3D Channel Tracking for UAV-Satellite Communications in Space-Air-Ground Integrated Networks

Jiadong Yu, *Graduate Student Member, IEEE*, Xiaolan Liu, *Graduate Student Member, IEEE*, Yue Gao, *Senior Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*— The space-air-ground integrated network (SAGIN) aims to provide seamless wide-area connections, high throughput and strong resilience for 5G and beyond communications. Acting as a crucial link segment of the SAGIN, unmanned aerial vehicle (UAV)-satellite communication has drawn much attention. However, it is a key challenge to track dynamic channel information due to the low earth orbit (LEO) satellite orbiting and three-dimensional (3D) UAV trajectory. In this paper, we explore the 3D channel tracking for a Ka-band UAV-satellite communication system. We firstly propose a statistical dynamic channel model called 3D two-dimensional Markov model (3D-2D-MM) for the UAV-satellite communication system by exploiting the probabilistic insight relationship of both hidden value vector and joint hidden support vector. Specifically, for the joint hidden support vector, we consider a more realistic 3D support vector in both azimuth and elevation direction. Moreover, the spatial sparsity structure and the time-varying probabilistic relationship between degree patterns named the spatial and temporal correlation, respectively, are studied for each direction. Furthermore, we derive a novel 3D dynamic turbo approximate message passing (3D-DTAMP) algorithm to recursively track the dynamic channel with the 3D-2D-MM priors. Numerical results show that our proposed algorithm achieves superior channel tracking performance to the state-of-the-art algorithms with lower pilot overhead and comparable complexity.**

***Index Terms*— Channel tracking, massive antenna array, message passing, UAV-satellite communication.**

I. Introduction

THE space-air-ground integrated network (SAGIN) [1] has drawn dramatic attention for its advantages such as seamless wide-area connections, high throughput, and strong resilience for 5G and beyond communications [2]. There are mainly three segments in the SAGIN multidimensional network: space segment with satellite network, air segment with aerial network and ground segment with terrestrial network [3]. Due to the satellite-ground links are very sensitive to terrestrial interference [4], it's necessary to use UAVs act as relays to improve the performance of communications with enhanced capacity for covered areas. Compared to traditional ground or satellite networks, unmanned aerial vehicle (UAV)-aided communication as a key part of SAGIN can be applied not only in the scenario that ground infrastructure cannot be accessed but also for emergency communications in crowded or disaster areas. Besides, with UAVs deployed in the middle of ground-space communications, shorter range line of sight (LoS) links provide a significant performance improvement over long-distance links [5]. For UAV-satellite links, to mitigate the problem of the spectrum crowding, millimeter wave (mmWave) Ka-Band (26.5-36GHz) offers additional frequency ranges with higher data rates [6], [7]. Although Ka-band is much more weather dependent due to the high path loss, an antenna array with a large number of antenna elements provides higher transmission gain and compensates the propagation losses [8]. In the above-mentioned scenarios, accurate channel estimation with a small pilot overhead is a crucial challenge to both UAV-satellite links (e.g., low earth orbit (LEO) and medium earth orbit (MEO) satellites [9]) and air-to-ground links over dynamic wireless channels due to the satellite orbiting and UAV three-dimensional (3D) trajectory [10], [11].

A. Related Work

For conventional channel estimation methods, such as least squares (LS) and minimum mean square error (MMSE), the number of the training pilots must be at least the number of antenna elements [12], which is impractical for large antenna array systems such as thousands of antenna elements in a mmWave massive MIMO system. By exploiting the limited scatterers of propagation environment, a lot of research works have developed different compressive sensing (CS) based algorithms, such as matching pursuit (MP) and orthogonal matching pursuit (OMP), to reduce the training overhead by exploring static and sparse nature of channels [13].

Apart from naively applying the conventional compressive sensing algorithms to solve the channel estimation problem, several structured sparse channel estimation algorithms have been developed by learning spatial correlation. The spatial

Manuscript received December 1, 2019; revised March 15, 2020; accepted April 28, 2020. Date of publication July 1, 2020; date of current version November 20, 2020. This work was supported by the Engineering and Physical Sciences Research Council, U.K., under Grant EP/R00711X/2. *(Corresponding author: Yue Gao.)*

Jiadong Yu and Xiaolan Liu are with the Department of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: jiadong.yu@qmul.ac.uk; xiaolan.liu@qmul.ac.uk).

Yue Gao is with the Department of Electrical and Electronic Engineering, University of Surrey, Surrey GU2 7XH, U.K. (e-mail: yue.gao@surrey.ac.uk).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2020.3005490

correlation is a feature that non-zero elements of an angular domain channel are concentrated in several bursts [14]–[17]. In [14], a joint orthogonal matching pursuit algorithm which exploited the hidden joint channel sparsity among multiple users was proposed with improved performance. Moreover, the sparsity of channel supports for the neighbouring antennas were similar as shown in [15]. Additionally, by obtaining the partial knowledge of the physical scattering structure channel, the proposed least absolute shrinkage and selection operator algorithm showed robust recovery performance [16]. In [17], the structured spatial domain sparsity was modeled as a Markov prior by considering the relationship of channel support vectors. However, all of these investigations were based on the static channel environment which could not be applied for the mobile scenario.

Different from the static sparse channel estimation, channel tracking was realized by considering prior information from the previously estimated channel. This temporal correlation can help to track the channel information with further reduced training overhead and dramatically improved real-time estimation performance [18]–[20]. In [18], by updating the channel variation based on the temporal correlation over time, the proposed differential OMP (D-OMP) tracked the dynamic sparse channels with improved accuracy. Similarly, [19] proposed a method called structured matching pursuit which simultaneously detected the common and the dynamic channel taps by looking at the insight of path delays in each time slot. By further excavating the physical direction between the base station and mobile user, and utilizing the obtained channel information from the previous time slot, a priori-aided tracking method was proposed in [20] to predict the information of beamspace channels for the next time slot.

In order to have better channel tracking performance, it is necessary to jointly consider both spatial correlation and temporal correlation of the dynamic channel [21]–[23]. The works in [21] and [22] decoupled the time-varying channel estimation procedure into angles of arrival/departure (AoAs/AoDs) and path gains estimation with [21] utilized unscented Kalman filter. Reference [22] explored the block-sparse problem of the angle, separately. Additionally, to accurately formulate the more realistic dynamic model of AoDs support in the flat surface over time, spatial sparsity structure and temporal transition probability during the dynamic evolution of the channel were captured in [23].

B. Motivation and Contribution

As discussed above, the channel estimation or tracking methods in [14]–[23] that consider the local ground environment characteristics can be directly applied to the air-to-ground links. The most widely considered mobile scenario is vehicle-to-vehicle (V2V) communications. Reference [24] has briefly summarized three different mmWave V2V channel modeling: stochastic channel models, geometry-based channel models, and deterministic channel models. Stochastic channel models detailed statistics such as delay, Doppler and angular domains. Geometry-based models [25] considered cluster number, position and nonisotropic scattering. Deterministic modeling addressed the issues of sparse multipath components. However, as another segment of the SAGIN, the UAV-satellite communication has some distinct properties: the high non-stationary channel condition and the dominant LoS ray [2]. Due to the limited scatterers of ground propagation environment, the angular domain structured sparsity of channels has been studied in [14]–[17]. This property was similar to UAV-satellite communication with one concentrate ray. However, in practice, it is non-trivial to obtain the real-time channel information due to the orbiting of the satellite and 3D UAV trajectories [26], which impose a formidable challenge on modeling and tracking the dynamic channel. By utilizing the advantage of the prior information of the channel, time-varying channel tracking has been considered in [18]–[20]. However, all these tracking algorithms have ignored the dynamic angular domain spatial correlation. Although both spatial and temporal correlation were explored in [21]–[23] for channel tracking, the premise of the system was that the communication was transmitted over the same flat surface. Additionally, the tracking methods mentioned above [18]–[22] lacks the consideration of the dynamic probability of the model. Although the spatial sparsity structure has been introduced in [23], the insight of a more complicated 3D dynamic spatial sparsity has not been investigated.

Motivated by the aforementioned literature review, to capture a more realistic UAV-satellite dynamic channel information, we consider a fading channel as dynamic probability model and explore the 3D dynamic angle domain sparsity structure over time in this paper. We study the UAV-satellite communication system with UAVs equipped with a uniform plane array (UPA) and propose a method to recursively track the 3D dynamic channel with lower pilot overhead, comparable complexity, and better recovery performance. The main contributions are summarized as follows:

- We develop a statistical 3D dynamic channel model for UAV-satellite communications called 3D space two-dimensional Markov Model (3D-2D-MM). The 3D-2D-MM captures the structured sparsity in both azimuth and elevation spatial domains as well as probabilistic temporal correlation.
- The dynamic spatial domain transition probabilities are designed for real-world UAV-satellite communication channel modeling to replace the expression of the complex continuous relative displacement between the 3D UAV trajectory and orbital satellite.
- We propose a novel 3D dynamic turbo approximate message passing (3D-DTAMP) algorithm which recursively tracks the 3D dynamic channel by exploiting the relationship of both azimuth and elevation angle supports.
- The proposed 3D-DTAMP algorithm achieves better reconstruction performance with lower pilot overhead and comparable complexity.

This paper is organized as follows. The UAV-satellite system model is introduced in Section II. In Section III, the 3D-2D-MM channel model is elaborated and the problem of channel tracking is formulated. The proposed algorithm 3D-DTAMP is presented in Section IV. Numerical results and conclusions are then provided in Section V and Section VI, respectively.

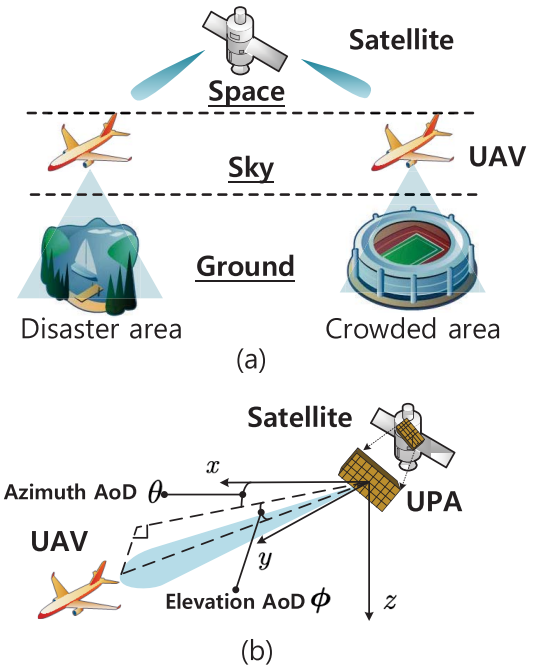

Fig. 1. (a) The space-air-ground integrated network. (b) The illustration of the UAV-satellite system model.

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 $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ respectively. $\boldsymbol{I}$ is the identity matrix. $\mathcal{CN}$ (mean, covariance) indicates a complex Gaussian random vector with defined mean and covariance. $\|\cdot\|$ is the l_2 norm.

II. System Model

The SAGIN and the proposed UAV-satellite models are in Fig. 1. Important notations are summarized in Table I. We consider a narrow-band channel model with OFDM transmission through a UAV-satellite communication downlink. Since array antennas have been widely employed on commercial satellites to improve the spectral efficiency by generating directional beams [27], the assumption in our model is that the LEO satellite is installed with uniform planar array (UPA) $M = N_x \times N_y$ antennas. Additionally, similar as in [28], a single omnidirectional antenna is implemented on UAVs. For the training procedure, the symbols $\boldsymbol{S} \in \mathbb{C}^{M\times N_{\mathrm{S}}}$ are transmitted. Then, the received signal at the UAV can be expressed as

$$\boldsymbol{y}_t^{\mathrm{RX}} = \boldsymbol{H}_t \boldsymbol{S} + \boldsymbol{n}_t^{\mathrm{RX}}, \tag{1}$$

and $\boldsymbol{n}_t^{\mathrm{RX}} \sim \mathcal{CN}(0, \sigma_n^2 \boldsymbol{I})$ is the additive white Gaussian noise.

For a fixed satellite communication system over Ka-band, the LoS ray is the dominant path [11]. Hence, the $\boldsymbol{H}_t$ denotes the frequency domain channel matrix with one LoS path, which is given by,

$$\boldsymbol{H}_t = \alpha_t \boldsymbol{A}^{\mathrm{H}}(\theta_t, \phi_t), \tag{2}$$

where α_t is the complex channel gain and $\boldsymbol{A}(\theta_t, \phi_t) \in \mathbb{C}^{(N_x\times N_y)\times 1}$ represents the transmit array response vectors, with θ_t and ϕ_t as azimuth and elevation directions of AoD at time t. We take the UPA into consideration and the steering vector can be written as

$$\boldsymbol{A}(\theta_t, \phi_t) = \frac{1}{\sqrt{N_x \times N_y}} [1, \ldots, e^{\jmath \frac{2\pi}{\lambda} d n_x \sin\theta_t \sin\phi_t + n_y \cos\phi_t} , \ldots, e^{\jmath \frac{2\pi}{\lambda} d N_x \sin\theta_t \sin\phi_t + N_y \cos\phi_t}]^{\mathrm{T}}, \tag{3}$$

where λ is the signal wavelength, with array size as $N_x \times N_y$ and $0 \leqslant n_x < N_x$ and $0 \leqslant n_y < N_y$ are the indices of an antenna element. In this paper, we consider the case of quantized AoD pair, i.e., $\theta_t \in \{0, \theta_{\text{range}}/N_x, \ldots, \theta_{\text{range}}(N_x - 1)/N_x\}$ and $\phi_t \in \{0, \phi_{\text{range}}/N_y, \ldots, \phi_{\text{range}}(N_y - 1)/N_y\}$ [1] which means AoD pair is chosen from the $N_x \times N_y$ uniform grid.[2]

For large number of antenna elements at the LEO satellite, spatial resolution of the angular basis extends. Hence, the channel is treated sparse and the channel model can be written as a virtual model and then processed by compressive sensing techniques, which is presented as

$$\boldsymbol{H}_t = \boldsymbol{g}_t \tilde{\boldsymbol{A}}^{\mathrm{H}}(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{4}$$

where $\boldsymbol{g}_t \in \mathbb{C}^{1\times M}$ is a sparse matrix with one non-zero elements (which means only one dominant path), $\tilde{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi}) \in \mathbb{C}^{(N_x\times N_y)\times M}$ is the array response vector dictionary matrix with spatial grids M.

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_S	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
θ_t, ϕ_t	Azimuth and elevation directions of AoD at time slot t
θ_{range}	The range of azimuth directions of AoD
ϕ_{range}	The range of elevation directions of AoD
α_t	Complex channel gain at time slot t
$\boldsymbol{A}(\theta_t, \phi_t)$	Transmit array response vector at time slot t
$\boldsymbol{H}_t$	Frequency domain channel model at time slot t
$\boldsymbol{S}$	Transmit pilot symbols
$\boldsymbol{y}_t^{\mathrm{RX}}, \boldsymbol{y}_t$	Received signal at UAV at time slot t
$\boldsymbol{y}^{(t)}$	Total received signal at UAV for period t
$\boldsymbol{n}_t^{\mathrm{RX}}, \boldsymbol{n}_t$	Additive white Gaussian noise at time slot t
$g_{t,m}$	Channel gain element for the m^{th} AoD at time slot t
$\boldsymbol{g}_t$	Sparse matrix complex channel gain at time slot t
$\boldsymbol{g}^{(t)}, \boldsymbol{g}^{(T)}$	Total sparse matrix complex channel gain for period t or T
$b_{t,m}$	Hidden support element for the m^{th} AoD at time slot t
$\boldsymbol{b}_t$	Hidden support vector at time slot t
$\boldsymbol{b}^{(t)}, \boldsymbol{b}^{(T)}$	Hidden support vector for total time slot period t or T
$a_{t,m}$	Azimuth support element for the m^{th} AoD at time slot t
$e_{t,m}$	Elevation support element for the m^{th} AoD at time slot t
$\boldsymbol{a}_t, \boldsymbol{e}_t$	Azimuth and elevation support vector at time slot t
$\boldsymbol{a}^{(t)}, \boldsymbol{a}^{(T)}$	Azimuth support vector for total time slot period t or T
$\boldsymbol{e}^{(t)}, \boldsymbol{e}^{(T)}$	Elevation support vector for total time slot period t or T
$\vartheta_{t,m}$	Hidden value element for the m^{th} AoD at time slot t
$\boldsymbol{\vartheta}_t$	Hidden value vector at time slot t
$\boldsymbol{\vartheta}^{(t)}, \boldsymbol{\vartheta}^{(T)}$	Hidden value vector for total time slot period t or T

[1] Due to the long distance between the UAV and the satellite, the moving range of both azimuth and elevation angle would be limited.

[2] In this paper, the quantized AoDs are considered, leave the continuous condition for future work.

III. Channel Model and Problem Formulation

In this section, the 3D-2D-MM channel model is introduced and the problem formulation of the proposed dynamic channel tracking is presented.

A. 3D-2D-MM Channel Model

The dynamic angular domain channel $\boldsymbol{g}^{(T)} = \{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_T\}$ can be modeled as a probabilistic signal model with two hidden random processes, $\boldsymbol{b}^{(T)} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_T\}$ and $\boldsymbol{\vartheta}^{(T)} = \{\boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_T\}$. Here, $\boldsymbol{b}_t = \left[b_{t,1,1}, \ldots, b_{t,N_x,N_y}\right] = [b_{t,1}, \ldots, b_{t,M}] \in \{0,1\}^M$ represents the joint hidden support vector at time t, which indicates the channel sparsity with $b_{t,m} \in \{1,0\}$.[3] The hidden value vector $\boldsymbol{\vartheta}_t = [\vartheta_{t,1}, \ldots, \vartheta_{t,M}] \in \mathbb{C}^M$ represents the temporal correlation of channel coefficients. The dynamic channel element for the m^{th} AoD direction in the time t can be written as

$$g_{t,m} = b_{t,m}\vartheta_{t,m}, \tag{5}$$

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

To formulate the more realistic dynamic channel (i.e., the non-activated AoD path at the current time slot[4] t may be activated at time slot $t+1$), a probabilistic channel model with channel prior distribution can be formed as

$$p\left(\boldsymbol{g}^{(T)}, \boldsymbol{b}^{(T)}, \boldsymbol{\vartheta}^{(T)}\right) = p\left(\boldsymbol{g}^{(T)} \mid \boldsymbol{b}^{(T)}, \boldsymbol{\vartheta}^{(T)}\right) p\left(\boldsymbol{b}^{(T)}\right) p\left(\boldsymbol{\vartheta}^{(T)}\right), \tag{6}$$

where the channel vector conditional prior is

$$p\left(\boldsymbol{g}^{(T)} \mid \boldsymbol{b}^{(T)}, \boldsymbol{\vartheta}^{(T)}\right) = \prod_{t=1}^{T}\prod_{m=1}^{M} p\left(g_{t,m} \mid b_{t,m}, \vartheta_{t,m}\right) = \prod_{t=1}^{T}\prod_{m=1}^{M} \delta\left(g_{t,m} - b_{t,m}\vartheta_{t,m}\right), \tag{7}$$

with $\delta(\cdot)$ is the Dirac delta function. The factor graph of the combined channel can be found in Fig. 2, with $\pi_{t,m}$ as the conditional prior $p\left(g_{t,m} \mid b_{t,m}, \vartheta_{t,m}\right)$.

The joint hidden support vector of azimuth and elevation directions and the hidden value vector are detailed as follows:

1) Joint Hidden Support Vector of Azimuth and Elevation: The relationship of the azimuth support $\boldsymbol{a}^{(T)}$ and elevation support $\boldsymbol{e}^{(T)}$ can be found in Fig. 3a, where only one active path is considered.

In our model, the known LEO satellite orbiting and the uncertainty UAV mobility are considered as the relative angular variation rather than the certain location. To be specific, because of the 3D UAV trajectory [26], together with the fixed orbital voyage of the LEO satellite, the relative displacement of the UAV and LEO satellite is continuous. This means for a time-varying case, the azimuth and elevation angles would only change to adjacent degree in the angular domain

[3]To be noticed, $m = (n_x - 1) \times N_x + n_y$

[4]In this paper, each time slot indicates a single estimation transmit frame.

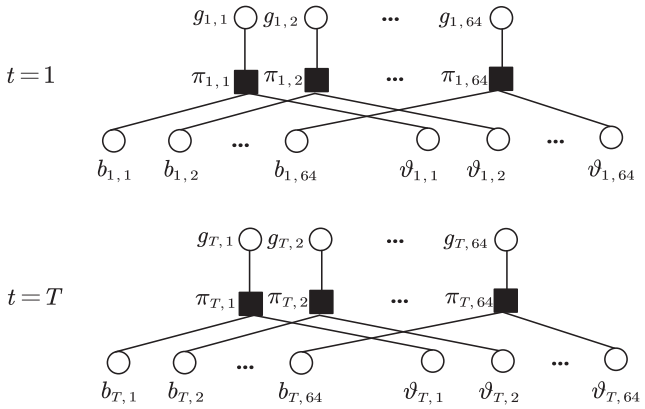

Fig. 2. Factor graph of 3D-2D-MM of the channel when $N_x = 8$ and $N_y = 8$. The detailed model of the joint hidden support vector $\boldsymbol{b}_t$ and hidden value vector $\boldsymbol{\vartheta}_t$ are illustrated in Fig. 3 and Fig. 4, separately.

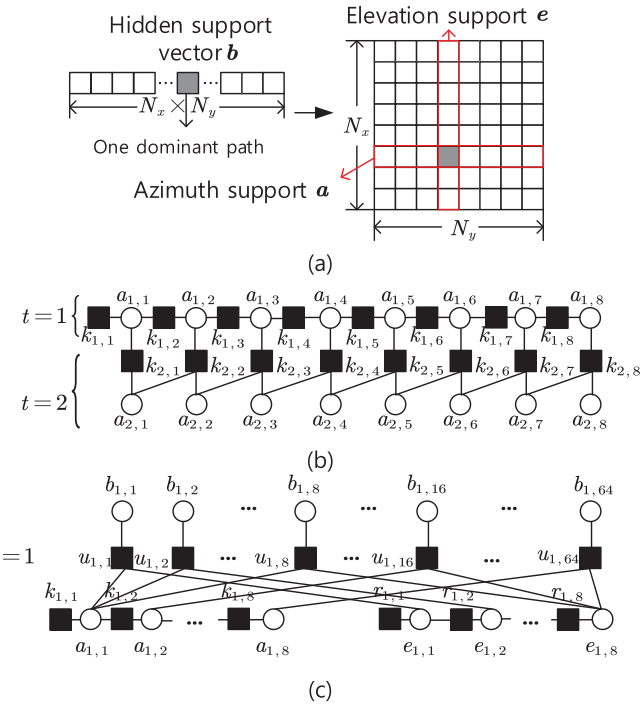

Fig. 3. (a) Hidden support matrix illustration. (b) Factor graph of 2D-MM of the azimuth support vectors when $T = 2$ and $N_x = 8$. (c) Factor graph of the 3D-2D-MM joint hidden support vector $\boldsymbol{b}_t$ when $T = 1$, $N_x = 8$, $N_y = 8$.

with certain probabilities. These probabilities can be treated as an expression of the mobility for both the UAV and satellite. Practically, parameters can be trained in advance and stored for real-time channel tracking based on the expectation maximization algorithm proposed in [29].

In other words, the spatial sparsity pattern will act highly related with the prior adjacent pattern, i.e., a_{t+1,n_x} has higher probability depends on the adjacent a_{t,n_x-1} in previous slot [16]. It is also verified that the support changes slowly over time [18], which indicates a_{t+1,n_x} has higher probability depends on the a_{t,n_x}. The spatial sparsity and the probabilistic relationship passing through time are called spatial correlation and temporal correlation, respectively. Combining both spatial

and temporal correlation is the reason for naming the azimuth support vector model as 2D-MM. The model can be expressed as:

$$p\left(\boldsymbol{a}^{(T)}\right) = \underbrace{p\left(a_{1,1}\right)}_{q_{a_{1,1}}} \prod_{n_x=2}^{N_x} \underbrace{p\left(a_{1,n_x} \mid a_{1,n_x-1}\right)}_{q^{(AS)}_{a_{1,n_x},a_{1,n_x-1}}} \times \prod_{t=2}^{T} \left(\underbrace{p\left(a_{t,1} \mid a_{t-1,1}\right)}_{q^{(AT)}_{a_{t,1},a_{t-1,1}}} \prod_{n_x=1}^{N_x} \underbrace{p(a_{t,n_x} \mid a_{t-1,n_x-1}, a_{t-1,n_x})}_{q^{(A)}_{a_{t,n_x},a_{t-1,n_x-1},a_{t-1,n_x}}} \right), \quad (8)$$

where $q_{a_{1,1}}$, $q^{(AS)}_{a_{1,n_x},a_{1,n_x-1}}$, $q^{(AT)}_{a_{t,1},a_{t-1,1}}$ and $q^{(A)}_{a_{t,n_x},a_{t-1,n_x-1},a_{t-1,n_x}}$ are the transition probabilities that equal to the factor nodes $k_{1,1}$, k_{1,n_x}, $k_{t,1}$ and k_{t,n_x} in Fig. 3b and in Fig. 3c of the azimuth directions. To be noticed, the azimuth supports $a_{t,n_x} \in \{0,1\}$. The illustration of this 2D-MM is in the factor graph in Fig. 3b. The support of the elevation direction of AoD $\boldsymbol{e}^{(T)}$ cis modeled in the similar way.

Hence, the total 3D-2D-MM of the AoD supports, which combines both azimuth and elevation direction supports is in Fig. 3c. To be noticed, each hidden support element $b_{t,m}$ in hidden support vector $\boldsymbol{b}_t$ is actually the combination of the n_x^{th} azimuth and n_y^{th} elevation support vectors. The joint conditional prior of channel support probability is given by

$$p\left(\boldsymbol{b}^{(T)} | \boldsymbol{a}^{(T)}, \boldsymbol{e}^{(T)}\right) = \prod_{t,n_x,n_y}^{T,N_x,N_y} p\left(b_{t,n_x,n_y} | a_{t,n_x}, e_{t,n_y}\right). \quad (9)$$

2) Gaussian Markov Model of Hidden Value Vector: The illustration of the factor graph of the hidden value vector $\boldsymbol{\vartheta}_t$ and the Gaussian-Markov model is in Fig. 4a and Fig. 4b, respectively. Due to the path gains change smoothly over time [18], the hidden value vector can be formulated as Gauss-Markov processes [29]

$$\vartheta_{t,m} = (1-\beta)\left(\vartheta_{t-1,m} - \mu\right) + \beta\omega_{t,m} + \mu, \quad (10)$$

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

$$p\left(\vartheta_{t,m} \mid \vartheta_{t-1,m}\right) \sim \mathcal{CN}\left(\vartheta_{t,m}; (1-\beta)\,\vartheta_{t-1,m} + \beta\mu, \beta^2\zeta\right). \quad (11)$$

The distribution of the steady state of the process is $\vartheta_{t,m} \sim \mathcal{CN}(\mu, \sigma^2 = \frac{\beta\zeta}{2-\beta})$. And the joint distribution of $\boldsymbol{\vartheta}^{(T)}$ can be formulated as

$$p\left(\boldsymbol{\vartheta}^{(T)}\right) = \prod_{m=1}^{M} p\left(\vartheta_{1,m}\right) \prod_{t=2}^{T} p\left(\vartheta_{t,m} \mid \vartheta_{t-1,m}\right), \quad (12)$$

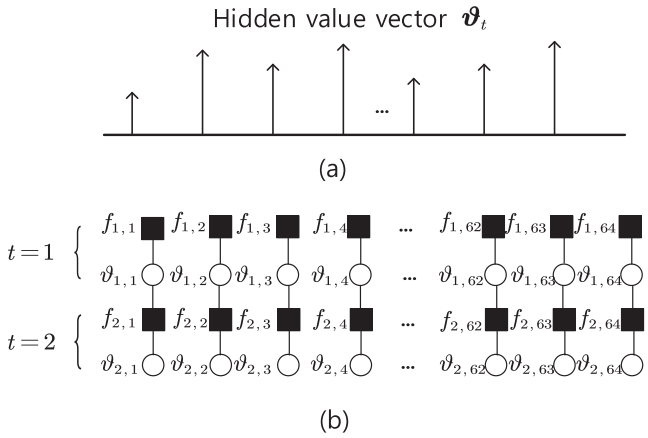

Fig. 4. (a) Hidden value vector illustration at a single time. (b) Factor graph of Gauss-Markov model of the hidden value vectors with $T = 2$, $M = 64$.

where $p\left(\vartheta_{1,m}\right)$ and $p\left(\vartheta_{t,m} \mid \vartheta_{t-1,m}\right)$ are the transition probabilities that equal to the factor nodes $f_{1,m}$ and $f_{t,m}$ in Fig. 4b.

B. Dynamic Sparse Channel Tracking Problem Formulation

By taking the angular domain channel (4) into consideration, the received signal can be written as

$$\boldsymbol{y}_t^{\mathrm{RX}} = \boldsymbol{g}_t \tilde{\boldsymbol{A}}^H(\boldsymbol{\theta}, \boldsymbol{\phi})\,\boldsymbol{S} + \boldsymbol{n}_t^{\mathrm{RX}}. \quad (13)$$

with $\boldsymbol{y}_t^{\mathrm{RX}} \in \mathbb{C}^{1\times N_S}$. Then, the received signal can be rewritten into the standard CS model

$$\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{g}_t + \boldsymbol{n}_t, \quad (14)$$

with $\boldsymbol{y}_t \in \mathbb{C}^{N_S\times 1}$, $\boldsymbol{g}_t = \boldsymbol{g}_t^T$, and the sensing matrix $\boldsymbol{\Phi} = \boldsymbol{S}^T(\tilde{\boldsymbol{A}}^H)^T \in \mathbb{C}^{N_S\times M}$. To be noticed, the transmitted pilot is defined as $\boldsymbol{S} \in \{0,1\}^{N_S\times M}$ with randomly selected and reordered N_S rows of the $M \times M$ identity matrix (In this way, $\boldsymbol{\Phi}$ can be treated as an approximate partial orthogonal sensing matrix and provides good performance).

Given the received signals of the t time period $\boldsymbol{y}^{(t)}$, the purpose is to track the dynamic channel vector $\boldsymbol{g}_t$ at time t. The channel vector can be estimated as $\hat{g}_{t,m} = E\left(g_{t,m} | \boldsymbol{y}^{(t)}\right)$, where the expectation is over the marginal posterior

$$p\left(g_{t,m} | \boldsymbol{y}^{(t)}\right) \propto \sum_{\boldsymbol{b}^{(t)}} \int_{\boldsymbol{g}^{(t)}_{-(t,m)}, \boldsymbol{\vartheta}^{(t)}} p(\boldsymbol{g}^{(t)}, \boldsymbol{b}^{(t)}, \boldsymbol{\vartheta}^{(t)}, \boldsymbol{y}^{(t)}) \quad (15)$$

with $\boldsymbol{g}^{(t)}_{-(t,m)}$ denotes the vector $\boldsymbol{g}^{(t)}$ without the element $g_{t,m}$. Our goal is to obtain the MMSE of the estimated $\boldsymbol{g}$.

IV. Proposed 3D Dynamic Turbo Approximate Message Passing Algorithm

The factor graph of the total 3D-2D-MM distribution $p(\boldsymbol{g}^{pri}_{\mathcal{B},t}, \boldsymbol{b}^{(t)}, \boldsymbol{\vartheta}^{(t)}, \boldsymbol{y}^{(t)})$ is shown in Fig. 5. This factor graph contains two sub-graphs: the joint hidden support $\boldsymbol{b}^{(T)}$ (on the left-hand side of Fig. 5) and the hidden value vector $\boldsymbol{\vartheta}^{(T)}$ (on the right-hand side of Fig. 5) through time T. The structure of the total hidden support $\boldsymbol{b}^{(T)}$ can be further extended as a 3D-2D-MM factor graph which contains both azimuth support

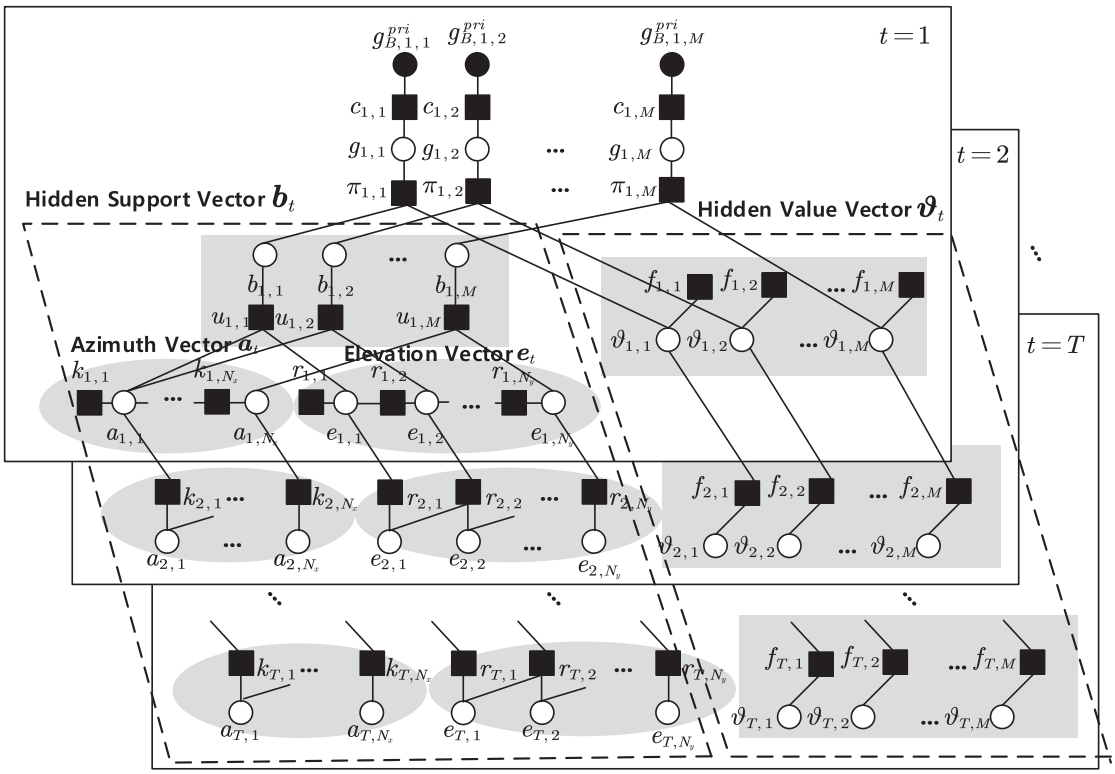

Fig. 5. Factor graph representation of the proposed 3D-2D-MM.

TABLE II

TABLE OF FACTORS, DISTRIBUTIONS AND FUNCTIONAL FORMS ASSOCIATED WITH OUR 3D-2D-MM

Factor	Distribution	Function Form
$c_{t,m}\left(g_{t,m}, g^{pri}_{\mathcal{B},t,m}\right)$	$p(g^{pri}_{\mathcal{B},t,m}\|g_{t,m})$	$\mathcal{CN}\left(\boldsymbol{g}_t; \boldsymbol{g}^{pri}_{\mathcal{B},t}, v^{pri}_{\mathcal{B},t}\boldsymbol{I}\right)$
$\pi_{t,m}\left(g_{t,m}, b_{t,m}, \vartheta_{t,m}\right)$	$p\left(g_{t,m}\|b_{t,m}, \vartheta_{t,m}\right)$	$\delta\left(g_{t,m} - b_{t,m}\vartheta_{t,m}\right)$
$f_{1,m}\left(\vartheta_{1,m}\right)$	$p\left(\vartheta_{1,m}\right)$	$\mathcal{CN}\left(\vartheta_{1,m}; \mu, \sigma^2\right)$
$f_{t,m}\left(\vartheta_{t,m}, \vartheta_{t-1,m}\right)$	$p\left(\vartheta_{t,m}\|\vartheta_{t-1,m}\right)$	$\mathcal{CN}\left(\vartheta_{t,m}; (1-\beta)\,\vartheta_{t-1,m} + \beta\mu, \beta^2\zeta\right)$
$k_{1,1}(a_{1,1})$	$p(a_{1,1})$	$(1-\lambda)^{1-a_{1,1}}\lambda^{a_{1,1}}$
$k_{1,m}(a_{1,m}, a_{1,m-1})$	$p(a_{1,m}\|a_{1,m-1})$	$\begin{cases}\left(q^{(AS)}_{0,1}\right)^{a_{1,m}}\left(1-q^{(AS)}_{0,1}\right)^{(1-a_{1,m})}, a_{1,m-1}=0\\ \left(1-q^{(AS)}_{1,0}\right)^{a_{1,m}}\left(q^{(AS)}_{1,0}\right)^{(1-a_{1,m})}, a_{1,m-1}=1\end{cases}$
$k_{t,1}(a_{t,1}, a_{t-1,1})$	$p(a_{t,1}\|a_{t-1,1})$	$\begin{cases}\left(q^{(AT)}_{0,1}\right)^{a_{t,1}}\left(1-q^{(AT)}_{0,1}\right)^{(1-a_{t,1})}, a_{t-1,1}=0\\ \left(1-q^{(AT)}_{1,0}\right)^{a_{t,1}}\left(q^{(AT)}_{1,0}\right)^{(1-a_{t,1})}, a_{t-1,1}=1\end{cases}$
$k_{t,m}(a_{t,m}, a_{t,m-1}, a_{t-1,m})$	$p(a_{t,m}\|a_{t,m-1}, a_{t-1,m})$	$\left(q^{(A)}_{a_{t,m-1},a_{t-1,m},1}\right)^{a_{t,m}}\left(1-q^{(A)}_{a_{t,m-1},a_{t-1,m},1}\right)^{(1-a_{t,m})}, a_{t,m-1}, a_{t-1,m}\in\{0,1\}$
$r_{1,1}(a_{1,1})$	$p(e_{1,1})$	$(1-\lambda)^{1-e_{1,1}}\lambda^{e_{1,1}}$
$r_{1,m}(e_{1,m}, e_{1,m-1})$	$p(e_{1,m}\|e_{1,m-1})$	$\begin{cases}\left(q^{(ES)}_{0,1}\right)^{e_{1,m}}\left(1-q^{(ES)}_{0,1}\right)^{(1-e_{1,m})}, e_{1,m-1}=0\\ \left(1-q^{(ES)}_{1,0}\right)^{e_{1,m}}\left(q^{(ES)}_{1,0}\right)^{(1-e_{1,m})}, e_{1,m-1}=1\end{cases}$
$r_{t,1}(e_{t,1}, e_{t-1,1})$	$p(e_{t,1}\|e_{t-1,1})$	$\begin{cases}\left(q^{(ET)}_{0,1}\right)^{e_{t,1}}\left(1-q^{(ET)}_{0,1}\right)^{(1-e_{t,1})}, e_{t-1,1}=0\\ \left(1-q^{(ET)}_{1,0}\right)^{e_{t,1}}\left(q^{(ET)}_{1,0}\right)^{(1-e_{t,1})}, e_{t-1,1}=1\end{cases}$
$r_{t,m}(e_{t,m}, e_{t,m-1}, e_{t-1,m})$	$p(e_{t,m}\|e_{t,m-1}, e_{t-1,m})$	$\left(q^{(E)}_{e_{t,m-1},e_{t-1,m},1}\right)^{e_{t,m}}\left(1-q^{(E)}_{e_{t,m-1},e_{t-1,m},1}\right)^{(1-e_{t,m})}, e_{t,m-1}, e_{t-1,m}\in\{0,1\}$
$u_{t,m}(b_{t,m}, a_{t,n_x}, e_{t,n_y})$	$p(b_{1,m}\|a_{t,n_x}, e_{t,n_y})$	$a_{t,n_x}e_{t,n_y}(\gamma_{\boldsymbol{b}_t})^{b_{t,m}}(1-\gamma_{\boldsymbol{b}_t})^{(1-b_{t,m})} + (1-a_{t,n_x}e_{t,n_y})\delta(b_{t,m})$

vector $\boldsymbol{a}^{(T)}$ and elevation support vector $\boldsymbol{e}^{(T)}$ through time T. The expression of function nodes is listed in Table II. The details of all the nodes are elaborated in Section III-A.

There are two reasons for utilizing the AMP framework in our model: rigorous analysis and extremely fast runtimes [30]. Conventional AMP runs with the static priors. However, for our special designed factor graph, the prior distribution knowledge for the AMP algorithm is changed over time. Due to the complex structure in the total factor graph of the 3D-2D-MM, it is challenging to achieve the exact marginal

posterior in (15). The extended version called turbo-AMP was proposed in [31] which offered a way to further improve the performance. Conventional turbo compressed sensing is an iterative algorithm which follows the two-step iteration framework [32]: linear minimum mean-square error (LMMSE) estimator and i.i.d. prior-based MMSE estimator. In this section, to perfectly exploit the dynamic channel structure, we elaborate the details of the proposed algorithm based on turbo iterative framework with the proposed 3D-2D-MM MMSE denoiser. The proposed algorithm will be introduced in three parts: Module $\mathcal{A}$ LMMSE Estimator, Module $\mathcal{B}$ message passing MMSE denoiser, and message passing through time.

A. Module $\mathcal{A}$: LMMSE Estimator

In Module $\mathcal{A}$, the angular channel vector $\boldsymbol{g}_t$ is estimated based on the received signal $\boldsymbol{y}_t$ with the prior distribution $\mathcal{CN}\left(\boldsymbol{g}_t; \boldsymbol{g}_{\mathcal{A},t}^{pri}, v_{\mathcal{A},t}^{pri}\boldsymbol{I}\right)$, where the extrinsic mean and variance are defined as [32]

$$\boldsymbol{g}_{\mathcal{A},t}^{post} = \boldsymbol{g}_{\mathcal{A},t}^{pri} + \frac{v_{\mathcal{A},t}^{pri}}{v_{\mathcal{A},t}^{pri} + \sigma_e^2}\boldsymbol{\Phi}_t^H\left(\boldsymbol{y}_t - \boldsymbol{\Phi}\boldsymbol{g}_{\mathcal{A},t}^{pri}\right) \tag{16}$$

and

$$v_{\mathcal{A},t}^{post} = v_{\mathcal{A},t}^{pri} - \frac{N_S}{M}\frac{\left(v_{\mathcal{A},t}^{pri}\right)^2}{v_{\mathcal{A},t}^{pri} + \sigma_e^2}, \tag{17}$$

respectively. The extrinsic distribution of $\boldsymbol{g}_t$ is given by

$$\begin{aligned}&\mathcal{CN}\left(\boldsymbol{g}_t; \boldsymbol{g}_{\mathcal{A},t}^{post}, v_{\mathcal{A},t}^{post}\boldsymbol{I}\right)\\ &\quad\propto \mathcal{CN}\left(\boldsymbol{g}_t; \boldsymbol{g}_{\mathcal{A},t}^{pri}, v_{\mathcal{A},t}^{pri}\boldsymbol{I}\right)\mathcal{CN}\left(\boldsymbol{g}_t; \boldsymbol{g}_{\mathcal{A},t}^{ext}, v_{\mathcal{A},t}^{ext}\boldsymbol{I}\right),\end{aligned} \tag{18}$$

where the extrinsic mean and variance are respectively given by

$$\boldsymbol{g}_{\mathcal{A},t}^{ext} = v_{\mathcal{A},t}^{ext}\left(\frac{\boldsymbol{g}_{\mathcal{A},t}^{post}}{v_{\mathcal{A},t}^{post}} - \frac{\boldsymbol{g}_{\mathcal{A},t}^{pri}}{v_{\mathcal{A},t}^{pri}}\right) \tag{19}$$

$$v_{\mathcal{A},t}^{ext} = \left(\frac{1}{v_{\mathcal{A},t}^{post}} - \frac{1}{v_{\mathcal{A},t}^{pri}}\right)^{-1}. \tag{20}$$

B. Module $\mathcal{B}$: Message Passing MMSE Denoiser

In module $\mathcal{B}$, the message passing MMSE denoiser is executed by exploiting the spatial sparsity details of the proposed 3D-2D-MM.

As the core structure of the turbo compressed sensing, the extrinsic mean and variance are transmitted from the LMMSE estimator in module $\mathcal{A}$, which means $\boldsymbol{g}_{\mathcal{B},t}^{pri} = \boldsymbol{g}_{\mathcal{A},t}^{ext}$ and $v_{\mathcal{B},t}^{pri} = v_{\mathcal{A},t}^{ext}$ [32]. For message passing algorithm, the basic assumption is that $\boldsymbol{g}_{\mathcal{B},t}^{pri} = \boldsymbol{g}_t + \boldsymbol{z}_t$ where $\boldsymbol{z}_t \sim \mathcal{CN}(0, v_{\mathcal{B},t}^{pri}\boldsymbol{I})$ is independent of $\boldsymbol{g}_t$ [32]. Aiming to calculate the approximate posterior distributions of $p(g_{t,m}|\boldsymbol{g}_{\mathcal{B}}^{pri})$ instead of (15), the sum-product message passing rule is used, which carefully follows the message passing structure shown in Fig. 5. The formulation details can be found in Appendix A and the procedure is summarized as follows:

- Firstly, the message passing over the Gaussian hidden value vector $\boldsymbol{\vartheta}_t$ is given by (38) with μ and σ^2 as the input to update the message $\upsilon_{\vartheta_{t,m}\to\pi_{t,m}}$.
- Secondly, the message is passed over the path $g_{t,m} \to \pi_{t,m} \to b_{t,m} \to u_{t,m} \to a_{t,n_x}$ using (39) to (41), with the input $\boldsymbol{g}_{\mathcal{B},t}^{pri}$ and $v_{\mathcal{B},t}^{pri}$ from module $\mathcal{A}$.
- Then the forward-backward message passing is performed over the 3D-2D-MM with sequence three paths: the 2D-MM of the azimuth support $\boldsymbol{a}_t$, path $a_{t,n_x} \to u_{t,m} \to e_{t,n_y}$ and the 2D-MM of the elevation support $\boldsymbol{e}_t$. The details of the 3D-2D-MM azimuth and elevation support nonzero probability estimation are summarized in **Algorithm 1**.
- The message is finally passed back over the path $e_{t,n_y} \to u_{t,m} \to b_{t,m} \to \pi_{t,m} \to g_{t,m}$ using (60) to (63).

Based on the calculated updated messages $\upsilon_{\pi_{t,m}\to g_{t,m}}$, the posterior distributions can be written as

$$p\left(g_{t,m}|\boldsymbol{g}_{\mathcal{B},t}^{pri}\right) \propto \upsilon_{\pi_{t,m}\to g_{t,m}}(g_{t,m})\,\upsilon_{c_{t,m}\to g_{t,m}}(g_{t,m}), \tag{21}$$

where $\upsilon_{c_{t,m}\to g_{t,m}}(g_{t,m}) = \mathcal{CN}\left(g_{t,m}; g_{\mathcal{B},t,m}^{pri}, v_{\mathcal{B},t}^{pri}\right)$. Then the posterior mean and variance of each element of $\boldsymbol{g}_t$ can be updated as

$$g_{\mathcal{B},t,m}^{post} = \int_{g_{t,m}} g_{t,m} p\left(g_{t,m}|\boldsymbol{g}_{\mathcal{B},t}^{pri}\right) \tag{22}$$

and

$$v_{\mathcal{B},t}^{post} = \frac{1}{M}\sum_{m=1}^{M}\mathrm{Var}\left(g_{t,m}|\boldsymbol{g}_{\mathcal{B},t}^{pri}\right), \tag{23}$$

where $\mathrm{Var}\left(g_{t,m}|\boldsymbol{g}_{\mathcal{B},t}^{pri}\right)$ denotes the conditional variance of $g_{t,m}$ given $\boldsymbol{g}_{\mathcal{B},t}^{pri}$. Then similar to (19) and (20), the extrinsic update mean and covariance can be written as

$$\boldsymbol{g}_{\mathcal{A},t}^{pri} = \boldsymbol{g}_{\mathcal{B},t}^{ext} = v_{\mathcal{A},t}^{pri}\left(\frac{\boldsymbol{g}_{\mathcal{B},t}^{post}}{v_{\mathcal{B},t}^{post}} - \frac{\boldsymbol{g}_{\mathcal{B},t}^{pri}}{v_{\mathcal{B},t}^{pri}}\right) \tag{24}$$

and

$$v_{\mathcal{A},t}^{pri} = v_{\mathcal{B},t}^{ext} = \left(\frac{1}{v_{\mathcal{B},t}^{post}} - \frac{1}{v_{\mathcal{B},t}^{pri}}\right)^{-1}. \tag{25}$$

C. Message Passing Through Time

For channel tracking, temporal correlation can be utilized for better recovery with prior information provided. Therefore, we also consider the message passing through time. There are mainly three paths passing through time: the path of the Gaussian hidden value vector $\boldsymbol{\vartheta}_t$, the path of the azimuth support $\mathbf{a}_t$ and the path of the elevation support $\mathbf{e}_t$. The sequence paths can be observed in Fig. 6.

1) The Gaussian Hidden Value Vector $\boldsymbol{\vartheta}_t$ Passing Through Time: The message passing from factor node $\pi_{t,m}$ to variable node $\vartheta_{t,m}$ should be

$$\begin{aligned}&\upsilon_{\pi_{t,m}\to\vartheta_{t,m}}^{true}(\vartheta_{t,m})\\ &\quad= \rho_{t,m}^{out}\mathcal{CN}\left(\vartheta_{t,m}; \boldsymbol{g}_{\mathcal{B},t}^{pri}, v_{\mathcal{B},t}^{pri}\right)\\ &\qquad+ \left(1-\rho_{t,m}^{out}\right)\mathcal{CN}\left(0; \boldsymbol{g}_{\mathcal{B},t}^{pri}, v_{\mathcal{B},t}^{pri}\right).\end{aligned} \tag{26}$$

Algorithm 1 3D-2D-MM Azimuth and Elevation Support Estimation

Input: $\rho_{t,n_x}^{(A)in}$, $\rho_{t,n_x}^{(A)dyna}$, $\rho_{t,n_y}^{(E)in}$, $\rho_{t,n_y}^{(E)dyna}$.
Output: $\rho_{t,n_x}^{(A)out}$, $\rho_{t,n_y}^{(E)out}$.

% Azimuth Estimation:
if $t=1$ **then**
  Calculate $\upsilon_{u_{t,m}\to a_{t,n_x}}$ as in (44)
  Initialization: $\gamma_{t,1}^{(A)f}=\lambda$, $\gamma_{t,N_x}^{(A)b}=\frac{1}{2}$
  for $n_x=2,\ldots,N_x$ **do**
    Use (45) to calculate $\gamma_{t,n_x}^{(A)f}$.
  end for
  for $n_x=1,\ldots,N_x-1$ **do**
    Use (46) to calculate $\gamma_{t,n_x}^{(A)b}$.
  end for
else if $t>1$ **then**
  Initialization:$\gamma_{t,1}^{f}=q_{0,1}^{(T)}(1-\rho_{t,1}^{ac})+\left(1-q_{1,0}^{(T)}\right)\rho_{t,1}^{ac}$, $\gamma_{t,N_x}^{b}=\frac{1}{2}$.
  for $n_x=2,\ldots,N_x$ **do**
    Calculate the $\gamma_{t,n_x}^{(A)f}$ as in (47)
  end for
  for $n_x=1,\ldots,N_x-1$ **do**
    Calculate $\gamma_{t,n_x}^{(A)b}=\frac{1}{1+\gamma^{(A)}}$ use (48).
  end for
end if
Calculate AoD azimuth support vector $\rho_{t,n_x}^{(A)out}$ as in (49).
%Message Passing Over Azimuth to Elevation support:
Calculate individual message $\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out}$ using (51).
Calculate message $\upsilon_{a_{t,n_x}\to u_{t,m}}(a_{t,n_x})$ and $\upsilon_{u_{t,m}\to e_{t,n_y}}(e_{t,n_y})$ as in (50) and (52), respectively.
% Elevation Estimation:
if $t=1$ **then**
  Calculate $\upsilon_{u_{t,m}\to e_{t,n_y}}$ as in (54)
  Initialization: $\gamma_{t,1}^{(E)f}=\lambda$, $\gamma_{t,N_y}^{(E)b}=\frac{1}{2}$
  for $n_x=2,\ldots,N_y$ **do**
    Use (55) to calculate $\gamma_{t,n_y}^{(E)f}$.
  end for
  for $n_y=1,\ldots,N_y-1$ **do**
    Use (56) to calculate $\gamma_{t,n_y}^{(E)b}$.
  end for
else if $t>1$ **then**
  Initialization:$\gamma_{t,N_y}^{(E)b}=\frac{1}{2}$, $\gamma_{t,1}^{(E)f}=q_{0,1}^{(ET)}(1-\rho_{t,1}^{(E)dyna})+\left(1-q_{1,0}^{(ET)}\right)\rho_{t,1}^{(E)dyna}$.
  for $n_y=2,\ldots,N_y$ **do**
    Calculate the $\gamma_{t,n_y}^{(E)f}$ as in (57)
  end for
  for $n_y=1,\ldots,N_y-1$ **do**
    Calculate $\gamma_{t,n_y}^{(E)b}=\frac{1}{1+\gamma^{(E)}}$ use (58).
  end for
end if
Calculate AoD elevation support vector $\rho_{t,n_y}^{(E)out}$ as in (59).

To be notice, if $b_{t,m}=0$, $g_{t,m}$ would be 0 in (5), which makes $\vartheta_{t,m}$ unobservable. To solve this kind of problem, similar in [29], a threshold which is slightly less than 1 is introduced.

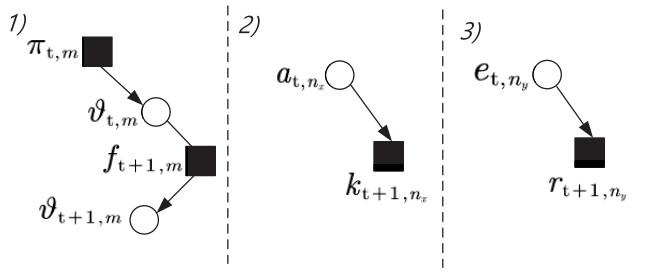

Fig. 6. The illustration of the message passing sequence paths through time, which shows details of the equation (27) to (36).

Hence, the modified message passing is given by

$$\upsilon_{\pi_{t,m}\to\vartheta_{t,m}}(\vartheta_{t,m})=\mathcal{CN}\left(\vartheta_{t,m};\mu_{t,m}^{in},(\sigma_{t,m}^{in})^2\right), \quad (27)$$

with

$$\left(\mu_{t,m}^{in},(\sigma_{t,m}^{in})^2\right)=\begin{cases}\left(\frac{1}{\epsilon}\boldsymbol{g}_{\mathcal{B},t}^{pri},\frac{1}{\epsilon^2}\upsilon_{\mathcal{B},t}^{pri}\right) & \rho_{t,m}^{out}\leqslant \text{Th}, \quad (28)\\ \left(\boldsymbol{g}_{\mathcal{B},t}^{pri},\upsilon_{\mathcal{B},t}^{pri}\right) & \rho_{t,m}^{out}>\text{Th}, \quad (29)\end{cases}$$

where ϵ is a small positive value close to 0. Then the message passing from factor node $f_{t+1,m}$ to variable node $\vartheta_{t+1,m}$ is formed as

$$\upsilon_{f_{t+1,m}\to\vartheta_{t+1,m}}(\vartheta_{t+1,m})=\mathcal{CN}\left(\vartheta_{t+1,m};\mu_{t+1,m}^{dyna},\left(\sigma_{t+1,m}^{dyna}\right)^2\right), \quad (30)$$

where

$$\mu_{t+1,m}^{dyna}=(1-\beta)\left(\frac{\left(\sigma_{t,m}^{dyna}\right)^2\left(\sigma_{t,m}^{in}\right)^2}{\left(\sigma_{t,m}^{dyna}\right)^2+\left(\sigma_{t,m}^{in}\right)^2}\right)\times\left(\frac{\mu_{t,m}^{dyna}}{\left(\sigma_{t,m}^{dyna}\right)^2}+\frac{\mu_{t,m}^{in}}{\left(\sigma_{t,m}^{in}\right)^2}\right)+\beta\mu \quad (31)$$

and

$$\left(\sigma_{t+1,m}^{dyna}\right)^2=(1-\beta)^2\left(\frac{\left(\sigma_{t,m}^{dyna}\right)^2\left(\sigma_{t,m}^{in}\right)^2}{\left(\sigma_{t,m}^{dyna}\right)^2+\left(\sigma_{t,m}^{in}\right)^2}\right)+\beta^2\zeta. \quad (32)$$

2) The Azimuth Support $\mathbf{a}_t$ Passing Through Time: The azimuth support prior information passing to the next time $\upsilon_{a_{t,n_x}\to k_{t+1,n_x}}(a_{t,n_x})$ can be given as

$$\upsilon_{a_{t,n_x}\to k_{t+1,n_x}}(a_{t,n_x})=\rho_{t+1,n_x}^{(A)dyna}\delta(a_{t,n_x}-1)+\left(1-\rho_{t+1,n_x}^{(A)dyna}\right)\delta(a_{t,n_x}), \quad (33)$$

with

$$\rho_{t+1,n_x}^{(A)dyna}=\frac{\gamma_{t,n_x}^{(A)f}\gamma_{t,n_x}^{(A)b}\rho_{t,n_x}^{(A)in}}{\gamma_{t,n_x}^{(A)f}\gamma_{t,n_x}^{(A)b}\rho_{t,n_x}^{(A)in}+(1-\gamma_{t,n_x}^{(A)f})(1-\gamma_{t,n_x}^{(A)b})(1-\rho_{t,n_x}^{(A)in})}. \quad (34)$$

Algorithm 2 3D-DTAMP Dynamic Channel Tracking

Input: $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_T$, sensing matrix $\boldsymbol{\Phi}$.
Output: $\hat{\boldsymbol{g}}_1,\ldots,\hat{\boldsymbol{g}}_T$
1: **for** t=1,...,T **do**
2: **Initialization**:$\boldsymbol{g}_{\mathcal{A},t}^{pri}=\boldsymbol{0}$, $v_{\mathcal{A},t}^{pri}=\sigma^2$
3: **while** not converge **do**
%Module $\mathcal{A}$ LMMSE estimator:
4: Update $\boldsymbol{g}_{\mathcal{A},t}^{post}$ and $v_{\mathcal{A},t}^{post}$ in (16) and (17), respectively.
5: Calculate $\boldsymbol{g}_{\mathcal{B},t}^{pri}=\boldsymbol{g}_{\mathcal{A},t}^{ext}$, $v_{\mathcal{B},t}^{pri}=v_{\mathcal{A},t}^{ext}$ using (19) and (20).
%Module $\mathcal{B}$ 3D-2D-MM MMSE denoiser:
6: Message passing over Gaussian hidden value vector $\boldsymbol{\vartheta}_t$ using (38).
7: Message passing over $g_{t,m}\rightarrow\pi_{t,m}\rightarrow b_{t,m}\rightarrow u_{t,m}\rightarrow a_{t,n_x}$ using (39)-(43)
8: AoD azimuth support $\boldsymbol{a}_t$ and elevation support $\boldsymbol{e}_t$ estimation in **Algorithm 1**.
9: Message passing over elevation support $e_{t,n_y}\rightarrow u_{t,m}\rightarrow b_{t,m}\rightarrow\pi_{t,m}\rightarrow g_{t,m}$ using (60)-(63)
10: Calculate the approximate posterior distributions $p\left(g_{t,m}|\boldsymbol{g}_{\mathcal{B},t}^{pri}\right)$ using (21).
11: Update $\boldsymbol{g}_{\mathcal{B},t,m}^{post}$ and $v_{\mathcal{B},t}^{post}$ using (22) and (23).
12: Update $\boldsymbol{g}_{\mathcal{A},t}^{pri}=\boldsymbol{g}_{\mathcal{B},t}^{ext}$ and $v_{\mathcal{A},t}^{pri}=v_{\mathcal{B},t}^{ext}$ using (24) and (25).
13: Repeat previous steps until convergence.
14: **end while**
15: Update dynamic messages to the next time slot $(t+1)$: $\upsilon_{a_{t,n_x}\rightarrow k_{t+1,n_x}}$, $\upsilon_{e_{t,n_y}\rightarrow r_{t+1,n_y}}$ and $\upsilon_{f_{t+1,m}\rightarrow\vartheta_{t+1,m}}$ using (33), (35) and (30).
16: **end for**

3) The Elevation Support $\boldsymbol{e}_t$ Passing Through Time: Similarly, the elevation support prior information passing to the next time can be written as

$$\upsilon_{e_{t,n_y}\rightarrow r_{t+1,n_y}}(e_{t,n_y}) = \rho_{t+1,n_y}^{(E)dyna}\delta\left(e_{t,n_y}-1\right)+\left(1-\rho_{t+1,n_y}^{(E)dyna}\right)\delta\left(e_{t,n_y}\right), \quad (35)$$

with

$$\rho_{t+1,n_y}^{(E)dyna} = \frac{\gamma_{t,n_y}^{(E)f}\gamma_{t,n_y}^{(E)b}\rho_{t,n_y}^{(E)in}}{\gamma_{t,n_y}^{(E)f}\gamma_{t,n_y}^{(E)b}\rho_{t,n_y}^{(E)in}+(1-\gamma_{t,n_y}^{(E)f})(1-\gamma_{t,n_y}^{(E)b})(1-\rho_{t,n_y}^{(E)in})}. \quad (36)$$

Finally, the overall proposed 3D-DTAMP is summarized in **Algorithm 2**.

V. Numerical Results

In this section, the complexity and the recovery performance of the proposed algorithm are evaluated. For UAV-satellite communication system, the sparsity ratio for both azimuth and elevation directions is $\lambda=\frac{1}{N_x}=\frac{1}{N_y}$ which indicates only one dominant LoS ray. The parameters are set as follows[5]: the azimuth spatial correlation parameters $\{q_{0,1}^{(AS)},q_{1,0}^{(AS)}\}=\{0.05,0.05\}$, the azimuth temporal correlation parameters $\{q_{0,1}^{(AT)},q_{1,0}^{(AT)}\}=\{0.05,0.05\}$, the azimuth joint parameters $\{q_{1,1,1}^{(A)},q_{0,0,1}^{(A)},q_{0,1,1}^{(A)},q_{1,0,1}^{(A)}\}=\{0.9958,0.0013,0.3276,0.3936\}$, the elevation spatial correlation parameters $\{q_{0,1}^{(ES)},q_{1,0}^{(ES)}\}=\{0.025,0.025\}$, the elevation temporal correlation parameters $\{q_{0,1}^{(ET)},q_{1,0}^{(ET)}\}=\{0.025,0.175\}$, the elevation joint parameters $\{q_{1,1,1}^{(E)},q_{0,0,1}^{(E)},q_{0,1,1}^{(E)},q_{1,0,1}^{(E)}\}=\{0.99546,0.0007,0.5,0.1078\}$, the hidden value vector parameters $\{\mu,\sigma^2,\beta,\zeta\}=\{0,\frac{1}{3},0.5,1\}$, Th $=1-10^{-2}$, $\gamma_{\boldsymbol{b}_t}=1$ and $\epsilon=10^{-7}$. The benchmark algorithms which can be employed in our 3D channel model are sparse Bayesian learning (SBL) and D-OMP proposed in [33] and [18], respectively.

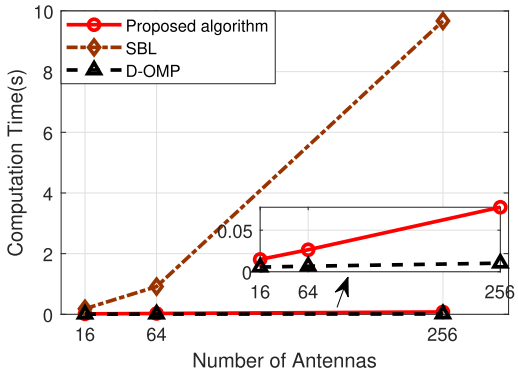

Fig. 7. Computation time of various algorithms versus the number of antennas when SNR $=15$dB, $N_s=10, T=50, M=16,64,256$ and $N_x=N_y=4,8,16$.

A. Complexity Analysis

The complexity of the LMMSE estimator in module $\mathcal{A}$ is $\mathcal{O}(MN_s)$ which is mainly determined by the multiplication of the matrix. The complexity of the proposed MMSE denoiser in module $\mathcal{B}$ is the simple sum-product message passing rule with the $\mathcal{O}(N_x+N_y)$ complexity. To sum up, the total complexity of the proposed algorithm over time T is $\mathcal{O}(T(MN_s+N_x+N_y))$ for each iteration.

As shown in Fig. 7, the SBL takes longer computation time, especially for larger antenna elements. This is caused by the required higher computational complexity of the multiplication and inversion of the matrix. Compared with SBL, the simulation time of both the proposed algorithm and D-OMP drops dramatically. To be noticed, the complexity of D-OMP is $\mathcal{O}(T(N_s(M+1+1)))$ over total time period T. With the number of antennas increase, the proposed algorithm shows comparable computation time.

B. Simulation Results

To evaluate the recovery performance of the algorithms, we refer the performance metric which is called the time-averaged normalized mean square error (TNMSE) as follows [29]

$$\text{TNMSE}\overset{\text{def}}{=}\frac{1}{T}\sum_{t}^{T}\frac{\|\boldsymbol{g}_t-\hat{\boldsymbol{g}}_t\|^2}{\|\boldsymbol{g}_t\|^2}, \quad (37)$$

where $\hat{\boldsymbol{g}}_t$ is the estimated result of $\boldsymbol{g}_t$ in time t.

[5]The expectation maximization algorithm is proposed for training parameters in [29]. Since our paper is focused on the channel tracking based on known probabilities, we recommend readers to refer to [29] for details of the parameters trained algorithm. The parameters in this paper are referred 2D settings in [23].

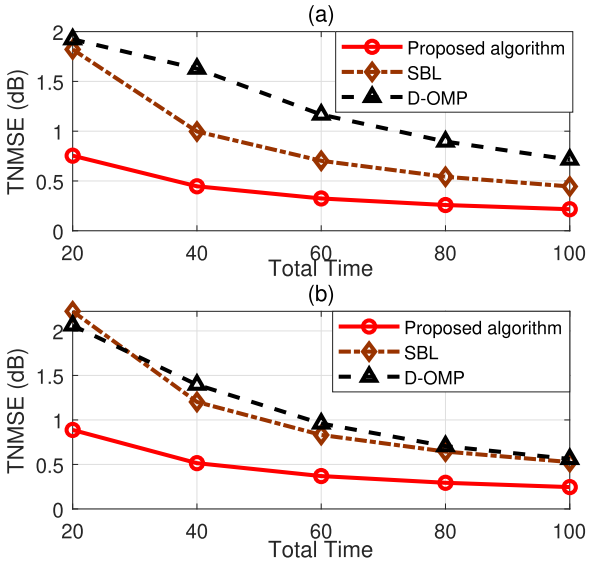

Fig. 8. TNMSE versus total time when SNR $= 15$dB, $N_s = 22$. (a) $M = 64, N_x = N_y = 8$. (b) $M = 256, N_x = N_y = 16$.

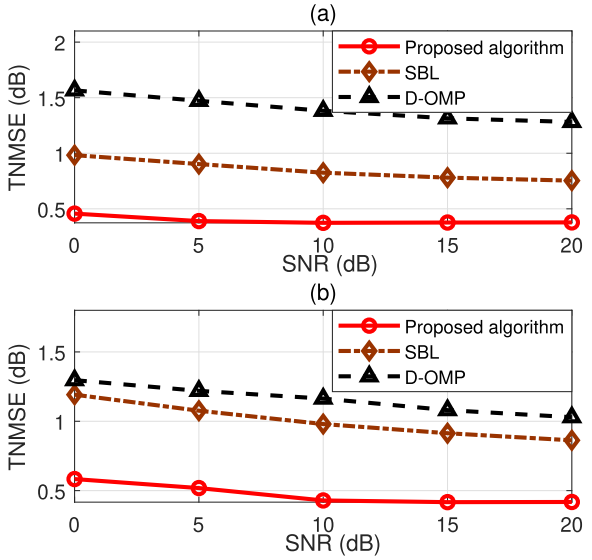

Fig. 9. TNMSE versus SNR when $N_s = 22, T = 50$. (a) $M = 64, N_x = N_y = 8$. (b) $M = 256, N_x = N_y = 16$.

In Fig. 8, the value of the TNMSE shows a downward trend with the total number of time increasing for all the methods. The reason is that the accumulated prior signal collected from the total period of time can provide additional information for accurate channel tracking. For different sizes of the antenna arrays, the performance metric trend maintains the same. It is easy to be noticed that the performance of the proposed algorithm is obviously improved compared with the benchmark algorithms SBL and D-OMP.

In Fig. 9, we further compare the TNMSE performance versus SNR when the total number of antenna array elements is $M = 64$ and 256, respectively. It can be observed that under the different numbers of antennas, the proposed algorithm achieves sufficient performance gain over the benchmark algorithms SBL and D-OMP. This proves the advantage of effectively exploiting the detail dynamic azimuth and elevation spatial sparsity structure of the channel in the proposed algorithm. At the same time, due to this strength, the size of the antenna has a rare influence on the proposed algorithm compared with the other two algorithms.

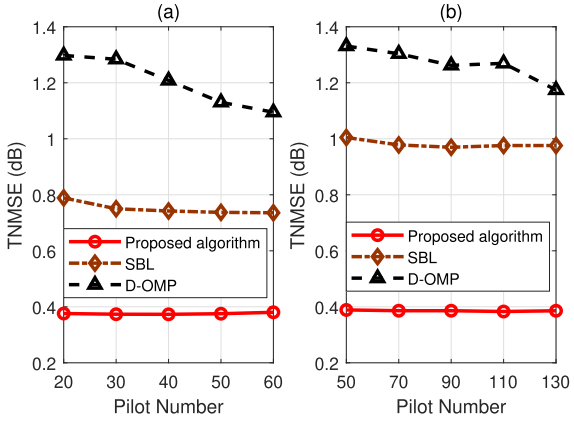

Fig. 10. TNMSE versus total pilot number when SNR $= 15$dB, $T = 50$. (a) $M = 64, N_x = N_y = 8$. (b) $M = 256, N_x = N_y = 16$.

The impact of the pilot number is further compared in Fig. 10. It is obvious that although both SBL and D-OMP have better recovery performance with the pilot number increase, the proposed algorithm always remains the superior performance gain. Moreover, different from the other two algorithms, the proposed algorithm shows a near horizontal performance with the increase pilot size. This demonstrates that the proposed algorithm can achieve a solid performance with a smaller size of the pilot overhead. The reason is that conventionally, increasing pilot number can improve the recovery performance based on multiple estimations through time, while our proposed algorithm can efficiently achieve good accuracy by exploring the dynamic channel structure.

VI. Conclusion

In this paper, we investigated the 3D channel modelling and tracking for Ka-band UAV-satellite communications with the orbit movement of a LEO satellite and 3D UAV trajectory in SAGIN. The spatial d on the continuous relative displacement between the UAV and satellite, and the temporal correlation of the dynamic channel were explored. We proposed a statistical dynamic channel model called 3D-2D-MM in which the insight probability relationship of both hidden value vector and joint hidden support vector have been developed. We further designed a novel 3D-DTAMP algorithm to recursively track the dynamic channel with additional 3D-2D-MM prior information. The benefits of the proposed algorithm can be summarized in three aspects. Firstly, the message passing rule has taken the more realistic 3D channel environment into account. Secondly, the consideration of the sparsity structure of the dynamic channel from both spatial and temporal domains have enhanced the recovery performance. Finally, it was more comparable in tracking a longer time series dynamic channel with lower pilot overhead. Analytical and numerical results have shown that the proposed algorithm achieved the superior recovery performance against two other benchmark algorithms as well as required less pilot overhead and time complexity.

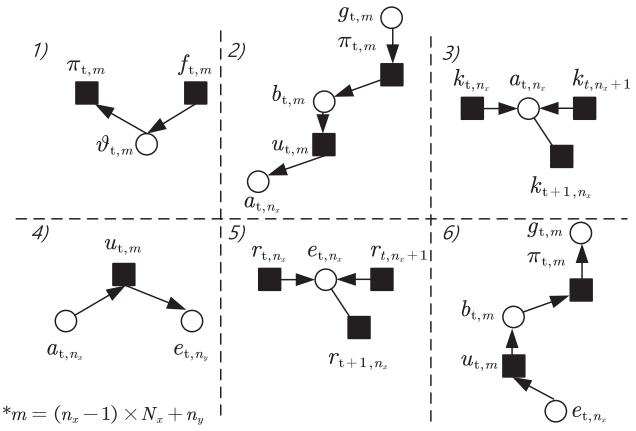

Fig. 11. The message passing sequence paths in Module $\mathcal{B}$.

APPENDIX A
3D-2D-MM MESSAGE PASSING IN MODULE $\mathcal{B}$

The details of the message passing sequence paths are shown in Fig. 11.

1) Message Passing Over Gaussian Hidden Value Vector $\boldsymbol{\vartheta}_t$: The message $\upsilon_{f_{t,m}\to\vartheta_{t,m}}$ passed from the last time slot can be formed as $\mathcal{CN}(\vartheta_{t,m};\mu_{t,m}^{dyna},(\sigma_{t,m}^{dyna})^2)$. The message passed from variable node $\vartheta_{t,m}$ to factor node $\pi_{t,m}$ can be formed as

$$\upsilon_{\vartheta_{t,m}\to\pi_{t,m}} = \mathcal{CN}(\vartheta_{t,m};\mu_{t,m}^{out},(\sigma_{t,m}^{out})^2), \tag{38}$$

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

2) Message Passing Over $g_{t,m}\to\pi_{t,m}\to b_{t,m}\to u_{t,m}\to a_{t,n_x}$: The message from variable node $g_{t,m}$ to factor node $\pi_{t,m}$ is

$$\upsilon_{g_{t,m}\to\pi_{t,m}}(g_{t,m}) = \mathcal{CN}\left(g_{t,m};g_{\mathcal{B},t,m}^{pri},v_{\mathcal{B},t,m}^{pri}\right). \tag{39}$$

The message from factor node $\pi_{t,m}$ to variable node $b_{t,m}$ is

$$\upsilon_{\pi_{t,m}\to b_{t,m}}(b_{t,m}) = \rho_{b,t,m}^{in}\delta(b_{t,m}-1) + \left(1-\rho_{b,t,m}^{in}\right)\delta(b_{t,m}), \tag{40}$$

where $\rho_{\boldsymbol{b}_t,t,m}^{in} = \left(1+\frac{\mathcal{CN}(0;g_{\mathcal{B},t,m}^{pri},v_{\mathcal{B},t}^{pri})}{\mathcal{CN}(0;g_{\mathcal{B},t,m}^{pri}-\mu_{t,m}^{out},v_{\mathcal{B},t}^{pri}+(\sigma_{t,m}^{out})^2)}\right)^{-1}$. The message $\upsilon_{b_{t,m}\to u_{t,m}}$ from variable node $b_{t,m}$ to factor node $u_{t,m}$ is the same as $\upsilon_{\pi_{t,m}\to b_{t,m}}(b_{t,m})$. The message from factor node $u_{t,m}$ to variable node a_{t,n_x} is

$$\upsilon_{u_{t,m}\to a_{t,n_x}}(a_{t,n_x}) = \rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in}\delta(a_{t,n_x}-1) + \left(1-\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in}\right)\delta(a_{t,n_x}), \tag{41}$$

where

$$\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in} = \frac{\rho_{\boldsymbol{a}_t}}{1+\rho_{\boldsymbol{a}_t}-\rho_{\boldsymbol{b}_t,t,m}^{in}}, \tag{42}$$

with

$$\rho_{\boldsymbol{a}_t} = \gamma_{\boldsymbol{b}_t}\rho_{\boldsymbol{b}_t,t,m}^{in}\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out} + (1-\gamma_{\boldsymbol{b}_t})(1-\rho_{\boldsymbol{b}_t,t,m}^{in})\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out} + (1-\rho_{\boldsymbol{b}_t,t,m}^{in})(1-\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out}) \tag{43}$$

and $\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out} = \upsilon_{e_{t,ny}\to u_{t,m}}(e_{t,n_y}=1)$.

3) Message Passing Over the 2D-MM of the Azimuth Support $\boldsymbol{a}_t$: We firstly introduce the details of the 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

$$\upsilon_{u_{t,m}\to a_{t,n_x}} = \rho_{t,n_x}^{(A)in} = \frac{\prod_{n_y}\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in}}{\prod_{n_y}\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in}+\prod_{n_y}\left(1-\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{in}\right)}, \tag{44}$$

which is deriviated based on the the matched dimension M from the total gain, n_x from the azimuth support and n_y from the elevation support.

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

$$\gamma_{t,n_x}^{(A)f} = \frac{q_{0,1}^{(AS)}\left(1-\rho_{t,n_x-1}^{(A)in}\right)\left(1-\gamma_{t,n_x-1}^{(A)f}\right)+q_{1,1}^{(AS)}\rho_{t,n_x-1}^{(A)in}\gamma_{t,n_x-1}^{(A)f}}{\left(1-\rho_{t,n_x-1}^{(A)in}\right)\left(1-\gamma_{t,n_x-1}^{(A)f}\right)+\rho_{t,n_x-1}^{(A)in}\gamma_{t,n_x-1}^{(A)f}}. \tag{45}$$

The azimuth backward parameter $\gamma_{t,n_x}^{(A)b}$ can be found in (46), as shown at the bottom of the page. When $t>1$,

$$\gamma_{t,n_x}^{(A)b} = \frac{q_{1,0}^{(AS)}\left(1-\rho_{t,n_x+1}^{(A)in}\right)\left(1-\gamma_{t,n_x+1}^{(A)b}\right)+\left(1-q_{1,0}^{(S)}\right)\rho_{t,n_x+1}^{(A)in}\gamma_{t,n_x+1}^{(A)b}}{\left(q_{0,0}^{(AS)}+q_{1,0}^{(S)}\right)\left(1-\rho_{t,n_x+1}^{(A)in}\right)\left(1-\gamma_{t,n_x+1}^{(A)b}\right)+\left(q_{1,1}^{(AS)}+q_{0,1}^{(AS)}\right)\rho_{t,n_x+1}^{(A)in}\gamma_{t,n_x+1}^{(A)b}} \tag{46}$$

$$\gamma_{t,n_x}^{(A)f} = \frac{q_{1,1,1}^{(A)}\rho_{t,n_x}^{dyna}+q_{1,0,1}^{(A)}\left(1-\rho_{t,n_x}^{(A)dyna}\right)}{1+\left(\left(\rho_{t,n_x-1}^{(A)in}\right)^{-1}-1\right)\left(\left(\gamma_{t,n_x-1}^{(A)f}\right)^{-1}-1\right)} + \frac{q_{0,1,1}^{(A)}!\rho_{t,n_x}^{(A)dyna}+q_{0,0,1}^{(A)}\left(1-\rho_{t,n_x}^{(A)dyna}\right)}{\left(\left(\rho_{t,n_x-1}^{(A)in}\right)^{-1}-1\right)^{-1}\left(\left(\left(\gamma_{t,n_x-1}^{(A)f}\right)^{-1}-1\right)\right)^{-1}+1} \tag{47}$$

$$\gamma^{(A)} = \frac{\left(1-\rho_{t,n_x+1}^{(A)in}-\gamma_{t,n_x+1}^{(A)b}\right)\left[q_{0,1,0}^{(A)}\rho_{t,n_x+1}^{(A)dyna}+q_{0,0,0}^{(A)}\left(1-\rho_{t,n_x+1}^{(A)dyna}\right)\right]+\rho_{t,n_x+1}^{(A)in}\gamma_{t,n_x+1}^{(A)b}}{\left(1-\rho_{t,n_x+1}^{(A)in}-\gamma_{t,n_x+1}^{(A)b}\right)\left[q_{1,1,0}^{(A)}\rho_{t,n_x+1}^{(A)dyna}+q_{1,0,0}^{(A)}\left(1-\rho_{t,n_x+1}^{(A)dyna}\right)\right]+\rho_{t,n_x+1}^{(A)in}\gamma_{t,n_x+1}^{(A)b}} \tag{48}$$

the azimuth forward parameter $\gamma_{t,n_x}^{(A)f}$ can be found in (47), as shown at the bottom of the previous page. The azimuth backward parameter $\gamma_{t,n_x}^{(A)b} = \frac{1}{1+\gamma^{(A)}}$ is given in (48), as shown at the bottom of the previous page.

The final azimuth message which represents the nonzero probability can be written as

$$\begin{aligned} \upsilon_{a_{t,n_x}\to u_{t,m}} &= \rho_{t,n_x}^{(A)out} \\ &= \frac{\gamma_{t,n_x}^{(A)f}\gamma_{t,n_x}^{(A)b}}{\gamma_{t,n_x}^{(A)f}\gamma_{t,n_x}^{(A)b} + \left(1-\gamma_{t,n_x}^{(A)f}\right)\left(1-\gamma_{t,n_x}^{(A)b}\right)}. \end{aligned} \quad (49)$$

4) Message Passing Over the Path $a_{t,n_x} \to u_{t,m} \to e_{t,n_y}$: The message from factor node a_{t,n_x} to variable node $u_{t,m}$ is

$$\begin{aligned} &\upsilon_{a_{t,n_x}\to u_{t,m}}(a_{t,n_x}) \\ &= \rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out}\delta(a_{t,n_x}-1) + (1-\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out})\delta(a_{t,n_x}), \end{aligned} \quad (50)$$

where

$$\begin{aligned} &\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out} \\ &= \frac{\rho_{t,n_x}^{(A)out}}{\rho_{t,n_x}^{(A)out} + \left(1-\rho_{t,n_x}^{(A)out}\right)\prod_{n_y'\neq n_y}\left(\left(\rho_{\boldsymbol{a}_t,t,n_x,n_y'}^{in}\right)^{-1}-1\right)}. \end{aligned} \quad (51)$$

To be noticed, $\rho_{t,n_x}^{(A)out}$ indicates the message passing over the n_x^{th} azimuth support element and $\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out}$ indicates the matching relationship between n_x^{th} azimuth support element and n_y^{th} elevation support element. This condition is similar for elevation support message passing equations. The message from factor node $u_{t,m}$ to variable node e_{t,n_y} is

$$\begin{aligned} &\upsilon_{u_{t,m}\to e_{t,n_y}}(e_{t,n_y}) \\ &= \rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in}\delta\left(e_{t,n_y}-1\right) + \left(1-\rho_{\boldsymbol{e},t,n_x,n_y}^{in}\right)\delta\left(e_{t,n_y}\right), \end{aligned} \quad (52)$$

where $\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in}$ can be found in (53), as shown at the bottom of the page.

5) Message Passing Over the 2D-MM of the Elevation Support $\boldsymbol{e}_t$: The details can be derived similarly as the message passing process of the azimuth support $\mathbf{a}_t$ estimation. The message passing from factor node $u_{t,m}$ to variable node e_{t,n_y} is

$$\begin{aligned} \upsilon_{u_{t,m}\to e_{t,n_y}} &= \rho_{t,n_y}^{(E)in} \\ &= \frac{\prod_{n_x}\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in}}{\prod_{n_x}\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in} + \prod_{n_x}\left(1-\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in}\right)}. \end{aligned} \quad (54)$$

When $t=1$, the elevation forward parameter is

$$\begin{aligned} &\gamma_{t,n_y}^{(E)f} \\ &= \frac{q_{0,1}^{(ES)}\left(1-\rho_{t,n_y-1}^{(E)in}\right)\left(1-\gamma_{t,n_y-1}^{(E)f}\right) + q_{1,1}^{(ES)}\rho_{t,n_y-1}^{(E)in}\gamma_{t,n_y-1}^{(E)f}}{\left(1-\rho_{t,n_y-1}^{(E)in}\right)\left(1-\gamma_{t,n_y-1}^{(E)f}\right) + \rho_{t,n_y-1}^{(E)in}\gamma_{t,n_y-1}^{(E)f}}. \end{aligned} \quad (55)$$

The elevation backward parameter $\gamma_{t,n_y}^{(E)b}$ can be found in (56), as shown at the bottom of the page. When $t>1$, the elevation forward parameter $\gamma_{t,n_y}^{(E)f}$ is given in (57), as shown at the bottom of the page. And the elevation backward parameter $\gamma_{t,n_y}^{(E)b} = \frac{1}{1+\gamma^{(E)}}$ with $\gamma^{(E)}$ given as (58), as shown at the bottom of the page.

The final elevation nonzero probability output can be written as

$$\begin{aligned} \upsilon_{e_{t,n_x}\to u_{t,m}} &= \rho_{t,n_y}^{(E)out} \\ &= \frac{\gamma_{t,n_y}^{(E)f}\gamma_{t,n_y}^{(E)b}}{\gamma_{t,n_y}^{(E)f}\gamma_{t,n_y}^{(E)b} + \left(1-\gamma_{t,n_y}^{(E)f}\right)\left(1-\gamma_{t,n_y}^{(E)b}\right)}. \end{aligned} \quad (59)$$

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

$$\begin{aligned} &\upsilon_{e_{t,n_y}\to u_{t,m}}(e_{t,n_y}) \\ &= \rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out}\delta\left(e_{t,n_y}-1\right) + \left(1-\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{out}\right)\delta\left(e_{t,n_y}\right), \end{aligned} \quad (60)$$

$$\rho_{\boldsymbol{e}_t,t,n_x,n_y}^{in} = \frac{\gamma_{\boldsymbol{b}_t}\rho_{\boldsymbol{b}_t,t,m}^{in}\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out} + (1-\gamma_{\boldsymbol{b}_t})(1-\rho_{\boldsymbol{b}_t,t,m}^{in})\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out} + (1-\rho_{\boldsymbol{b}_t,t,m}^{in})(1-\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out})}{\gamma_{\boldsymbol{b}_t}\rho_{\boldsymbol{b}_t,t,m}^{in}\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out} + (1-\gamma_{\boldsymbol{b}_t})(1-\rho_{\boldsymbol{b}_t,t,m}^{in})\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out} + (1-\rho_{\boldsymbol{b}_t,t,m}^{in})(1-\rho_{\boldsymbol{a}_t,t,n_x,n_y}^{out}) + 1 - \rho_{\boldsymbol{b}_t,t,m}^{in}} \quad (53)$$

$$\gamma_{t,n_y}^{(E)b} = \frac{q_{1,0}^{(ES)}\left(1-\rho_{t,n_y+1}^{(E)in}\right)\left(1-\gamma_{t,n_y+1}^{(E)b}\right) + \left(1-q_{1,0}^{(ES)}\right)\rho_{t,n_y+1}^{(E)in}\gamma_{t,n_y+1}^{(E)b}}{\left(q_{0,0}^{(ES)}+q_{1,0}^{(ES)}\right)\left(1-\rho_{t,n_y+1}^{(E)in}\right)\left(1-\gamma_{t,n_y+1}^{(E)b}\right) + \left(q_{1,1}^{(ES)}+q_{0,1}^{(ES)}\right)\rho_{t,n_y+1}^{(E)in}\gamma_{t,n_y+1}^{(E)b}} \quad (56)$$

$$\gamma_{t,n_y}^{(E)f} = \frac{q_{1,1,1}^{(E)}\rho_{t,n_y}^{(E)dyna} + q_{1,0,1}^{(E)}\left(1-\rho_{t,n_x}^{(E)dyna}\right)}{1+\left(\left(\rho_{t,n_y-1}^{(E)in}\right)^{-1}-1\right)\left(\left(\gamma_{t,n_y-1}^{(E)f}\right)^{-1}-1\right)} + \frac{q_{0,1,1}^{(E)}\rho_{t,n_y}^{(E)dyna} + q_{0,0,1}^{(E)}\left(1-\rho_{t,n_y}^{(E)dyna}\right)}{\left(\left(\rho_{t,n_y-1}^{(E)in}\right)^{-1}-1\right)^{-1}\left(\left(\left(\gamma_{t,n_y-1}^{(E)f}\right)^{-1}-1\right)\right)^{-1}+1} \quad (57)$$

$$\gamma^{(E)} = \frac{\left(1-\rho_{t,n_y+1}^{(E)in}-\gamma_{t,n_y+1}^{(E)b}\right)\left[q_{0,1,0}^{(E)}\rho_{t,n_y+1}^{(E)dyna} + q_{0,0,0}^{(E)}\left(1-\rho_{t,n_y+1}^{(E)dyna}\right)\right] + \rho_{t,n_y+1}^{(E)in}\gamma_{t,n_y+1}^{(E)b}}{\left(1-\rho_{t,n_y+1}^{(E)in}-\gamma_{t,n_y+1}^{(E)b}\right)\left[q_{1,1,0}^{(E)}\rho_{t,n_y+1}^{(E)dyna} + q_{1,0,0}^{(E)}\left(1-\rho_{t,n_y+1}^{(E)dyna}\right)\right] + \rho_{t,n_y+1}^{(E)in}\gamma_{t,n_y+1}^{(E)b}} \quad (58)$$

where

$$\rho^{out}_{\boldsymbol{e}_t,t,n_x,n_y} = \frac{\rho^{(E)out}_{t,n_y}}{\rho^{(E)out}_{t,n_y} + \left(1-\rho^{(E)out}_{t,n_y}\right)\prod_{n'_x\neq n_x}\left(\left(\rho^{in}_{\boldsymbol{e}_t,t,n'_x,n_y}\right)^{-1}-1\right)}. \tag{61}$$

The message from factor node $u_{t,m}$ to variable node $b_{t,m}$ is

$$\upsilon_{u_{t,m}\to b_{t,m}}(b_{t,m}) = \rho^{out}_{\boldsymbol{b}_t,t,m}\delta(b_{t,m}-1) + \left(1-\rho^{out}_{\boldsymbol{b}_t,t,m}\right)\delta(b_{t,m}), \tag{62}$$

where $\rho^{out}_{\boldsymbol{b}_t,t,m} = \rho^{out}_{\boldsymbol{a}_t,t,n_x,n_y}\rho^{out}_{\boldsymbol{e}_t,t,n_x,n_y}\gamma_{\boldsymbol{b}_t}$. The message $\upsilon_{b_{t,m}\to\pi_{t,m}}$ from variable node $b_{t,m}$ to factor node $\pi_{t,m}$ is the same as $\upsilon_{u_{t,m}\to b_{t,m}}(b_{t,m})$. The message from factor node $\pi_{t,m}$ back to variable node $g_{t,m}$ is

$$\upsilon_{\pi_{t,m}\to g_{t,m}}(g_{t,m}) = \rho^{out}_{\boldsymbol{b}_t,t,m}\mathcal{CN}\left(g_{t,m};\mu^{out}_{t,m},(\sigma^{out}_{t,m})^2\right) + \left(1-\rho^{out}_{\boldsymbol{b}_t,t,m}\right)\delta(g_{t,m}). \tag{63}$$

REFERENCES

[1] Z. Li *et al.*, "Energy efficient resource allocation for UAV-assisted space-air-ground Internet of remote things networks," *IEEE Access*, vol. 7, pp. 145348–145362, 2019.

[2] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2714–2741, Dec. 2018.

[3] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 101–109, Jul. 2017.

[4] N. Hosseini and D. W. Matolak, "Software defined radios as cognitive relays for satellite ground stations incurring terrestrial interference," in *Proc. IEEE CCAA*, Jun. 2017, pp. 1–4.

[5] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[6] S. M. R. Jones, R. E. Sheriff, and P. A. Watson, "Ka-band mobile-satellite service propagation environment," in *Proc. IEE Colloq. Future Ka Band Satell. Commun.*, Nov. 1993, pp. 1–7.

[7] S. Gong, D. Wei, X. Xue, and M. Y. Chen, "Study on the channel model and BER performance of single-polarization satellite-Earth MIMO communication systems at ka band," *IEEE Trans. Antennas Propag.*, vol. 62, no. 10, pp. 5282–5297, Oct. 2014.

[8] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[9] X. Cheng *et al.*, "Space/aerial-assisted computing offloading for IoT applications: A learning-based approach," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 5, pp. 1117–1129, May 2019.

[10] A. Varasteh *et al.*, "Toward optimal mobility-aware VM placement and routing in space-air-ground integrated networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2019, pp. 1–2.

[11] J. Zhao, F. Gao, Q. Wu, S. Jin, Y. Wu, and W. Jia, "Beam tracking for UAV mounted SatCom on-the-move with massive antenna array," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 2, pp. 363–375, Feb. 2018.

[12] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[13] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 164–174, Nov. 2010.

[14] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.

[15] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104–118, Jan. 2015.

[16] A. Liu, V. K. N. Lau, and W. Dai, "Exploiting burst-sparsity in massive MIMO with partial channel support information," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7820–7830, Nov. 2016.

[17] L. Chen, A. Liu, and X. Yuan, "Structured turbo compressed sensing for massive MIMO channel estimation using a Markov prior," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4635–4639, May 2018.

[18] X. Zhu, L. Dai, G. Gui, W. Dai, Z. Wang, and F. Adachi, "Structured matching pursuit for reconstruction of dynamic sparse channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–5.

[19] X. Zhu, L. Dai, W. Dai, Z. Wang, and M. Moonen, "Tracking a dynamic sparse channel via differential orthogonal matching pursuit," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2015, pp. 792–797.

[20] X. Gao, L. Dai, Y. Zhang, T. Xie, X. Dai, and Z. Wang, "Fast channel tracking for terahertz beamspace massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5689–5696, Jul. 2017.

[21] J. Zhao, F. Gao, W. Jia, S. Zhang, S. Jin, and H. Lin, "Angle domain hybrid precoding and channel tracking for millimeter wave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6868–6880, Oct. 2017.

[22] Q. Qin, L. Gui, P. Cheng, and B. Gong, "Time-varying channel estimation for millimeter wave multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9435–9448, Oct. 2018.

[23] L. Lian, A. Liu, and V. K. N. Lau, "Exploiting dynamic sparsity for downlink FDD-massive MIMO channel tracking," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2007–2021, Apr. 2019.

[24] R. He *et al.*, "Propagation channels of 5G millimeter-wave vehicle-to-vehicle communications: Recent advances and future challenges," *IEEE Veh. Technol. Mag.*, vol. 15, no. 1, pp. 16–26, Mar. 2020.

[25] R. He, B. Ai, G. L. Stuber, G. Wang, and Z. Zhong, "Geometrical-based modeling for millimeter-wave MIMO mobile-to-mobile channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 2848–2863, Apr. 2018.

[26] W. Shi *et al.*, "Multi-drone 3D trajectory planning and scheduling in drone assisted radio access networks," 2019, *arXiv:1906.00777*. [Online]. Available: http://arxiv.org/abs/1906.00777

[27] H. Fenech, S. Amos, and T. Waterfield, "The role of array antennas in commercial telecommunication satellites," in *Proc. 10th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1–4.

[28] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8052–8066, Dec. 2017.

[29] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5270–5284, Nov. 2013.

[30] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

[31] P. Schniter, "Turbo reconstruction of structured sparse signals," in *Proc. 44th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2010, pp. 1–6.

[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.

Jiadong Yu (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Dalian Maritime University, and the M.S. degree (Hons.) in wireless communication from the University of Southampton. She is currently pursuing the Ph.D. degree with the Department of Electronic Engineering and Computer Science, Queen Mary University of London. Her research interests include compressive sensing and millimeter wave communications.

Xiaolan Liu (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in communication engineering from Jilin University. She is currently pursuing the Ph.D. degree with the Department of Electronic Engineering and Computer Science, Queen Mary University of London. Her research interests include energy harvesting and machine learning for IoT networks with low-power communication technology, and millimeter wave communications.

Yue Gao (Senior Member, IEEE) received the Ph.D. degree from the Queen Mary University of London (QMUL), U.K., in 2007.

He was a Lecturer, a Senior Lecturer, and a Reader in antennas and signal processing with QMUL. He is currently a Professor of wireless communications with the Institute for Communication Systems, University of Surrey, U.K. He also 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 more than 180 peer-reviewed journal articles and conference papers, three patents, one book, and five book chapters and three best paper awards. He is also 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 has served as the Signal Processing for Communications Symposium Co-Chair for the IEEE ICCC 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 also 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, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the IEEE TRANSACTIONS ON COGNITIVE NETWORKS.

Xuemin (Sherman) Shen (Fellow, IEEE) 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 member of IEEE Fellow Selection Committee. He 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. He received the R. A. Fessenden Award in 2019 from the IEEE, Canada, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) presents in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015 and the Education Award in 2017 from the IEEE Communications Society, and the Technical Recognition Award from Wireless Communications Technical Committee in 2019 and AHSN Technical Committee in 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 has 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 Symposia Chair for the IEEE ICC'10, and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications. He is the Elected IEEE Communications Society Vice President for Technical and Educational Activities, the Vice President for Publications, a Member-at-Large on the Board of Governors, and the Chair of the Distinguished Lecturer Selection Committee. He was/is the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, the IEEE NETWORK, *IET Communications*, and *Peer-to-Peer Networking and Applications*.