as a small positive value close to 0. Then the message passing from factor node \( f_{k,t,1,m} \) to variable node
\( \vartheta_{k,t+1,m} \) is formed as
\[
\mathbb{C}N \left( \vartheta_{k,t+1,m}; \mu_{k,t+1,m}^{\text{dyna}}, \sigma_{k,t+1,m}^{\text{dyna}} \right)
\]
where
\[
\mu_{k,t+1,m}^{\text{dyna}} = (1 - \beta) \left( \frac{\sigma_{k,t,m}^{\text{dyna}}}{\sigma_{k,t,m}^{\text{in}}} + \frac{\sigma_{k,t,m}^{\text{in}}}{\sigma_{k,t,m}^{\text{dyna}}} \right) \left( \frac{\sigma_{k,t,m}^{\text{in}}}{\sigma_{k,t,m}^{\text{dyna}}} \right)
\]
\[
\times \left( \frac{\mu_{k,t,m}^{\text{dyna}}}{\sigma_{k,t,m}^{\text{in}}} + \frac{\mu_{k,t,m}^{\text{in}}}{\sigma_{k,t,m}^{\text{dyna}}} \right) + \beta \mu
\]
and
\[
\left( \sigma_{k,t+1,m}^{\text{dyna}} \right)^2 = (1 - \beta)^2 \left( \frac{\sigma_{k,t,m}^{\text{dyna}}}{\sigma_{k,t,m}^{\text{in}}} + \frac{\sigma_{k,t,m}^{\text{in}}}{\sigma_{k,t,m}^{\text{dyna}}} \right) + \beta^2 \zeta.
\]

2) The MU Azimuth Support \( \tilde{a}_i \) Passing Through Time:
The azimuth support prior information passing to the next time
\[
v_{\tilde{a}_i,n_x} \rightarrow v_{\tilde{a}_i,n_x} \left( \tilde{a}_i,n_x \right)
\]
can be given as
\[
v_{\tilde{a}_i,n_x} \rightarrow v_{\tilde{a}_i,n_x} \left( \tilde{a}_i,n_x \right) = \tilde{p}_{t+1,n_x}^{(A)} \delta \left( \tilde{a}_i,n_x - 1 \right) + \left( 1 - \tilde{p}_{t+1,n_x}^{(A)} \right) \delta \left( \tilde{a}_i,n_x \right),
\]
with
\[
\tilde{p}_{t+1,n_x}^{(A)} = \tilde{\gamma}_{t,n_x} \left( \tilde{\alpha}_{t+1,n_x} \right)^{\tilde{\beta}_{t+1,n_x}} \tilde{\rho}_{t,n_x}^{\tilde{\gamma}_{t,n_x}^{(A)} \tilde{\alpha}_{t+1,n_x}} + (1 - \tilde{\gamma}_{t,n_x}) \left( 1 - \tilde{\alpha}_{t+1,n_x} \right)^{\tilde{\beta}_{t+1,n_x}}
\]

3) The MU Elevation Support \( \tilde{e}_i \) Passing Through Time:
Similarly, the elevation support prior information passing to the next time can be written as
\[
v_{\tilde{e}_i,n_y} \rightarrow v_{\tilde{e}_i,n_y} \left( \tilde{e}_i,n_y \right) = \tilde{p}_{t+1,n_y}^{(E)} \delta \left( \tilde{e}_i,n_y - 1 \right) + \left( 1 - \tilde{p}_{t+1,n_y}^{(E)} \right) \delta \left( \tilde{e}_i,n_y \right),
\]
with
\[
\tilde{p}_{t+1,n_y}^{(E)} = \tilde{\gamma}_{t,n_y} \left( \tilde{\alpha}_{t+1,n_y} \right)^{\tilde{\beta}_{t+1,n_y}} \tilde{\rho}_{t,n_y}^{\tilde{\gamma}_{t,n_y}^{(E)} \tilde{\alpha}_{t+1,n_y}} + (1 - \tilde{\gamma}_{t,n_y}) \left( 1 - \tilde{\alpha}_{t+1,n_y} \right)^{\tilde{\beta}_{t+1,n_y}}
\]

D. 3D Off-grid Estimation

As there is no closed-form expression due to the complex MD-MM structure of \( \theta_k^{(t)} \), it’s hard to directly maximize the log-likelihood function \( \ln p \left( y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right) \) in (21). We extend the proposed in-exact majorization-minimization (MM) algorithm in [38] to help us solve the off-grid problem. As MM is the generalization of the expectation maximization method, the convergence of expectation maximization algorithm can prove that the in-exact MM can converge to a stationary solution [39]. Firstly, we iteratively construct the lower bound for \( \ln p \left( y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right) \) which can be named as surrogate function, i.e., for any fixed azimuth and elevation offset pairs \( \left( \Delta \theta_k^{(t)}, \Delta \phi_k^{(t)} \right) \), the surrogate function can be defined as [38]
\[
\mathcal{L} \left( \Delta \theta_k, \Delta \phi_k \right) = \int p \left( g_k, y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right) \ln p \left( g_k, y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right) dg_k
\]
which satisfies the following properties:
\[
\mathcal{L} \left( \Delta \theta_k, \Delta \phi_k \right) \Delta \theta_k, \Delta \phi_k \right) \leq \ln p \left( y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right), \forall \Delta \theta_k, \Delta \phi_k,
\]
\[
\mathcal{L} \left( \Delta \theta_k, \Delta \phi_k \right) \Delta \theta_k, \Delta \phi_k \right) = \ln p \left( y_k^{(t)}, \Delta \theta_k, \Delta \phi_k \right)
\]
Then \( \Delta \theta_k, \Delta \phi_k \) can be updated at each iteration as
\[
\Delta \theta_k^{(t+1)} = \arg \max_{\Delta \theta_k} \mathcal{L} \left( \Delta \theta_k, \Delta \phi_k^{(t+1)} \right)
\]
\[
\Delta \phi_k^{(t+1)} = \arg \max_{\Delta \phi_k} \mathcal{L} \left( \Delta \theta_k^{(t+1)}, \Delta \phi_k \right)
\]
with \((\cdot)^{(i)}\) represents for the \(i^{th}\) iteration update. To be noticed, for the \((i + 1)^{th}\) iteration elevation update, we use the updated \((i + 1)^{th}\) azimuth offset as fixed known knowledge in (48). However, the problems in (47) and (48) are non-convex. Hence the gradient update can be applied as follows:

\[
\begin{align*}
\Delta \theta_{k,t}^{(i+1)} &= \Delta \theta_{k,t}^{(i)} + \\
\Delta \phi_{k,t}^{(i)} &= \frac{\partial L}{\partial \Delta \theta_{k,t}}|_{\Delta \theta_{k,t}=\Delta \theta_{k,t}^{(i)}} \\
\Delta \phi_{k,t}^{(i+1)} &= \Delta \phi_{k,t}^{(i)} + \\
\Delta \Delta \phi_{k,t}^{(i)} &= \frac{\partial L}{\partial \Delta \phi_{k,t}}|_{\Delta \phi_{k,t}^{(i)}} \\
\end{align*}
\]

(49)

and

\[
\begin{align*}
\Delta \phi_{k,t}^{(i+1)} &= \Delta \phi_{k,t}^{(i)} + \\
\Delta \Delta \phi_{k,t}^{(i)} &= \frac{\partial L}{\partial \Delta \phi_{k,t}}|_{\Delta \phi_{k,t}^{(i)}} \\
\end{align*}
\]

(50)

where \(\Delta \theta^{(i)}\) and \(\Delta \phi^{(i)}\) are the stepsizes. The off-grid estimation algorithm is summarized in Algorithm 2 and the detailed equation can be found in Appendix C. Finally, the overall proposed MD-DTAMP algorithm can be summarized in Algorithm 3.

**VI. NUMERICAL RESULTS**

In this section, the complexity and the estimation performance of the proposed algorithm MD-DTAMP with and without 3D off-grid estimation are evaluated. We assume that the proposed algorithm is performed in the distributed UAVs. For the space-air link of each single UAV, the sparsity ratio for both azimuth and elevation directions is \(\lambda = 1/\sqrt{N}\) which indicates only one dominant LoS ray for each UAV-satellite link. The parameters are set as follows\(^3\): We assume that there are total 3 UAVs in the system. The MU azimuth and elevation spatial correlation parameters \(\{\rho_{0,1}, \rho_{1,0}, \rho_{0,0}, \rho_{1,1}, \rho_{1,0}\}\) = \(\{0.05, 0.05, 0.025, 0.025\}\), the MU azimuth and elevation temporal correlation parameters \(\{\rho_{0,1}, \rho_{1,0}, \rho_{0,0}, \rho_{1,1}, \rho_{1,0}\}\) = \(\{0.05, 0.05, 0.025, 0.175\}\), the MU azimuth and elevation joint parameters \(\{\rho_{0,1}, \rho_{1,0}, \rho_{0,0}, \rho_{1,1}, \rho_{1,0}\}\) = \(\{0.9958, 0.0013, 0.3276, 0.3936, 0.9956, 0.0007, 0.5, 0.1078\}\), the hidden value vector parameters for \(k^{th}\) UAV \(\{\mu_k, \sigma^2_k, \beta_k, \zeta_k, \gamma_k, \gamma_{k,t,1,1}, \gamma_{k,t,1,1}\}\) = \(\{0.1, 0.5, 0.5, 0.9375, 0.9375\}\), \(\mu = 1, \sigma = 10^{-2}, \epsilon = 10^{-7}, \gamma_{bk,t} = 1\), the azimuth range \(\text{range} = \pi\) and the elevation range \(\text{range} = \pi/4\). 3D Off-grid azimuth and elevation update step sizes are \(\Delta \theta^{(i)} = \pi/(100 \times M)\) and \(\Delta \phi^{(i)} = \pi/(4 \times 100 \times M)\).

To be noticed, in our simulation, our proposed algorithm MD-DTAMP with 3D off-grid estimation is presented in Algorithm 3. MD-DTAMP on grid which is Algorithm 3 without step 15. The benchmark algorithms which can be employed in our MU 3D channel model are sparse Bayesian learning (SBL) and D-OMP proposed in [38] and [17], respectively. All simulations are averaged over 500 times.

**A. Complexity Analysis**

The computational complexity of the LMMSE estimator in module \(A\) and MD-MM MMSE denoiser in module \(B\) are \(O(MN_S)\) and \(O(N_A + N_B)\), respectively. Hence, the total complexity of the proposed MD-DTAMP without 3D off-grid estimation over time \(T\) is \(O(T \times (MN_S + N_A + N_B))\) for each iteration. For off-grid estimation, with defined stepsize \(\Delta \theta^{(i)}\), the complexity for update azimuth offset \(\Delta \theta_{k,t}^{(i+1)}\) is \(O(N_S M^2)\) per iteration. The reason why the complexity is not \(O(N_S N^2)\) is that the original dimension has to be extended for paring the azimuth and elevation angles. Similarly, the complexity for update elevation offset \(\Delta \phi_{k,t}^{(i+1)}\) is also

---

\(^3\)The parameters can be trained through expectation maximization algorithm which can refer to [30] for details. The MU azimuth and elevation related parameters for multiple UAVs system in this paper are referred 2D settings in [20].
achieve a solid performance by taking the dynamic channel pilot overhead. This proves that our proposed algorithms can maintain a good performance with smaller offset.

Improve the performance through gradient update of the 3D tracking performance and the training pilot overhead. It is easy to notice that our proposed algorithm with or without off-grid estimation consumes less computation time for benchmark algorithms. Moreover, the off-grid can further improve the performance through gradient update of the 3D offset.

In Fig. 9, we further consider the relationship between the tracking performance and the training pilot overhead. It is easy to be noticed that our proposed algorithm with or without off-grid estimation can maintain a good performance with smaller pilot overhead. This proves that our proposed algorithms can achieve a solid performance by taking the dynamic channel spatial and temporal correlation into consideration.

To be notice, empirically, the proposed algorithm normally converges within 30 iterations.

B. Simulation Results

We choose time-averaged normalized mean square error (TNMSE) as the performance metric as [30]

$$\text{TNMSE} = \frac{1}{T} \sum_{t}^{T} \frac{\|g_t - \hat{g}_t\|^2}{\|g_t\|^2},$$

where $\hat{g}_t$ is the estimated result of $g_t$ in time $t$. This performance metric is to evaluate the recovery performance of various algorithms.

In Fig. 8, we further compare the TNMSE performance metric versus SNR among various algorithms with antenna array elements set as $M = 64, 256$. With SNR increases, all four algorithms show better recovery performance. To be noticed, our proposed MD-DTAMP on-grid and off-grid tracking algorithms always show superior performance compared to benchmark algorithms. Moreover, the off-grid can further improve the performance through gradient update of the 3D offset.

In Fig. 9, we further consider the relationship between the tracking performance and the training pilot overhead. It is easy to be noticed that our proposed algorithm with or without off-grid estimation can maintain a good performance with smaller pilot overhead. This proves that our proposed algorithms can achieve a solid performance by taking the dynamic channel spatial and temporal correlation into consideration.

$\mathcal{O}(N_S M^2)$ per iteration$^4$.

Fig. 7 shows the computation time of various algorithms versus the number of antennas. It can be observed that the proposed on-grid MD-DTAMP shows comparable complexity compared with D-OMP. Furthermore, compared with SBL algorithm which requires matrix multiplication and inversion with high computation complexity, the proposed MD-DTAMP with off-grid estimation consumes less computation time for different number of antennas.
The proposed on-grid and off-grid algorithms are compared in Fig. 10. The off-grid estimation achieves sufficient performance gain over the on-grid algorithm. In Fig. 10 (a), the performance over continuous total time slot is compared. It is obvious that, with time increases, the performance show superior gain for both on-grid and off-grid algorithms. Moreover, it can be observed from Fig. 10 (a) and (b) that with $M$ increases from 64 to 256, the tracking performance also improves. This is because with a larger antenna array and denser virtual grid, and the on-grid 3D tracking angle can be estimated with a narrower grid and closer to the true angle.

VII. CONCLUSION

This paper exploits the on-grid and off-grid 3D channel modelling and tracking for space-air communications with multiple UAVs and an LEO satellite in the system. The relative displacement between the orbit movement of the LEO satellite and the 3D trajectory of multiple UAVs was modelled based on the statistical dynamic channel model called MD-MM. The proposed MD-MM was composed of the insight probability relationship of the spatial and temporal correlation of the MU hidden support vector, SU joint hidden support vector, and SU hidden value vector. We further developed a novel MD-DTAMP algorithm to track the dynamic channel recursively. Furthermore, the gradient update for the 3D off-grid estimation scheme was proposed, which can iteratively find the azimuth and elevation offset. In addition, we compared the computation complexity between benchmark algorithms and the proposed algorithm with and without the off-grid scheme. Simulations verified that the proposed algorithm derived from the MD-MM dynamic channel model can better exploit the more realistic space-air 3D dynamic channel to achieve more accurate recovery performance over other benchmark algorithms with smaller pilot overhead.

ACKNOWLEDGEMENTS

The work in this paper was supported by the Engineering and Physical Sciences Research Council in the United Kingdom, under the Grant EP/R00711X/2.

APPENDIX A

TRAINING PILOT MATRIX DESIGN

It has been proved that a partial DFT sensing matrix outperforms the independent identically distributed sensing matrix in solving the CS problem with AMP algorithm [11], [32]. Hence, in our proposed model, the sensing matrix $\Phi_{k,t} (\Delta \theta_{k,t}, \Delta \phi_{k,t})$ is referred to be a partial DFT random permutation matrix as $\Phi_{k,t} (\Delta \theta_{k,t}, \Delta \phi_{k,t}) \approx S_{SEL} S_{DFT} S_{RAND}$, where $S_{SEL}$ is a selection matrix consisting of randomly selected and reordered rows of $N_S \times M$ identity matrix, $S_{DFT}$ is the $M \times M$ DFT matrix, and $S_{RAND}$ is random permutation matrix generated by a randomly reordered $M \times M$ identity matrix. For large $M$ antenna elements array, $S_{SEL} S_{DFT} S_{RAND} = S_{DFT} \left( \tilde{A}^H \left( \tilde{\theta}_{k,t} + \Delta \theta_{k,t}, \tilde{\phi}_{k,t} + \Delta \phi_{k,t} \right) \right)$. Hence, for simplicity with fixed on-grid and approximate partial orthogonal, the pilot matrix $S_{k,t}^T = S_{SEL} S_{DFT} S_{RAND} \left( \tilde{A}^H \left( \tilde{\theta}_{k,t} + \Delta \theta_{k,t}, \tilde{\phi}_{k,t} + \Delta \phi_{k,t} \right) \right)^{-1}$.

APPENDIX B

MD-MM MESSAGE PASSING IN MODULE B

1) Message Passing Over Gaussian Hidden Value Vector $\tilde{\theta}_{k,t}$: The message $\nu_{k,t,m} \rightarrow \tilde{\theta}_{k,t,m}$ passed from the last time slot can be formed as $CN(\tilde{\theta}_{k,t,m}; \mu_{k,t,m}^{\text{dyna}}, \sigma_{k,t,m}^{\text{dyna}})$, and $\tilde{\theta}_{k,t,m}$ followed by a randomly reordered $\pi_{k,t,m}$. Moreover, it is obvious that, with time increases, the performance show superior gain for both on-grid and off-grid algorithms. Moreover, it can be observed from Fig. 10 (a) and (b) that with $M$ increases from 64 to 256, the tracking performance also improves. This is because with a larger antenna array and denser virtual grid, and the on-grid 3D tracking angle can be estimated with a narrower grid and closer to the true angle.

$\nu_{k,t,m} \rightarrow \tilde{\theta}_{k,t,m} = CN(\tilde{\theta}_{k,t,m}; \mu_{k,t,m}^{\text{out}}, \sigma_{k,t,m}^{\text{out}})$, (52)

with $(\mu_{k,t,m}^{\text{out}}, \sigma_{k,t,m}^{\text{out}}) = (\mu_{k,t,m}, \sigma_{k,t,m}^{\text{out}})$. When $t = 1$, we set $(\mu_{1,m}, \sigma_{1,m}^{\text{out}}) = (\mu, \sigma^2)$

2) Message Passing Over $g_k,t,m \rightarrow \pi_{k,t,m}$: The message from channel $g_k,t,m$ to factor node $\pi_{k,t,m}$ is $\nu_{g_k,t,m} \rightarrow \pi_{k,t,m} (g_k,t,m) = CN(g_k,t,m; g_{B k,t,m}^{\text{pri}}, \mu_{B k,t,m}^{\text{pri}}, v_{B k,t,m}^{\text{pri}})$. (53)

The message from factor node $\pi_{k,t,m}$ to variable node $b_k,t,m$ is $\nu_{\pi_{k,t,m} \rightarrow b_k,t,m} (b_k,t,m) = \rho_{b_{k,t,m}}^{\text{in}} \delta (b_k,t,m - 1) + (1 - \rho_{b_{k,t,m}}^{\text{in}}) \delta (b_k,t,m)$, (54)

where $\rho_{b_{k,t,m}} = \frac{CN (0; \mu_{B k,t,m}^{\text{pri}}, v_{B k,t,m}^{\text{pri}})}{CN (0; \mu_{B k,t,m}^{\text{out}}, v_{B k,t,m}^{\text{out}} + (\sigma_{B k,t,m}^{\text{out}})^2)}$. (55)

The message from factor node $\pi_{k,t,m}$ to variable node $b_k,t,m$ from variable node $u_k,t,m$ to factor node $u_k,t,m$ is the same as $\nu_{\pi_{k,t,m} \rightarrow b_k,t,m} (b_k,t,m)$. The message from factor node $u_k,t,m$ to variable node $a_k,t,n_z$ is $\nu_{u_k,t,m} \rightarrow a_k,t,n_z (a_k,t,n_z) = \rho_{a_{k,t,n_z}}^{\text{in}} \delta (a_k,t,n_z - 1) + (1 - \rho_{a_{k,t,n_z}}^{\text{in}}) \delta (a_k,t,n_z)$, (56)

where $\rho_{a_{k,t,n_z}}^{\text{in}} = \frac{\rho_{a_{k,t}}}{1 + \rho_{a_{k,t}} - \rho_{b_{k,t,m}}}$, (57)

with $\rho_{a_{k,t}} = \gamma_{b_{k,t,m}} \rho_{b_{k,t,m}}^{\text{out}} + (1 - \gamma_{b_{k,t,m}}) (1 - \rho_{b_{k,t,m}}^{\text{in}})$

$\rho_{b_{k,t,m}}^{\text{out}} + (1 - \rho_{b_{k,t,m}}^{\text{in}}) (1 - \rho_{b_{k,t,m}}^{\text{out}})$

and $\rho_{c_{k,t,n_z}}^{\text{out}} = \nu_{u_k,t,n_z} \rightarrow u_k,t,n_z (c_{k,t,n_z}) = 1$.

3) Message Passing Over $a_k,t,n_z \rightarrow \tilde{\theta}_{k,t,n_z} \rightarrow \tilde{a}_{t,n_z}$, the 2D-MM of the MU Azimuth Support $\tilde{\theta}_{k,t,n_z}$, and $\tilde{a}_{t,n_z} \rightarrow \tilde{\phi}_{k,t,n_z} \rightarrow a_k,t,n_z$: The forward-backward message
passing is performed over the HMM $a_{k,t}$ for given input messages $u_{k,t,m,n_{y}} \rightarrow a_{k,t,n_{x}}$.

We firstly introduce the details of the MU azimuth support estimation. The condition can be divided into two section, one is when $t = 1$, the other is when $t > 1$. To be notice, the message from factor node $u_{t,m}$ to variable node $a_{t,n_{x}}$ is

$$ u_{k,t,m} \rightarrow a_{k,t,n_{x}} = \frac{\rho^{(A)in}_{a_{t,k},t,n_{x},n_{y}}}{\prod_{n_{y}} \rho^{in}_{a_{t,k},t,n_{x},n_{y}}} ,$$

(59)

which is derivated based on the matched dimension $M$ from the total gain, $n_{x}$ from the azimuth support and $n_{y}$ from the elevation support.

The message from factor node $\phi_{k,t,n_{x}}$ to variable node $\tilde{a}_{t,n_{x}}$ is

$$ u_{k,t,n_{x}} \rightarrow \tilde{a}_{t,n_{x}} = \sum_{a_{k,t,n_{x}}} \prod_{n_{x}} u_{k,t,m} \rightarrow a_{k,t,n_{x}} (a_{k,t,n_{x}}) \zeta^{(A)}_{k,t,n_{x}} a_{t,n_{x}} = \rho^{in}_{a_{t,k},t,n_{x},n_{y}} \delta (\tilde{a}_{t,n_{x}} - 1) + (1 - \rho^{in}_{a_{t,k},t,n_{x},n_{y}}) \tilde{a}_{t,n_{x}} ,$$

(60)

where

$$ \rho^{in}_{a_{t,k},t,n_{x},n_{y}} = \frac{(A)_{k,t,n_{x}} (A)_{k,t,n_{x}} \rho^{in}_{k,t,n_{x},n_{y}} + (1 - (A)_{k,t,n_{x}}) (1 - \rho^{in}_{k,t,n_{x},n_{y}})}{(A)_{k,t,n_{x}} (A)_{k,t,n_{x}} + (1 - (A)_{k,t,n_{x}}) (1 - \rho^{in}_{k,t,n_{x},n_{y}})} ,$$

(61)

When $t = 1$, the MU azimuth forward parameter is

$$ z^{(A)f}_{l,t,n_{x}} = \frac{\zeta^{(A)f}_{l,t,n_{x}} (\zeta^{(A)f}_{l,t,n_{x}}) (\zeta^{(A)f}_{l,t,n_{x}}) (\zeta^{(A)f}_{l,t,n_{x}})}{\rho_{l,t,n_{x}} \rho_{l,t,n_{x}} + \prod_{n_{y}} (1 - \rho^{in}_{a_{t,k},t,n_{x},n_{y}})} ,$$

(62)

where

$$ \rho_{l,t,n_{x}} = \frac{\prod_{n_{y}} \rho^{in}_{a_{t,k},t,n_{x},n_{y}} + \prod_{n_{y}} (1 - \rho^{in}_{a_{t,k},t,n_{x},n_{y}})}{\prod_{n_{y}} \rho^{in}_{a_{t,k},t,n_{x},n_{y}} + \prod_{n_{y}} (1 - \rho^{in}_{a_{t,k},t,n_{x},n_{y}})} .$$

The azimuth backward parameter $z^{(A)b}_{l,n_{x}}$ can be found in (63). When $t > 1$, the azimuth forward parameter $z^{(A)f}_{l,t,n_{x}}$ can be found in (64). The final MU azimuth message which represents the non-zero probability can be written as

$$ u_{k,t,n_{x}} \rightarrow a_{k,t,n_{x}} = \rho^{out}_{a_{t,k},t,n_{x},n_{y}} \delta (\tilde{a}_{t,n_{x}} - 1) + (1 - \rho^{out}_{a_{t,k},t,n_{x},n_{y}}) \tilde{a}_{t,n_{x}} ,$$

(64)

where $\rho^{out}_{a_{t,k},t,n_{x},n_{y}}$ can be seen in (65). The message from factor node $\phi_{k,t,n_{x}}$ to variable node $a_{k,t,n_{x}}$ is

$$ u_{k,t,n_{x}} \rightarrow a_{k,t,n_{x}} (a_{k,t,n_{x}}) = \rho^{out}_{k,t,n_{x}} (\delta (a_{k,t,n_{x}} - 1) + (1 - \rho^{out}_{a_{t,k},t,n_{x},n_{y}}) \rho^{out}_{a_{t,k},t,n_{x},n_{y}} a_{k,t,n_{x}} ) ,$$

(65)

where $\rho^{out}_{k,t,n_{x}} = \prod_{n_{y}} (1 - \rho^{in}_{a_{t,k},t,n_{x},n_{y}})$.
backward parameter $z_{t,n_y}^{(E) b}$ can be found in $(77)$. When $t > 1$, the azimuth forward parameter $z_{t,n_y}^{(E) f}$ can be found as

$$z_{t,n_y}^{(E) f} = \frac{\gamma_{t,n_y}^{(E) dyna}}{1 + \left(\gamma_{t,n_y}^{(E) dyna} - \gamma_{t,n_y}^{(E) in}\right)^{-1} \left(\gamma_{t,n_y}^{(E) dyna} - \gamma_{t,n_y}^{(E) in}\right)^{-1}} + \gamma_{t,n_y}^{(E) dyna} + \gamma_{t,n_y}^{(E) in} \left(1 - \gamma_{t,n_y}^{(E) dyna} - \gamma_{t,n_y}^{(E) in}\right)^{-1} \left(\gamma_{t,n_y}^{(E) dyna} - \gamma_{t,n_y}^{(E) in}\right)^{-1} + 1.$$  

(78)

The azimuth backward parameter $z_{t,n_y}^{(E) b}$ is given in $(79)$. The final MU elevation message which represents the non-zero probability can be written as

$$\psi_{t,n_y}^{\nu, \pi, e_{k,t,n_y}} = \rho_{e_{k,t,n_y}}^{\nu, \pi, e_{k,t,n_y}} \delta(e_{k,t,n_y} - 1) + (1 - \rho_{e_{k,t,n_y}}^{\nu, \pi, e_{k,t,n_y}}) \delta(e_{k,t,n_y}),$$  

(80)

where $\rho_{e_{k,t,n_y}}^{\nu, \pi, e_{k,t,n_y}}$ can be seen in $(81)$. The message from factor node $\pi_{k,t,n_y}$ to variable node $e_{k,t,n_y}$ is

$$\psi_{e_{k,t,n_y}}^{\nu, \pi, e_{k,t,n_y}} = \rho_{e_{k,t,n_y}}^{(E) out} \delta(e_{k,t,n_y} - 1) + (1 - \rho_{e_{k,t,n_y}}^{(E) out}) \delta(e_{k,t,n_y}),$$  

(82)

where $\rho_{e_{k,t,n_y}}^{(E) out} = \rho_{e_{k,t,n_y}}^{(E) out} \pi_{k,t,n_y}$.

6) **Message Passing Over the Path** $e_{k,t,n_y} \rightarrow u_{k,t,m} \rightarrow b_{k,t,m} \rightarrow \pi_{k,t,m} \rightarrow g_{k,t,m}$. The message from factor node $e_{k,t,n_y}$ to variable node $u_{k,t,m}$ is

$$\psi_{u_{k,t,m}}^{\nu, e_{k,t,n_y}} = \rho_{u_{k,t,m}}^{out} \delta(e_{k,t,n_y} - 1) + (1 - \rho_{u_{k,t,m}}^{out}) \delta(e_{k,t,n_y}),$$  

(83)

where $\rho_{u_{k,t,m}}^{out} = \rho_{u_{k,t,m}}^{out} \pi_{k,t,n_y}$. The message from factor node $u_{k,t,m}$ to variable node $b_{k,t,m}$ is

$$\nu_{u_{k,t,m}}^{\nu, b_{k,t,m}} = \rho_{b_{k,t,m}}^{out} \delta(b_{k,t,m} - 1) + (1 - \rho_{b_{k,t,m}}^{out}) \delta(b_{k,t,m}),$$  

(85)

where $\rho_{b_{k,t,m}}^{out} = \rho_{b_{k,t,m}}^{out} \pi_{k,t,m}$.

The azimuth offset $\Delta_{k,t}^{(i+1)}$ can be updated use $(49)$. The derivative $^5$To be noticed, in order to have azimuth and elevation one-one correspondence, we extend the original $N_x$ dimension of offset into $M$ dimension with azimuth and elevation pairing. The elevation offset can be written in the similar way.

\[ \Delta_{k,t}^{(i+1)} = \begin{cases} \frac{\partial L}{\partial \Delta_{k,t}^{(i)}} & \text{if } i \leq T \text{ and } k \neq 1, \\ 0 & \text{otherwise}. \end{cases} \]

\[ \frac{\partial L}{\partial \Delta_{k,t}^{(i)}} = \sum_{t=1}^{T} \left( y_{k,t} - \Phi_{k,t}(\Delta_{k,t}^{(i)}, \Phi_{k,t}) g_{k,t} \right) \Phi_{k,t}^{T} \Sigma^{-1} \Phi_{k,t}. \]

(87)

\[ \Delta_{k,t}^{(i+1)} = \begin{cases} \frac{\partial L}{\partial \Delta_{k,t}^{(i)}} & \text{if } i \leq T \text{ and } k \neq 1, \\ 0 & \text{otherwise}. \end{cases} \]

\[ \frac{\partial L}{\partial \Delta_{k,t}^{(i)}} = \sum_{t=1}^{T} \left( y_{k,t} - \Phi_{k,t}(\Delta_{k,t}^{(i)}, \Phi_{k,t}) g_{k,t} \right) \Phi_{k,t}^{T} \Sigma^{-1} \Phi_{k,t}. \]
The article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3121301, IEEE Transactions on Wireless Communications.
channel and covariance estimation in mmwave networks,” IEEE Trans. Commun., vol. 68, no. 6, pp. 3906–3921, 2020.

[24] H. Fenech, S. Amos, and T. Waterfield, “The role of array antennas in commercial telecommunication satellites,” in 2016 10th European Conference on Antennas and Propagation (EuCAP), Apr. 2016, pp. 1–4.

[25] Y. Xu and C. Ruan, “A novel design of circularly polarized omnidirectional antenna for Ka-Band,” in 2008 Global Symposium on Millimeter Waves, 2008, pp. 378–379.

[26] Güngör and A. Ünal, “Design of a vertically polarized omni-directional antenna at Ka-Band,” in 2016 IEEE International Symposium on Antennas and Propagation (APSURSI), 2016, pp. 129–130.

[27] J. Zhao, F. Gao, Q. Wu, S. Jin, Y. Wu, and W. Jia, “Beam tracking for edge computing paradigms,” IEEE J. Sel. Areas Commun., vol. 36, no. 2, pp. 363–375, Feb. 2018.

[28] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, “Channel estimation with arbitrary 2D-array geometry,” IEEE Trans. Signal Process., vol. 61, no. 21, pp. 5270–5284, Nov. 2013.

[29] J. Zhao, F. Gao, Q. Wu, S. Jin, Y. Wu, and W. Jia, “Beam tracking for wireless communication optimization,” IEEE Trans. Signal Process., vol. 68, pp. 4379–4393, 2020.

[30] Y. Gao and J. L. Taylor, “Multi-drone for wireless communication optimization,” IEEE Trans. Signal Process., vol. 61, no. 9, pp. 2388–2399, May 2013.

[31] J. Zhao, F. Gao, Q. Wu, S. Jin, Y. Wu, and W. Jia, “Beam tracking for edge computing and machine learning,” in 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), 2019, pp. 487–491.

[32] Z. Xue, J. Ma, and X. Yuan, “Denoising-based turbo compressed sensing,” IEEE Access, vol. 5, pp. 7193–7204, 2017.

[33] J. Dai, A. Liu, and V. K. N. Lau, “FDD massive MIMO channel estimation with arbitrary 2D-array geometry,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2584–2599, May 2018.

[34] C. J. Wu, “On the convergence properties of the EM algorithm,” The Annals of statistics, pp. 95–103, 1983.

Jiadong Yu received her B.S in Communication Engineering from Dalan Maritime University, and her M.S in Wireless Communication from University of Southampton with distinction. She is pursuing her Ph.D. degree with the Department of Electronic Engineering and Computer Science in Queen Mary University of London. Her research interests include machine learning for millimeter wave communications and Internet of Things.

Xiaolan Liu received the PhD degree from Queen Mary University of London, U.K. in 2021. She was a research assistant in King’s College London from 2020-2021. Since Oct. 2021, She has been with the Institute for Digital Technologies, Loughborough University London, U.K., where she is currently a lecturer. Her current research interests include wireless distributed learning, multi-agent reinforcement learning for edge computing and machine learning for wireless communication optimization.

Yue Gao (S’03–M’07–SM’13) is a Professor of Wireless Communications at Institute for Communication Systems, University of Surrey, United Kingdom. He received the Ph.D. degree from the Queen Mary University of London (QMUL), U.K., in 2007. Dr Gao was a Lecturer, Senior Lecturer and Reader in Antennas and Signal Processing with QMUL. He currently leads a team developing fundamental research into practice in the interdisciplinary area of smart antennas, signal processing, spectrum sharing, millimetre-wave and Internet of Things technologies in mobile and satellite systems. He has published over 180 peer-reviewed journal and conference papers, 3 patents, 1 book and 5 book chapters and 3 best paper awards. He is an Engineering and Physical Sciences Research Council Fellow from 2018 to 2023. He was a co-recipient of the EU Horizon Prize Award on Collaborative Spectrum Sharing in 2016, and shortlisted for the Newton Prize on IoT systems for smart farming in 2019.

He served as the Signal Processing for Communications Symposium Co-Chair for IEEE ICC 2016, the Publicity Co-Chair for the IEEE GLOBECOM 2016, the Cognitive Radio Symposium Co-Chair for the IEEE GLOBECOM 2017, and the General Chair of the IEEE WoWMoM and IWEM 2017. He is the Chair of the IEEE Technical Committee on Cognitive Networks, the Secretary of the IEEE ComSoc Technical Committee Wireless Communication and the IEEE Distinguished Lecturer of the Vehicular Technology Society. He is an Editor for the IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and IEEE TRANSACTIONS ON COGNITIVE NETWORKS.
Xuemin (Sherman) Shen (M’97–SM’02–F’09) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on network resource management, wireless network security, Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.

Dr. Shen received the R.A. Fessenden Award in 2019 from IEEE, Canada, Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) presents in 2019, James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society, and Technical Recognition Award from Wireless Communications Technical Committee (2019) and AHSN Technical Committee (2013). He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier’s Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He served as the Technical Program Committee Chair/Co-Chair for the IEEE Globecom’16, the IEEE Infocom’14, the IEEE VTC’10 Fall, the IEEE Globecom’07, the Symposium Chair for the IEEE ICC’10, and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications. Dr. Shen is the elected IEEE Communications Society Vice President for Technical Educational Activities, Vice President for Publications, Member-at-Large on the Board of Governors, Chair of the Distinguished Lecturer Selection Committee, Member of IEEE Fellow Selection Committee. He was/is the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, IEEE Network, IET Communications, and Peer-to-Peer Networking and Applications.