This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

Location Tracking for Reconfigurable Intelligent Surfaces Aided Vehicle Platoons: Diverse Sparsities Inspired Approaches

Yuanbin Chen, Ying Wang, *Member, IEEE,* Xufeng Guo, Zhu Han, *Fellow, IEEE,* and Ping Zhang, *Fellow, IEEE*

***Abstract*—In this paper, we investigate the employment of reconfigurable intelligent surfaces (RISs) into vehicle platoons, functioning in tandem with a base station (BS) in support of the high-precision location tracking. In particular, the use of a RIS imposes additional structured sparsity that, when paired with the initial sparse line-of-sight (LoS) channels of the BS, facilitates beneficial group sparsity. The resultant group sparsity significantly enriches the energies of the original direct-only channel, enabling a greater concentration of the LoS channel energies emanated from the same vehicle location index. Furthermore, the burst sparsity is exposed by representing the non-line-of-sight (NLoS) channels as their sparse copies. This thus constitutes the philosophy of the diverse sparsities of interest. Then, a diverse dynamic layered structured sparsity (DiLuS) framework is customized for capturing different priors for this pair of sparsities, based upon which the location tracking problem is formulated as a maximum a posterior (MAP) estimate of the location. Nevertheless, the tracking issue is highly intractable due to the ill-conditioned sensing matrix, intricately coupled latent variables associated with the BS and RIS, and the spatial-temporal correlations among the vehicle platoon. To circumvent these hurdles, we propose an efficient algorithm, namely DiLuS enabled spatial-temporal platoon localization (DiLuS-STPL), which incorporates both variational Bayesian inference (VBI) and message passing techniques for recursively achieving parameter updates in a turbo-like way. Finally, we demonstrate through extensive simulation results that the localization relying exclusively upon a BS and a RIS may achieve the comparable precision performance obtained by the two individual BSs, along with the robustness and superiority of our proposed algorithm as compared to various benchmark schemes.**

***Index Terms*—Reconfigurable intelligent surface, location tracking, vehicle networks, diverse sparsities.**

I. Introduction

With a tidal wave of technological innovation upon us, the sixth-generation (6G) communication system tends to be a game changer in terms of placating the extraordinary promises, particularly in support of enormous amounts of widely heterogeneous information exchange in a hyperfast, low-latency and ultra-reliable manner. A grand 6G paradigm needs to be materialized by revolution-natured technologies, such as terahertz (THz) communications, reconfigurable intelligent surfaces (RISs), and artificial intelligence, in an effort to advance the existing mechanisms and accommodate the newly encountered particularities [1], [2]. Reaping plethora of benefits in 6G, connected autonomous vehicles (CAVs) have the potential to provide unparalleled quality of experiences, tremendously improved traffic conditions, road safety, and an abundance of cutting-edge vehicular applications [3]. This relies significantly upon a high-definition degree of situational awareness with the physical environment, i.e., the ability to identify its own position. In response, standardization lays the foundation of localization use cases and specifications for sidelinks (e.g., vehicle platoons and coordinated maneuvers), including the recently frozen 3rd generation partnership project (3GPP) Release 17, in an effort to further improve the fifth-generation (5G) localization for fulfilling more stringent criteria, e.g., centimeter-level accuracy [4], [5]. The existing vehicle localization technologies, such as global positioning system (GPS), radar, lidar and inertial sensors, however, fall well short of achieving these specifications, thus mandating rethinking ways and new contributing technologies that go beyond the present ones.

The common localization techniques can be informally categorized into indirect approaches and direct thereof. In particular, the indirect localization aims to obtain the intermediate parameters, e.g., angles of arrival (AoA), time of arrival (ToA), and received signal strength (RSS), from the received signals before applying the trick of triangulation to determine user's location. For instance, the data-driven methods including deep learning and fingerprinting are employed in [6] and [7], respectively, to estimate the RSS in a massive multiple-input multiple-output (MIMO) system. It is investigated in [8] that a joint estimate of AoA and ToA is considered for establishing user's location. The direct localization, by contrast, recovers the associated positions straightforwardly from the received signal, which is applicable exclusively in the pure line-of-sight (LoS) environment but can also be extended to the multi-path scenario, based upon which such a geometry-based approach allows the user's location to be constructed by taking full advantage of various LoS and non-line-of-sight (NLoS) components [9], [10]. Furthermore,

Manuscript received 31 August 2022; revised 31 March 2023; accepted 3 May 2023. This work was supported by Beijing Natural Science Foundation under Grant 4222011, and in part by NSF CNS-2107216, CNS-2128368, CMMI-2222810, US Department of Transportation, Toyota and Amazon. *(Corresponding author: Ying Wang.)*

Yuanbin Chen, Ying Wang, Xufeng Guo, and Ping Zhang are with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: chen_yuanbin@163.com; wangying@bupt.edu.cn; brook1711@bupt.edu.cn; pzhang@bupt.edu.cn).

Zhu Han is with the Department of Electrical and Computer Engineering in the University of Houston, Houston, TX 77004, USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea, 446-701 (e-mail: hanzhu22@gmail.com).

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

the application of the two localization techniques facilitates the location tracking in the context of vehicular networks [11]–[13], including the geometry-based cooperative localization [11], indirect approach-based tracking [12], and data-driven learning-based method [13]. Owing to the high degrees of freedom and large apertures in the array equipped at the base station (BS), the localization accuracy may be improved if used in conjunction with an appropriate LoS/NLoS model that is crafted to appropriately characterize the vehicular multi-path environment. While these initiatives share common facts with the collaborative efforts of multiple BSs, they may, in practice, achieve so at the expense of unacceptable signaling overhead and infrastructure costs.

Benefiting from the digitally controlled meta-surfaces, RIS has been hailed as an impressive technology capable of transforming the conventional uncontrollable wireless channels into their controllable counterparts, thus materializing the concept of "smart and customizable radio environment" [14]–[16]. Intriguingly, a tight amalgamation of RIS with localization is capable of making significant contributions to several challenging issues, with a number of fledgling research efforts already available to date [17]–[24]. More explicitly, in [17], the use of a RIS enables the exact construction of distinct RSS values across adjacent locations by appropriately adjusting the phase shifts of the RIS in the presence of only one access point (AP) indoors. It is investigated in [18] that a joint localization and synchronization paradigm is conceived based on an optimized design of the BS active precoding and RIS passive beamforming. A RIS is deployed in [19] for investigating the capability of positioning a single-antenna user by taking into account the mobility of the user and spatial-wideband effects in a single-input and single-output (SISO) system. Furthermore, RIS can be treated as a known anchor point that may be utilized for multi-path augmentation in pure LoS scenarios for the improved localization accuracy [20]–[22], while also functioning excellently in NLoS scenarios in the presence of obstructions by introducing beneficial LoS paths [23], [24]. Therefore, as can be succinctly concluded from the above attempts, RIS-aided localization is primarily characterized by the various advantages provided by the varied roles that RIS plays: i) as a known anchor point, a RIS has the potential to introduce additional location reference, based upon which the established priori facilitates the retrieval of the exact posterior associated with user's location using the Bayesian framework [22], [25], and ii) due to its ability to integrate into the wireless environment functioning as any scatterer or reflector, a RIS is capable of delivering additional observations, independent of the uncontrolled and passive multiple paths, which relaxes the dependence of scatterers that is required for the standard multi-path-based positioning [26].

In spite of the potential benefits of integrating RISs in vehicular networks, achieving the high-accuracy and real-time localization remains an open issue, as evidenced by the following facts. Firstly, since the LoS channels between the vehicle and different infrastructures (e.g., BSs and RISs) originate from the same vehicle location, the LoS channel energies associated with the BS and RIS are all concentrated on the same location index, which induces a group sparsity structure of the LoS channels. The burst sparsity structure inherent to the angular NLoS channels also provides additional angle references contributing to the improved localization accuracy. Therefore, an appropriate probabilistic model is required for characterizing the heterogeneous LoS/NLoS channels with diverse sparsities in support of the high-precision localization. Secondly, concerning that vehicle platoons are typically featured by vehicles advancing shoulder-to-shoulder at approximately the same speed, vehicle positions are highly spatial-correlated. With the passage of time, the previously obtained vehicle position delivers rich priors for the present tracking, and hence the use of such probability-dominated random correlations has the potential to achieve a high-accuracy location tracking with a reduced overhead. Furthermore, this pair of critical facts aforementioned motivate us to use sparse Bayesian-like approaches for the real-time location tracking. Nevertheless, the conventional Bayesian-based methods, e.g., variational Bayesian inference [27] and pure approximate message passing (AMP) [28], may be fruitless to cope with the elusive priors due to the tight coupling of the direct and cascaded location supports with various structured sparsities. Accordingly, the entries in the BS- and RIS-associated sensing matrices are not perfectly orthogonal, which renders them ill-conditioned. Consequently, a location tracking scheme has to be tailored for encapsulating LoS/NLoS channels with various structured sparsities, sufficiently robust in the presence of the sophisticated priors and highly tractable for enabling efficient algorithm design. In a nutshell, we target these non-trivial challenges in this paper to conceive a robust and high-accuracy tracking strategy for the RIS-aided vehicle platoon system.

Geared towards the challenges mentioned above, the current work presents the following advancements in the state-of-the-art:

- We investigate a scenario of vehicle platoon with a BS and a RIS functioning in tandem for the real-time and high-accuracy localization. The employment of a RIS imposes additional structured sparsity, which, when paired with the initial sparse LoS channels of the BS, facilitates the group sparsity. The group sparsity contributed by both enriches the energies of the original direct-only channel, resulting in a greater concentration of the LoS channel energies associated with the BS and RIS at the same location index that is based on a set of uniformed road grids prescribed by the off-grid basis. By contrast, the NLoS channels can be represented as their sparse copies, revealing the burst sparsity. This thus explicates the philosophy of the diverse sparse sparsities, which allows the original location tracking problem to be pruned as a joint estimate of the off-grid location and angle offsets. Additionally, by leveraging the geometric dilution of precision (GDOP) metric, we evince that the application of RISs can reduce the estimation error to some extent.
- A hierarchical framework, namely **D**iverse dynam**i**c **L**ayer str**u**ctured **S**parsity (DiLuS), is crafted to concurrently capture diverse sparsities while taking into account the spatial-temporal correlations among the vehicles. We

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

use a pair of layered probabilistic models to characterize these diverse sparsities inherent to the LoS/NLoS channels, each of which functions independently before being assimilated into the conditional probability of the received localization signal, thus facilitating high-accuracy location tracking using the adequate priors available. To this end, we aim at obtaining the maximum likelihood estimation (MLE) of the off-grid location and AoA offsets as well as the minimum mean square error (MMSE) estimates of the sparse LoS/NLoS channel vectors.

- The investigated problem is a quasi-compressive sensing (CS) copy, but it differs significantly from the standard CS-based problem due to its ill-conditioned sensing matrices (containing the offsets to be estimated) and dynamic sparse priors furnished by DiLuS, demonstrating a high level of intractability. To attain a stable solution for this problem, we develop a DiLuS-STPL algorithm that blends the techniques of variational Bayesian inference (VBI) and message passing to recursively achieve the offset updates in a turbo-like way while taking into account the spatial-temporal correlations of the vehicle platoon. The proposed DiLuS-STPL offers the capability of fully harnessing the diverse dynamic sparse priors as enveloped by DiLuS despite the presence of uncertain sensing matrices.
- Simulation results reveal that the RIS is a powerful facilitator for robust and high-accuracy localization by contributing to beneficial location references, in particular for its nature of low cost and ease of implementation. We elucidate that, under the same configurations, the cooperation between a BS and a RIS achieves the comparable localization precision obtained by the two individual BSs as measured by GDOP. In addition, our proposed DiLuS-STPL algorithm performs desirably even at a coarse grid resolution, and we show that it can mitigate the NLoS-induced misleading effect while attempting to identify the LoS channels in the presence of a fluctuating number of NLoS paths, thus unveiling the critical robustness for high-accuracy localization.

The remainder of this paper is structured as follows. Section II introduces the system model and off-grid basis representation. In Section III, we present the DiLuS framework for capturing diverse sparse priors. In Section IV, the DiLuS-STPL algorithm is proposed to effectively resolve the formulated location tracking problem, followed by the simulation results provided in Section V. Finally, concluding remarks are drawn in Section VI.

Notation: Scalar variables, column vectors and matrices are represented by Italic letters, boldface lower-case and capital letters such as x, $\mathbf{x}$, and $\mathbf{X}$, respectively. The upper-case calligraphic letters such as $\mathcal{M}$ is employed to indicate finite and discrete sets. $\mathbb{C}^{M\times N}$ represents the space of a $M\times N$ complex-valued matrix. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ represent the conjugate, transpose, and Hermitian (conjugate transpose) operators of their arguments, respectively. $\|\cdot\|_2$ and $\|\cdot\|_F$ represent the spectrum norm and Frobenius norm of their arguments, respectively. For a vector $\mathbf{x}$, $\mathrm{diag}(\mathbf{x})$ returns a diagonal matrix with the entries of $\mathbf{x}$ located on its main diagonal. The symbol $\otimes$ stands for the Kronecker product. The imaginary unit is given by $\jmath=\sqrt{-1}$. The expectation operator is denoted by $\mathbb{E}\{\cdot\}$. Regarding $\mathcal{CN}(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$, we indicate that the complex random vector $\mathbf{x}$ is distributed in accordance with a Gaussian probability density function (PDF) with a mean vector $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$. The symbol $\sim$ stands for "distributed as".

II. System Model

A. Localization Scenario

We consider a vehicle platoon composed of M single-antenna vehicle user equipments (VUEs) indexed by $\mathcal{M}=\{1,...,M\}$, with these VUEs moving at approximately the same velocity. The vehicle platoon is within the communication coverage of a BS equipped with a uniform linear array (ULA) having K antennas, in which a RIS that is an N-element uniform planar array (UPA) takes over the role for enhancing the communication and localization services between the BS and VUEs. The system being investigated divides the continuous time into a number of slots indexed by $\mathcal{T}=\{1,...,T\}$. For a two-dimensional (2D) geographic area, the mth VUE is located at $\mathbf{p}_m^t=[x_m^t,y_m^t,z_m^t]^T$ along a road in the tth slot, and the center of the array gravity for the BS and the RIS are given by $\tilde{\mathbf{p}}^{\mathrm{BS}}=\left[x^{\mathrm{BS}},y^{\mathrm{BS}},z^{\mathrm{BS}}\right]^T$ and $\tilde{\mathbf{p}}^{\mathrm{RIS}}=\left[x^{\mathrm{RIS}},y^{\mathrm{RIS}},z^{\mathrm{RIS}}\right]^T$, respectively. Note that such a 2D area must be known as a priori such that the road map can be uniformed as grids to establish the exact location of each VUE. Due to the intricacy of the traffic situation, a road map suffices for the precise localization instead of the comprehensive information on the road's dynamics.

B. Channel Model

The general clustered delay line (CDL) model [29] is employed for the corresponding channels in our considered system[1]. With both highly elevated BS and RIS having a large number of antennas and elements, the spatial resolution increases significantly in the presence of the angular basis, and the channel becomes much more sparse in this case. The channels spanning from the mth VUE to the BS $\mathbf{h}_{m,b}^t\in\mathbb{C}^{K\times 1}$ and that spanning from the mth VUE to the RIS $\mathbf{h}_{m,r}^t\in\mathbb{C}^{N\times 1}$ are, respectively, given by

$$\mathbf{h}_{m,b}^t=\beta_m^t\mathbf{a}_K\left(\omega_m^{\mathrm{AoA}}\right)+\sum_{l=1}^{L_{m,b}}\tilde{\beta}_{m,l}^t\mathbf{a}_K\left(\omega_{m,l}^{\mathrm{AoA}}\right),\forall m,t, \quad (1)$$

[1]We consider a narrowband system in this paper, and the proposed algorithm can be readily extended to the wideband system. More precisely, the DiLuS paradigm still holds for the channels over each sub-carrier, and our proposed algorithm can be directly harnessed to track VUE's locations employing the measurements received from the BS at this specific sub-carrier. Furthermore, the algorithm can be extended to incorporate the ToA information in wideband system by introducing a 2D grid of both AoAs and ToAs, based upon which the ToA offsets can be likewise attained by MLE. One may refer to [30] for related details in wideband systems. However, in the mobility scenario, the increased grid dimension induced by ToA information significantly enlarges the search space and exacerbates the computational burden in a short coherent interval. To facilitate simplicity and clarity, we concentrate on a narrowband system in this paper to strive for a critical trade-off between the localization performance and complexity for the RIS-aided vehicle platoon system.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$$\mathbf{h}_{m,r}^{t}=\eta_{m}^{t}\mathbf{a}_{N}\left(\varphi_{m}^{\mathrm{AoA}},\vartheta_{m}^{\mathrm{AoA}}\right)+\sum_{l=1}^{L_{m,r}}\tilde{\eta}_{m,l}^{t}\mathbf{a}_{N}\left(\varphi_{m,l}^{\mathrm{AoA}},\vartheta_{m,l}^{\mathrm{AoA}}\right),\forall m,t. \tag{2}$$

In (1) and (2), $L_{m,b}$ and $L_{m,r}$ represent the number of propagation paths for their individual channels, respectively. For the channel spanning from the mth VUE to the BS, β_m^t and ω_m^{AoA} denote the complex gain and AoA of the LoS path, while $\tilde{\beta}_{m,l}^t$ and $\omega_{m,l}^{\mathrm{AoA}}$ indicate the counterparts of the lth NLoS path, respectively. Similarly, as for the channel spanning from the mth VUE to the RIS, η_m^t denotes the complex gain and $\left(\varphi_m^{\mathrm{AoA}},\vartheta_m^{\mathrm{AoA}}\right)$ denotes the azimuth-elevation AoA pair, whereas $\tilde{\eta}_{m,l}^t$ and $\left(\varphi_{m,l}^{\mathrm{AoA}},\vartheta_{m,l}^{\mathrm{AoA}}\right)$ represents the corresponding parameters of the lth NLoS path. The channel complex gains associated with the LoS channels are generated according to $\eta_m=\eta_0 d_{m,r}\exp\left(\jmath\frac{2\pi}{\lambda}d_{m,r}\right)$ and $\beta_m=\beta_0 d_{m,b}\exp\left(\jmath\frac{2\pi}{\lambda}d_{m,b}\right)$, respectively, where $d_{m,r}=\left\|\mathbf{p}_m^t-\tilde{\mathbf{p}}^{\mathrm{RIS}}\right\|^{-\zeta_{m,r}}$ and $d_{m,b}=\left\|\mathbf{p}_m^t-\tilde{\mathbf{p}}^{\mathrm{BS}}\right\|^{-\zeta_{m,b}}$ denote their individual link distances, η_0 and β_0 indicate the path-loss at the reference distance of one meter, as well as $\zeta_{m,r}$ and $\zeta_{m,b}$ for path-loss exponents. The NLoS counterparts are generated in compliance with Gaussian distributions with zero means and variances $\sigma_{\tilde{\eta}}$ and $\sigma_{\tilde{\beta}}$, i.e., $\tilde{\eta}_m^t\sim\mathcal{CN}\left(0,\sigma_{\tilde{\eta}}\right)$ and $\tilde{\beta}_m^t\sim\mathcal{CN}\left(0,\sigma_{\tilde{\beta}}\right)$, respectively [31]. The array steering vectors of the half-wavelength-space ULA and UPA at the BS and RIS are denoted by $\mathbf{a}_K\left(\omega\right)=\left[1,e^{-\jmath\pi\cos\omega},...,e^{-\jmath\pi(K-1)\cos\omega}\right]^T$ and $\mathbf{a}_N\left(\varphi,\vartheta\right)=\mathbf{a}_{N_x}\left(\varphi\right)\otimes\mathbf{a}_{N_y}\left(\vartheta\right)$, respectively. Furthermore, we denote $\mathbf{H}_{r,b}\in\mathbb{C}^{K\times N}$ by the channel spanning from the RIS to the BS which is available in advance due to their fixed positions. The mapping between the Cartesian coordinate and the spherical coordinate regulates the conversion between corresponding angles and coordinates, e.g., $\omega_m^{\mathrm{AoA}}\left(\mathbf{p}_m^t\right)$, $\varphi_m^{\mathrm{AoA}}\left(\mathbf{p}_m^t\right)$, and $\vartheta_m^{\mathrm{AoA}}\left(\mathbf{p}_m^t\right)$, which is omitted here for conciseness. Additionally, as for the narrowband system of interest in this paper, the location information of each VUE can still be gathered from the received signal at the BS, even if no information can be retrieved from the ToA. In this case, we just need to acquire the angle-related information to determine the exact location of each VUE by applying the trick of triangulation for such a non-synchronous system [26].

At each time slot t, all the VUEs simultaneously transmit G information-carrying pilot sequences $x_m^t\left(g\right),g=1,...,G,\forall t$, for localization, in which $x_m^t\left(g\right)$ represents the gth pilot sequence transmitted from the mth VUE in the tth slot. Thus, the signals propagating across the multi-path environment and being reflected by the RIS and eventually being received at the BS, are given by

$$\mathbf{y}^t\left(g\right)=\sum_{m=1}^{M}\left(\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{h}_{m,r}^t+\mathbf{h}_{m,b}^t\right)x_m^t\left(g\right)+\mathbf{n}^t\left(g\right),\quad g=1,...,G,\forall t, \tag{3}$$

where $\mathbf{n}^t\left(g\right)\sim\mathcal{CN}\left(\mathbf{0},\sigma^2\mathbf{I}\right)$ is the additive white Gaussian noise (AWGN) with power σ^2 at the BS. We denote $\mathbf{\Theta}=\mathrm{diag}\left(\boldsymbol{\theta}\right)\in\mathbb{C}^{N\times N}$ by the diagonal phase shift matrix at the RIS, where $\boldsymbol{\theta}\in\mathbb{C}^{N\times 1}$ is the phase shift vector with its nth entry satisfying $\left|\theta_n\right|^2=1,1\le n\le N$[2]. Additionally, it is assumed that the large-scale parameters, such as η_m^t, β_m^t, φ_m^{AoA}, $\vartheta_m^{\mathrm{AoA}}$, and ω_m^{AoA}, remain approximately constant within a slot of interest. This assumption is practically reasonable since in comparison to the nominal distance with a remote BS, the traveling distance of the vehicle within one slot is marginal, resulting in subtle variations in the geometry-associated parameters that impact the large-scale characteristics [29], [32]. This potentially relaxes the frequency of instantaneous channel information updates, thus achieving pilot savings. As regards the channel fluctuation on the slot-level of the order of hundreds of milliseconds, that of the VUE location, for instance, is on the scale of say 72 km/h (2 cm/ms).

C. Off-grid Basis Representation

Generally, the maximum likelihood (ML) or least square (LS) method may be used to directly recover the location of VUE from the received signal at the BS. However, such a strategy is arduous due to its non-convexity-induced high intractability and the possibility of falling into a number of local optima. To resolve this issue, the location of VUE can be determined via its maximum a posteriori (MAP) estimation by exploiting the sparsity inherent to the grid basis model [11]. Specifically, we introduce a uniform grid of U positions along the center of the road, denoted by $\mathcal{U}$, for the location of VUE, which is given by $\mathcal{U}=\{\mathbf{r}_1,...,\mathbf{r}_U\}$. Meanwhile, a uniform grid of $\tilde{K}$ AoAs at the BS over $[-\pi/2,\pi/2)$ are prescribed to take values from the discrete sets: $\left\{\tilde{\omega}_{m,k}^{\mathrm{AoA}}:\sin\left(\tilde{\omega}_{m,k}^{\mathrm{AoA}}\right)=\frac{2}{\tilde{K}}\left(k-\left\lfloor\frac{\tilde{K}-1}{2}\right\rfloor\right),k=0,...,\tilde{K}-1,\forall m\right\}$. Nevertheless, the true VUE location and AoAs do not coincide exactly with the grid points. To capture the attributes of mismatches between the true angles and the grid points, we adopt an off-grid basis for the associated sparse representation. More explicitly, the AoA offset vector at the BS for the mth VUE is defined as $\Delta\boldsymbol{\omega}_m^{\mathrm{AoA}}=\left[\Delta\omega_{m,1}^{\mathrm{AoA}},...,\Delta\omega_{m,\tilde{K}}^{\mathrm{AoA}}\right]^T,\forall m$ with $\Delta\omega_{m,k}^{\mathrm{AoA}}$ given by

$$\Delta\omega_{m,k}^{\mathrm{AoA}}=\begin{cases}\omega_{m,k}^{\mathrm{AoA}}-\tilde{\omega}_{m,k_l}^{\mathrm{AoA}},k=k_l,l=1,...,L_{m,b},\forall m,\\0,\qquad\qquad\text{otherwise},\end{cases} \tag{4}$$

where k_l indicates the index of the AoA grid point nearest to the AoA of the lth NLoS path between the mth VUE and the BS. Then, let $\Delta\mathbf{r}_m^t=\left[\Delta\mathbf{r}_{m,1}^t;...;\Delta\mathbf{r}_{m,U}^t\right],\forall m$ denote the location offset vector with $\Delta\mathbf{r}_{m,u}^t$ given by

$$\Delta\mathbf{r}_{m,u}^t=\begin{cases}\mathbf{p}_m^t-\mathbf{r}_{m,q}^t,u=q,\forall m,\\\left[0,0\right]^T,\qquad\text{otherwise},\end{cases} \tag{5}$$

where q indicates the index of the off-grid location nearest to the uth counterpart for the mth VUE, i.e., the square with the center of the qth off-grid location $\mathbf{r}_{m,q}^t$ in $\mathcal{U}$. Furthermore, the azimuth-elevation AoA pair $\left(\boldsymbol{\varphi}_m^{\mathrm{AoA}},\boldsymbol{\vartheta}_m^{\mathrm{AoA}}\right)$ associated with the mth VUE at the RIS takes the values from

[2]It is indeed difficult to determine a RIS phase profile in the absence of any prior channel information. What we try best is to fix it when doing channel estimation in accordance with practicable principles (e.g., random/directional RIS profiles presented in [19]).

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

a discrete set of size $\tilde{N}$: $\left\{\left(\varphi_{m,n}^{\mathrm{AoA}}, \vartheta_{m,n}^{\mathrm{AoA}}\right) : n = 1, ..., \tilde{N}\right\}$. Similarly, let $\Delta\boldsymbol{\varphi}_m^{\mathrm{AoA}} = \left[\Delta\varphi_{m,1}^{\mathrm{AoA}}, ..., \Delta\varphi_{m,\tilde{N}}^{\mathrm{AoA}}\right]^T, \forall m$ and $\Delta\boldsymbol{\vartheta}_m^{\mathrm{AoA}} = \left[\Delta\vartheta_{m,1}^{\mathrm{AoA}}, ..., \Delta\vartheta_{m,\tilde{N}}^{\mathrm{AoA}}\right]^T, \forall m$ denote the offset vectors of the azimuth-elevation AoA pair at the RIS, respectively.

With the above definitions of off-grid location and AoA offsets, the off-grid basis for the LoS/NLoS array response at the BS can be reformulated as

$$\mathbf{A}_K^{\mathrm{LoS}}\left(\Delta\mathbf{r}_m^t\right) = \left[\mathbf{a}_K\left(\tilde{\omega}_{m,1}^{\mathrm{AoA}}\left(\mathbf{r}_{m,1}^t + \Delta\mathbf{r}_{m,1}^t\right)\right), ..., \mathbf{a}_K\left(\tilde{\omega}_{m,U}^{\mathrm{AoA}}\left(\mathbf{r}_{m,U}^t + \Delta\mathbf{r}_{m,U}^t\right)\right)\right] \in \mathbb{C}^{K\times U}, \quad (6)$$

$$\mathbf{A}_K^{\mathrm{NLoS}}\left(\Delta\boldsymbol{\omega}_m^{\mathrm{AoA}}\right) = \left[\mathbf{a}_K\left(\tilde{\omega}_{m,1}^{\mathrm{AoA}} + \Delta\omega_{m,1}^{\mathrm{AoA}}\right), ..., \mathbf{a}_K\left(\tilde{\omega}_{m,\tilde{K}}^{\mathrm{AoA}} + \Delta\omega_{m,\tilde{K}}^{\mathrm{AoA}}\right)\right] \in \mathbb{C}^{K\times\tilde{K}}. \quad (7)$$

Similarly, the off-grid basis for the LoS/NLoS array response at the RIS can be given by

$$\begin{aligned}&\mathbf{A}_N^{\mathrm{LoS}}\left(\Delta\mathbf{r}_m^t\right)\\ &= \left[\mathbf{a}_N\left(\varphi_{m,1}^{\mathrm{AoA}}\left(\mathbf{r}_{m,1}^t + \Delta\mathbf{r}_{m,1}^t\right), \vartheta_{m,1}^{\mathrm{AoA}}\left(\mathbf{r}_{m,1}^t + \Delta\mathbf{r}_{m,1}^t\right)\right),\right.\\ &\quad \left...., \mathbf{a}_N\left(\varphi_{m,U}^{\mathrm{AoA}}\left(\mathbf{r}_{m,U}^t + \Delta\mathbf{r}_{m,U}^t\right), \vartheta_{m,U}^{\mathrm{AoA}}\left(\mathbf{r}_{m,U}^t + \Delta\mathbf{r}_{m,U}^t\right)\right)\right]\\ &\in \mathbb{C}^{N\times U}, \end{aligned} \quad (8a)$$

$$\begin{aligned}&\mathbf{A}_N^{\mathrm{NLoS}}\left(\Delta\boldsymbol{\varphi}_m^{\mathrm{AoA}}, \Delta\boldsymbol{\vartheta}_m^{\mathrm{AoA}}\right)\\ &= \left[\mathbf{a}_N\left(\tilde{\varphi}_{m,1}^{\mathrm{AoA}} + \Delta\varphi_{m,1}^{\mathrm{AoA}}, \tilde{\vartheta}_{m,1}^{\mathrm{AoA}} + \Delta\vartheta_{m,1}^{\mathrm{AoA}}\right),\right.\\ &\quad \left...., \mathbf{a}_N\left(\tilde{\varphi}_{m,\tilde{N}}^{\mathrm{AoA}} + \Delta\varphi_{m,\tilde{N}}^{\mathrm{AoA}}, \tilde{\vartheta}_{m,\tilde{N}}^{\mathrm{AoA}} + \Delta\vartheta_{m,\tilde{N}}^{\mathrm{AoA}}\right)\right] \in \mathbb{C}^{N\times\tilde{N}}. \end{aligned} \quad (8b)$$

Therefore, the signal received from M VUEs at the BS can be recast as a CS form with indeterminate parameters in the sensing matrix, i.e.,

$$\mathbf{y}^t = \left[\begin{array}{cc}\mathbf{F}^t\left(\Delta\mathbf{r}^t\right) & \boldsymbol{\Xi}^t\left(\Delta\boldsymbol{\omega}^{\mathrm{AoA}}, \Delta\boldsymbol{\varphi}^{\mathrm{AoA}}, \Delta\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)\end{array}\right]\left[\begin{array}{c}\mathbf{z}^t\\ \mathbf{v}^t\end{array}\right] + \mathbf{n}^t, \quad (9)$$

where $\mathbf{y}^t = [\mathbf{y}^t(1); ...; \mathbf{y}^t(G)] \in \mathbb{C}^{KG\times 1}$, $\mathbf{n}^t = [\mathbf{n}^t(1); ...; \mathbf{n}^t(G)] \in \mathbb{C}^{KG\times 1}$, $\mathbf{z}^t = \left[\mathbf{z}_{\mathrm{R},1}^t; ...; \mathbf{z}_{\mathrm{R},M}^t;\ \mathbf{z}_{\mathrm{B},1}^t; ...; \mathbf{z}_{\mathrm{B},M}^t\right] \in \mathbb{C}^{2UM\times 1}$, and $\mathbf{v}^t = \left[\mathbf{v}_{\mathrm{R},1}^t; ...; \mathbf{v}_{\mathrm{R},M}^t; \mathbf{v}_{\mathrm{B},1}^t; ...; \mathbf{v}_{\mathrm{B},M}^t\right] \in \mathbb{C}^{\left(\tilde{K}+\tilde{N}\right)M\times 1}$. The sensing matrices $\mathbf{F}^t\left(\Delta\mathbf{r}^t\right)$ and $\boldsymbol{\Xi}^t\left(\Delta\boldsymbol{\omega}^{\mathrm{AoA}}, \Delta\boldsymbol{\varphi}^{\mathrm{AoA}}, \Delta\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)$ are expressed as

$$\mathbf{F}^t\left(\Delta\mathbf{r}^t\right) = \left[\begin{array}{cc}\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{LoS}}(1) & \mathbf{A}_K^{\mathrm{LoS}}(1)\\ \vdots & \vdots\\ \mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{LoS}}(G) & \mathbf{A}_K^{\mathrm{LoS}}(G)\end{array}\right] \in \mathbb{C}^{KG\times 2UM}, \quad (10)$$

$$\begin{aligned}&\boldsymbol{\Xi}^t\left(\Delta\boldsymbol{\omega}^{\mathrm{AoA}}, \Delta\boldsymbol{\varphi}^{\mathrm{AoA}}, \Delta\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)\\ &= \left[\begin{array}{cc}\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{NLoS}}(1) & \mathbf{A}_K^{\mathrm{NLoS}}(1)\\ \vdots & \vdots\\ \mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{NLoS}}(G) & \mathbf{A}_K^{\mathrm{NLoS}}(G)\end{array}\right] \in \mathbb{C}^{KG\times\left(\tilde{K}+\tilde{N}\right)M}, \end{aligned} \quad (11)$$

where the entries $\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{LoS}}(g)$, $\mathbf{A}_K^{\mathrm{LoS}}(g)$, $\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{A}_N^{\mathrm{NLoS}}(g)$, and $\mathbf{A}_K^{\mathrm{NLoS}}(g)$ are shown in (12) at the top of the next page. Next, let us concentrate on the physical meaning of the two blocks, i.e., $\mathbf{z}^t$ and $\mathbf{v}^t$, in the sparse vector $\left[\begin{array}{cc}\mathbf{z}^t & \mathbf{v}^t\end{array}\right]^T$. As for $\mathbf{z}^t$, the vector $\mathbf{z}_{\mathrm{R},m}^t = \left[z_{\mathrm{R},m,1}^t, ..., z_{\mathrm{R},m,U}^t\right] \in \mathbb{C}^{U\times 1}$ represents the sparse LoS channel pertaining to the RIS-VUE link, the uth entry of which denote the complex gain of the LoS path associated with the off-grid location $\mathbf{r}_{m,u}^t + \Delta\mathbf{r}_{m,u}^t$ in the tth slot. The vector $\mathbf{z}_{\mathrm{B},m}^t \in \mathbb{C}^{U\times 1}$ represents the sparse LoS channel for the BS-VUE link. Note that each $\mathbf{z}_{\mathrm{R},m}^t$ and $\mathbf{z}_{\mathrm{B},m}^t$ only has one non-zero entry that points to the true location of the mth VUE, whose index of the non-zero entries of LoS channel vectors $\left[\mathbf{z}_{\mathrm{R},m}^t; \mathbf{z}_{\mathrm{B},m}^t\right], \forall m$, are the same, as given by q_m^t (which will be clear later). In other words, the LoS channel vectors $\left[\mathbf{z}_{\mathrm{R},m}^t; \mathbf{z}_{\mathrm{B},m}^t\right], \forall m$, can be grouped into U blocks with block size 2, with the uth block constituted by $\left[z_{\mathrm{R},m,u}^t, z_{\mathrm{B},m,u}^t\right], \forall m$, and only the uth block is non-zero. This implies that the sparse LoS channel vector $\mathbf{z}^t$ can be grouped into UM blocks with block size 2. By contrast, the sparse vectors $\mathbf{v}_{\mathrm{R},m}^t$ and $\mathbf{v}_{\mathrm{B},m}^t, \forall m$, contained in the block $\mathbf{v}^t$ denotes the complex gains of NLoS paths with respect to direct and cascaded channels, respectively. $\mathbf{v}_{\mathrm{R},m}^t = \left[\mathbf{v}_{\mathrm{R},m,1}^t, ..., \mathbf{v}_{\mathrm{R},m,\tilde{K}}^t\right] \in \mathbb{C}^{\tilde{K}\times 1}$ stands for the NLoS channel vector, the kth entry of which denotes the complex gain of the NLoS path arriving at the RIS with elevation-azimuth pair $\left(\tilde{\varphi}_{m,n}^{\mathrm{AoA}} + \Delta\varphi_{m,n}^{\mathrm{AoA}}, \tilde{\vartheta}_{m,n}^{\mathrm{AoA}} + \Delta\vartheta_{m,n}^{\mathrm{AoA}}\right)$. The NLoS channel vector $\mathbf{v}_{\mathrm{B},m}^t$ associated with the BS conform to the same definition given above. Note that there are $L_{m,r}$ and $L_{m,b}$ non-zero entries in $\mathbf{v}_{\mathrm{R},m}^t$ and $\mathbf{v}_{\mathrm{B},m}^t$, respectively, in which the indices of the non-zero entries indicate the significant AoAs associated with the NLoS paths. Therefore, the diverse sparsities inherent in (9) motivate us to conceive a customized tracking framework for the RIS-aided vehicle platoons.

D. Geometric Dilution of Precision (GDOP) Metric

GDOP has been commonly used as a metric for the accuracy evaluation of localization and tracking systems. Due to the fact that the high-accuracy localization requires both precise range measurement and a strong geometric association between the VUE and prescribed grids, the study of GDOP is a crucial aspect of assessing the performance of a localization system. Despite the fact that the Cramér-Rao lower bound (CRLB) metric may be applied to represent the minimum variance of the estimate error from any unbiased estimator, it fails to explicitly quantify the impact of system geometry on performance and also incorporates the effect of input noise [33]. Therefore, we elucidate that the integration of RIS into our considered vehicle platoon system has a beneficial GDOP performance by means of the concise GDOP analytical expressions for each VUE. We first present the GDOP in the instance of a ULA at the BS, and then extend it to the UPA case at the RIS, as these phases are pertinent to the discussion that follows.

1) ULA Case: Let $\mu_{\mathrm{B},k}$ denote the kth entry of $\mathbf{h}_{m,b}^t$ when there is no ambiguity, in which case the Fisher information matrix (FIM) associated with the AoA ω_m^{AoA} returns to a scalar

$$\mathrm{FIM}\left(\omega_m^{\mathrm{AoA}}\right) = \frac{2}{\sigma^2}\,\mathrm{Re}\left\{\sum_{k=1}^{K}\frac{\partial\mu_{\mathrm{B},k}^*}{\omega_m^{\mathrm{AoA}}}\frac{\partial\mu_{\mathrm{B},k}}{\omega_m^{\mathrm{AoA}}}\right\}, \quad (13)$$

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$$\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{LoS}}(g) = \left[x_1^t(g)\,\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{LoS}}\left(\Delta\mathbf{r}^t\right), ..., x_M^t(g)\,\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{LoS}}\left(\Delta\mathbf{r}^t\right)\right] \in \mathbb{C}^{K\times MU}, \tag{12a}$$

$$\mathbf{A}_K^{\text{LoS}}(g) = \left[x_1^t(g)\,\mathbf{A}_K^{\text{LoS}}\left(\Delta\mathbf{r}^t\right), ..., x_M^t(g)\,\mathbf{A}_K^{\text{LoS}}\left(\Delta\mathbf{r}^t\right)\right] \in \mathbb{C}^{K\times MU}, \tag{12b}$$

$$\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{NLoS}}(g) = \Big[x_1^t(g)\,\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{NLoS}}\left(\Delta\boldsymbol{\varphi}_1^{\text{AoA}}, \Delta\boldsymbol{\vartheta}_1^{\text{AoA}}\right), ..., x_M^t(g)\,\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{A}_N^{\text{NLoS}}\left(\Delta\boldsymbol{\varphi}_M^{\text{AoA}}, \Delta\boldsymbol{\vartheta}_M^{\text{AoA}}\right)\Big] \in \mathbb{C}^{K\times M\tilde{N}}, \tag{12c}$$

$$\mathbf{A}_K^{\text{NLoS}}(g) = \left[x_1^t(g)\,\mathbf{A}_K^{\text{NLoS}}\left(\Delta\boldsymbol{\omega}_1^{\text{AoA}}\right), ...,\ x_M^t(g)\,\mathbf{A}_K^{\text{NLoS}}\left(\Delta\boldsymbol{\omega}_{K\times\tilde{K}}^{\text{AoA}}\right)\right] \in \mathbb{C}^{K\times M\tilde{K}}. \tag{12d}$$

with $\frac{\partial\mu_{\text{B},k}}{\omega_m^{\text{AoA}}} = \eta_m\pi(k-1)\cos\omega_m^{\text{AoA}}\exp\left\{-\jmath\pi(k-1)\sin\omega_m^{\text{AoA}} - \jmath\frac{\pi}{2}\right\}$ and $\frac{\partial\mu_{\text{B},k}^*}{\omega_m^{\text{AoA}}} = \eta_m^*\pi(k-1)\cos\omega_m^{\text{AoA}}\exp\{-\jmath\pi(k-1)\times\sin\omega_m^{\text{AoA}} + \jmath\frac{\pi}{2}\}$. We can further structure (13) in the form of

$$\begin{aligned}\text{FIM}\left(\omega_m^{\text{AoA}}\right) &= \frac{2}{\sigma^2}\left(\pi\cos\omega_m^{\text{AoA}}\right)^2|\eta_m|^2\sum_{k=1}^{K}(k-1)^2\\ &\propto \frac{K^3}{3}\frac{2}{\sigma^2}\left(\pi\cos\omega_m^{\text{AoA}}\right)^2|\eta_m|^2.\end{aligned} \tag{14}$$

Thus, the CRLB associated with the AoA ω_m^{AoA} at the BS can be given by

$$\begin{aligned}\text{CRLB}\left(\omega_m^{\text{AoA}}\right) &= \text{FIM}^{-1}\left(\omega_m^{\text{AoA}}\right)\\ &= \frac{3\sigma^2}{2K^3(\pi\cos\omega_m^{\text{AoA}})^2|\eta_m|^2}.\end{aligned} \tag{15}$$

Furthermore, the GDOP remapped from the ω_m^{AoA}-based CRLB at the BS can be expressed as

$$\begin{aligned}\text{GDOP}_{\text{ULA}} &= \left(\cos\omega_m^{\text{AoA}}\right)^{-1}d_{m,b}^{-\zeta_{m,b}}\sqrt{\text{CRLB}\left(\omega_m^{\text{AoA}}\right)}\\ &= \sqrt{1.5}\sigma K^{-1.5}\eta_0^{-1}\pi^{-1}d_{m,b}^{1+\zeta_{m,b}}\left(\cos\omega_m^{\text{AoA}}\right)^{-1}.\end{aligned} \tag{16}$$

Without loss of generality, $\eta_0 = (c/4\pi f_c)^{\zeta_{m,b}}$ with c being the speed of light and f_c for the carrier frequency, and regarding the half-wavelength-space ULA, GDOP_{ULA} can be recast as

$$\begin{aligned}&\text{GDOP}_{\text{ULA}}\\ &= \sqrt{1.5}\pi^{-1}(c/4\pi)^{-\zeta_{m,b}}\sigma K^{-1.5}f_c^{\zeta_{m,b}}d_{m,b}^{1+\zeta_{m,b}}\left(\cos\omega_m^{\text{AoA}}\right)^{-1}\\ &\triangleq \text{GDOP}\left(\omega_m^{\text{AoA}}, d_{m,b}; \sigma, K, f_c, \zeta_{m,b}\right).\end{aligned} \tag{17}$$

As observed in (17), the GDOP is dependent on different factors, e.g., AoAs at the BS ω_m^{AoA}, the link distance $d_{m,b}$, the noise σ, the array size K, the carrier frequency f_c, and the path-loss exponent $\zeta_{m,b}$, etc. Next, we extend the aforementioned insights to the instance of a UPA at the RIS.

2) UPA Case: For an N-element UPA, we further denote N_h and N_v by the number of elements along the horizontal and vertical directions, respectively, i.e., $N = N_h \times N_v$, in which we can readily obtain the GDOP of the n_vth ULA with size N_h based on the argument in (17). In view of the independence among the observations achieved by each ULA, the observation of a UPA can be treated as a combination of N_v observations, thus yielding that

$$\begin{aligned}\text{GDOP}_{\text{UPA}} =& \sqrt{1.5}\pi^{-1}(c/4\pi)^{-\zeta_{m,r}}\sigma N_h^{-1.5}N_v^{-0.5}\\ &\times f_c^{\zeta_{m,r}}d_{m,r}^{1+\zeta_{m,r}}\left(\cos\varphi_m^{\text{AoA}}\cos\vartheta_m^{\text{AoA}}\right)^{-1}.\end{aligned} \tag{18}$$

3) GDOP Achieved by a BS Plus a RIS: In light of the facts in (17) and (18), the cascaded channel $\mathbf{H}_{r,b}\mathbf{\Theta}\mathbf{h}_{m,r}^t$-associated GDOP at the RIS can be given by

$$\begin{aligned}\text{GDOP}_{\text{RIS}} =& |\text{PL}_{r,b}|^{-1}K^{-0.5}N_v^{-0.5}\\ &\times \text{GDOP}\left(\varphi_m^{\text{AoA}}, \vartheta_m^{\text{AoA}}, d_{m,r}; \sigma, N_h, f_c, \zeta_{m,r}\right),\end{aligned} \tag{19}$$

where we use $\text{PL}_{r,b}$ here for representing the complex gain of the BS-RIS channel. One can refer to Appendix A for detailed derivation. Accordingly, we arrive at the closed-form GDOP expression with respect to the implementation of a BS plus a RIS

$$\text{GDOP}_{\text{BS\&RIS}} = \left(\text{GDOP}_{\text{BS}}^{-1} + \text{GDOP}_{\text{RIS}}^{-1}\right)^{-1}. \tag{20}$$

Given that GDOP_{RIS} is typically non-negative, it is evident that

$$\text{GDOP}_{\text{BS\&RIS}} \le \text{GDOP}_{\text{BS}}. \tag{21}$$

This implies that the application of RIS evidently improves localization precision owing to the richer localization information explored by RIS, i.e., additional LoS paths dominated by the significant angles, resulting in a reduction in error estimation (a lower GDOP).

Additionally, one may argue that the implementation of double BSs results in possibly better localization performance in comparison to the case of a BS plus a RIS. Despite its intuitive plausibility, achieving this typically necessitates a number of ideal conditions, such as the presence of strong direct links without any blockages, which may be impractical due to the complex nature of traffic patterns. Putting another BS in a single cell may introduce significant interference and additional complexity of the practical implementation since a BS is far more expensive than a RIS. From an economic standpoint, we would favor a more cost-effective solution if the two function comparably in the majority of circumstances. Then, to compare the localization performance of double BSs and a BS plus a RIS, we analyze their GDOPs in the following.

4) A BS Plus a RIS vs. Double BSs: The GDOP with respect to the implementation of double BSs can be explicitly written as

$$\text{GDOP} = \left(\text{GDOP}_{\text{BS}}^{-1} + \text{GDOP}_{\text{BS}}^{-1}\right)^{-1} = 0.5\text{GDOP}_{\text{BS}}. \tag{22}$$

Given the facts in (17) and (19), $\text{GDOP}_{\text{BS}}/\text{GDOP}_{\text{RIS}}$ can be given by (23) shown at the top of the next page. It is intuitive that the ratio $\text{GDOP}_{\text{BS}}/\text{GDOP}_{\text{RIS}}$ is subject to a number of parameters associated with the direct and cascaded channels, in particular for the array scale and the link distance. If we

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$$\frac{\mathrm{GDOP_{BS}}}{\mathrm{GDOP_{RIS}}}=\frac{\sqrt{1.5}\pi^{-1}(c/4\pi)^{-\zeta_{m,b}}\sigma K^{-1.5}f_c^{\zeta_{m,b}}d_{m,b}^{1+\zeta_{m,b}}\left(\cos\omega_m^{\mathrm{AoA}}\right)^{-1}}{|\mathrm{PL}_{r,b}|^{-1}K^{-0.5}N_v^{-0.5}\mathrm{GDOP}\left(\varphi_m^{\mathrm{AoA}},\vartheta_m^{\mathrm{AoA}},d_{m,r};\sigma,N_h,f_c,\zeta_{m,r}\right)}. \tag{23}$$

appropriately neglect some trivial parameters, (23) can be boldly trimmed to

$$\frac{\mathrm{GDOP_{BS}}}{\mathrm{GDOP_{RIS}}}\propto|\mathrm{PL}_{r,b}|\cdot K^{-1}N_h^{1.5}N_v^{0.5}. \tag{24}$$

It can be inferred that the localization performance achieved by a BS plus a RIS is superior to that of the double BSs in the event of the ratio $\mathrm{GDOP_{BS}}/\mathrm{GDOP_{RIS}}$ being greater than 1, i.e., $\frac{\mathrm{GDOP_{BS}}}{\mathrm{GDOP_{RIS}}}>1$. With a glance at (24), the ratio $\mathrm{GDOP_{BS}}/\mathrm{GDOP_{RIS}}$ is inversely proportional to K while proportionate to N_h and N_v. This connotes that the localization performance of RIS may be considerably enhanced by increasing its element number $N=N_h\times N_v$, since this may compensate for the multiplicative fading nature inherent to the cascaded channels [32]. Therefore, it is not wise to follow empiricism rather than observing their attributes with theoretical analyses. Under different simulation configurations, we compare localization performance attained by double BS and a BS plus RIS using GDOP metric, as will be elaborated in Sec. V. C.

III. Diverse Dynamic Layered Structured Sparsity (DiLuS) Framework

In regard to high-accuracy and robust VUE location tracking, we intend to recover both the sparse LoS and NLoS channel vectors, namely $\mathbf{z}^t$ and $\mathbf{v}^t$ as presented in (9), for each VUE of the considered vehicle platoon system from the observation $\mathbf{y}_m^t$. Although such a parameter-recovery problem seems straightforward, it poses a few challenges owing to a pair of peculiarities: i) the uncertainty of the sensing matrices containing the location and angle offsets to be determined, i.e., $\mathbf{F}_m^t$ and $\mathbf{\Xi}_m^t$, and ii) the unavailability of the exact distributions for the sparse vectors $\mathbf{z}^t$ and $\mathbf{v}^t$ with diverse sparsities (the same for the distribution of each entry within both vectors). Therefore, targeting the twofold issues, we customize a DiLuS framework as illustrated in Fig. 1, for completely capturing diverse sparsity structures. The proposed DiLuS, on the one hand, leverages spatial-temporal correlations associated with LoS/NLoS channels among the vehicle platoon, in a formation to facilitate the closed-form updates with reduced complexity and overhead. On the other hand, it can provide appropriate prior distributions for $\mathbf{z}^t$ and $\mathbf{v}^t$ in support of efficient and robust performance in location tracking across a range of channel distributions.

A. LoS Channel

Recall that $\mathbf{z}^t$ represents the sparse LoS channels incorporating the information of both direct and cascaded channels, and we let $\boldsymbol{\rho}^t=[\boldsymbol{\rho}_{\mathrm{R}}^t;\boldsymbol{\rho}_{\mathrm{B}}^t]\in\mathbb{C}^{2UM\times1}$ represent the precision vector of $\mathbf{z}^t$. In particular, $\boldsymbol{\rho}_{\mathrm{R}}^t=\left[\boldsymbol{\rho}_{\mathrm{R},1}^t;...;\boldsymbol{\rho}_{\mathrm{R},M}^t\right]\in\mathbb{C}^{UM\times1}$, $\boldsymbol{\rho}_{\mathrm{B}}^t=\left[\boldsymbol{\rho}_{\mathrm{B},1}^t;...;\boldsymbol{\rho}_{\mathrm{B},M}^t\right]\in\mathbb{C}^{UM\times1}$, $\boldsymbol{\rho}_{\mathrm{R},m}^t=\left[\rho_{\mathrm{R},m,1}^t,...,\rho_{\mathrm{R},m,U}^t\right]^T\in\mathbb{C}^{U\times1}$, and $\boldsymbol{\rho}_{\mathrm{B},m}^t=\left[\rho_{\mathrm{B},m,1}^t,...,\rho_{\mathrm{B},m,U}^t\right]^T\in\mathbb{C}^{U\times1}$, where $\rho_{\mathrm{R},m,u}^t$ indicates the precision of the entry $z_{\mathrm{R},m,u}^t$ involved in the uth block with $1/\rho_{\mathrm{R},m,u}^t$ being the variance of $z_{\mathrm{R},m,u}^t$, and $\rho_{\mathrm{B},m,u}^t$ for $z_{\mathrm{B},m,u}^t$. For ease of exposition, we denote a sequence of time-related variables $\{\mathbf{z}^\tau\}_{\tau=1}^t$ by $\mathbf{z}^{1:t}$, the shorthand of which is shared with others.

1) Dynamic Hidden Markov Model for the Vehicle Platoon: Given that q indicates the index of the off-grid location nearest to the uth prescribed grid for the mth VUE, we refine $q_m^t\in\{1,...,U\}$ as the off-grid location state in order to represent the coarse position of the mth VUE along the prescribed U location grids in the tth slot, with $\boldsymbol{q}^t=[q_1^t,...,q_M^t]^T,\forall t$. Taking into account the spatial-temporal attributes of vehicle platoons, a dynamic hidden Markov model is used to generate priors, i.e.,

$$\begin{aligned}p\left(\boldsymbol{q}^{1:t}\right)&=\prod_\tau p\left(\boldsymbol{q}^\tau|\boldsymbol{q}^{\tau-1}\right)\\&=\prod_\tau\prod_m\underbrace{p\left(q_{m+1}^\tau|q_m^\tau\right)}_{\triangleq h_{S,m}^\tau(\text{Spatial})}\underbrace{p\left(q_m^\tau|q_m^{\tau-1}\right)}_{\triangleq h_{T,m}^\tau(\text{Temporal})}.\end{aligned} \tag{25}$$

Particularly, we suppose that the spatial correlation between the mth and $(m+1)$th VUE, i.e., $p\left(q_{m+1}^\tau|q_m^\tau\right)$, is drawn from a sampled Gamma distribution [34]

$$p\left(q_{m+1}^\tau|q_m^\tau\right)\propto\frac{\left(q_m^\tau-q_0\right)^{\varpi-1}}{\lambda^\varpi\Gamma\left(\varpi\right)}\exp\left(-\frac{q_m^\tau-q_0}{\lambda}\right),q_m^\tau\geq q_0, \tag{26}$$

where q_0 denotes the minimum distance between vehicles, $\varpi\geq1$ is the common shape parameter, and λ is the scale parameter. Exploiting the collected data (received signals) and the EM technique, parameters ϖ and λ can be calculated in practice, which is a widely held belief in the academic literature confirmed by empirical research [11], [34]. Furthermore, the VUE's location in the previous slot provides a priori indicator of its location in the present slot, which may be achieved by message passing and will be expounded upon later in Sec. IV. B.

2) Layers II & III: A Pair of Conjugate Distributions: A Gamma prior distribution is imposed on the precision vector $\boldsymbol{\rho}^t$ of the off-grid location state $\boldsymbol{q}^t$

$$p\left(\boldsymbol{\rho}^t|\boldsymbol{q}^t\right)=\prod_{m,u}p\left(\rho_{\mathrm{R},m,u}^t|q_m^t\right)p\left(\rho_{\mathrm{B},m,u}^t|q_m^t\right), \tag{27}$$

where $p\left(\rho_{i,m,u}^t|q_m^t\right)=\prod_{m,u}\Gamma\big(\rho_{i,m,u}^t;\mathfrak{a}_{i,m}^t,\mathfrak{b}_{i,m}^t\big)^{\mathbb{I}(q=u)}\times\Gamma\big(\rho_{i,m,u}^t;\bar{\mathfrak{a}}_{i,m}^t,\bar{\mathfrak{b}}_{i,m}^t\big)^{\mathbb{I}(q\neq u)},i\in\{\mathrm{R},\mathrm{B}\}$, and $\Gamma\left(\rho;\mathfrak{a},\mathfrak{b}\right)$ is a Gamma hyperprior with the shape parameter $\mathfrak{a}$ and rate parameter $\mathfrak{b}$. For the mth VUE, $\mathfrak{a}_{i,m}$ and $\mathfrak{b}_{i,m}$ are the shape and rate parameters of the LoS channel precision $\rho_{i,m,u}^t$, conditioned on the fact that the uth off-grid location corresponds to the mth VUE's location, i.e., $q=u$. On the contrary, the parameter and rate parameters $\bar{\mathfrak{a}}_{i,m}$ and $\bar{\mathfrak{b}}_{i,m}$ should be delicately chosen on the event of $q\neq u$.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

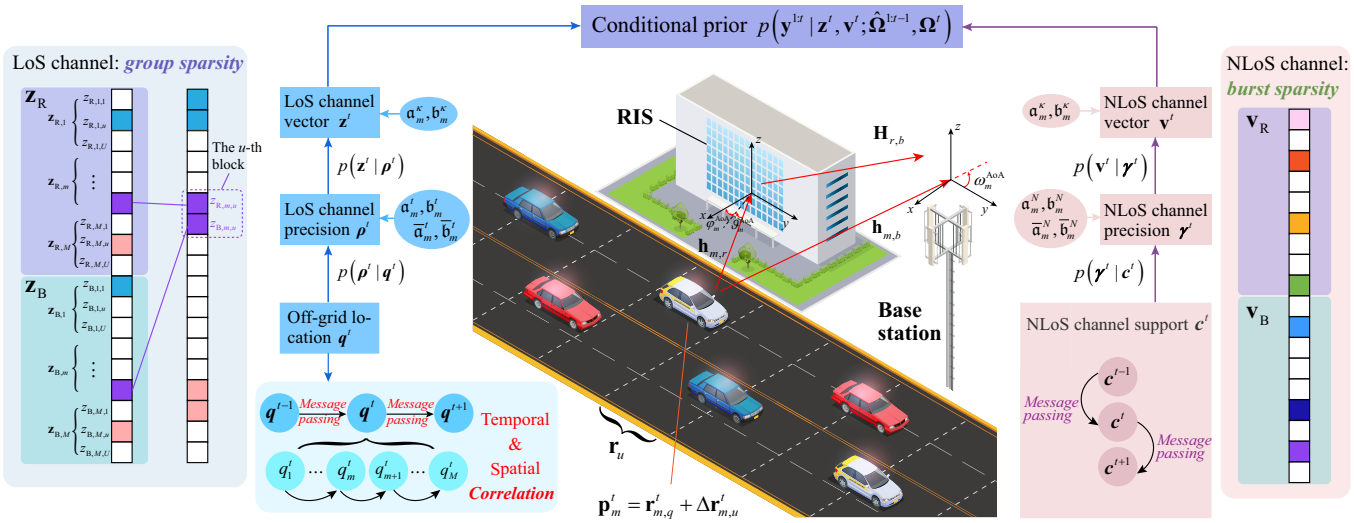


Fig. 1. The proposed DiLuS framework.

Regarding the LoS channel vector $\mathbf{z}^t$, each entry thereof is considered to have a non-stationary Gaussian prior distribution conditioned on a specified precision $\rho^t_{m,u}$, i.e.,

$$p\left(\mathbf{z}^t|\boldsymbol{\rho}^t\right)=\prod_{m,u}p\left(z^t_{\mathrm{R},m,u}|\rho^t_{\mathrm{R},m,u}\right)p\left(z^t_{\mathrm{B},m,u}|\rho^t_{\mathrm{B},m,u}\right), \quad (28)$$

in which $p\left(z^t_{\mathrm{R},m,u}|\rho^t_{\mathrm{R},m,u}\right)=\prod_{m,u}\mathcal{CN}\left(z^t_{\mathrm{R},m,u};0,\left(1/\rho^t_{\mathrm{R},m,u}\right)\right)$ and $p\left(z^t_{\mathrm{B},m,u}|\rho^t_{\mathrm{B},m,u}\right)=\prod_{m,u}\mathcal{CN}\left(z^t_{\mathrm{B},m,u};0,\left(1/\rho^t_{\mathrm{B},m,u}\right)\right)$.

3) Joint Distribution of the Layered Group Sparsity for LoS Channel: As regards the LoS channel, the joint distribution of $\boldsymbol{q}^{1:t}$, $\boldsymbol{\rho}^{1:t}$, and $\mathbf{z}^{1:t}$ is given by

$$p\left(\boldsymbol{q}^{1:t},\mathbf{z}^{1:t},\boldsymbol{\rho}^{1:t}\right)=\prod_{\tau}\underbrace{p\left(\boldsymbol{q}^{\tau}\right)}_{\text{Off-grid location}}\underbrace{p\left(\boldsymbol{\rho}^{\tau}|\boldsymbol{q}^{\tau}\right)}_{\text{Precision}}\underbrace{p\left(\mathbf{z}^{\tau}|\boldsymbol{\rho}^{\tau}\right)}_{\text{LoS channel}}. \quad (29)$$

Note that the LoS channel vector $\mathbf{z}^t$ can be segmented into UM blocks (U blocks for each VUE m) with the uth block carrying the channel energies of both direct and cascaded channels, as both energies are all concentrated on the same location index q^t_m. More explicitly, the conditional prior in (27) imparts fewer precisions to the blocks fulfilling $q=u$, forcing the entries thereof to depart from zero. For the opposite event of the blocks satisfying $q\neq u$, more precisions are assigned by the conditional prior in (27) to focus the block entries towards zero. Thus, the LoS channel vector contains separate blocks having distinct precisions for capturing different path gains, with the non-zero entry indexing the exact grid location of the mth VUE, termed as the *group sparsity*.

B. NLoS Channel

We also scrutinize NLoS channel modeling at the BS/RIS using a layered heterogeneous paradigm. For the sake of clarity, we use the RIS-VUE link as an instance employing only a set of self-contained variables to construct hierarchical priors for the exact channel distributions[3]. Thus in this instance, we just use $\mathbf{v}^t$ in place of $\mathbf{v}^t_{\mathrm{R}}$ or $\mathbf{v}^t_{\mathrm{B}}$ for notational brevity. Denote $\boldsymbol{c}^t=\left[\boldsymbol{c}^t_1;...;\boldsymbol{c}^t_M\right]^T\in\mathbb{C}^{\tilde{N}M\times 1},\forall t$ by the NLoS channel support, in which $\boldsymbol{c}^t_m=\left[\boldsymbol{c}^t_{m,1},...,\boldsymbol{c}^t_{m,\tilde{N}}\right]^T\in\mathbb{C}^{\tilde{N}\times 1},\forall m,t$ with $\boldsymbol{c}^t_{m,n}\in\{0,1\}$ indicating whether there is an active NLoS path arriving at the BS from the nth angular grid at the RIS. Then, let $\boldsymbol{\gamma}^t=\left[\boldsymbol{\gamma}^t_1;...;\boldsymbol{\gamma}^t_M\right]^T\in\mathbb{C}^{\tilde{N}M\times 1},\forall t$ denote the precise vector of the NLoS channel support with $\boldsymbol{\gamma}^t_m=\left[\gamma^t_{m,1},...,\gamma^t_{m,\tilde{N}}\right]^T\in\mathbb{C}^{\tilde{N}\times 1},\forall m,t$, for which $\gamma^t_{m,n}$ represents the precision of $v^t_{m,n}$, i.e., $1/\gamma^t_{m,n}$ is the variance of $v^t_{m,n}$.

1) Layer I: Markov Model for NLoS Channel Support: In a time-varying environment, the AoAs at the BS/RIS may change slowly on the timescale of slots, and thus we employ the following model to capture the temporal correlation of the non-zero entries in the NLoS channel support $\boldsymbol{c}^t$

$$c^t_{m,n}=\tilde{\varrho}_{m,n}c^{t-1}_{m,n}+\sqrt{1-\tilde{\varrho}^2_{m,n}}\mathcal{CN}\left(\alpha^t_{m,n};0,\tilde{\sigma}^2_{m,n}\right), \quad (30)$$

where $\tilde{\varrho}_{m,k,n,q}\in(0,1)$ characterizes the temporal correlation between two consecutive slots, termed as the temporal correlation coefficient, and $\mathcal{CN}\left(\alpha^t_{m,n};0,\tilde{\sigma}^2_{m,n}\right)$ is the added Gaussian perturbation with zero mean and variance $\tilde{\sigma}^2_{m,n}$. Hence, the NLoS channel support $\boldsymbol{c}^t$ takes the form of a conditional probability as follows

$$\begin{aligned}p\left(\boldsymbol{c}^{1:t}\right)&=p\left(\boldsymbol{c}^1\right)\prod_{\tau=2}^{t}p\left(\boldsymbol{c}^{\tau}|\boldsymbol{c}^{\tau-1}\right)\\&=\prod_{m,n}p\left(c^1_{m,n}\right)\prod_{\tau=2}^{t}p\left(c^{\tau}_{m,n}|c^{\tau-1}_{m,n}\right).\end{aligned} \quad (31)$$

Concerning the smooth dynamic behaviors inherent in the time-varying channels, we are allowed to identify the follow-

[3]Actually, there is still NLoS channel support over the BS-VUE link. The hierarchical priors of the BS-VUE NLoS channel continues to share with those presented in the RIS-VUE NLoS channel.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

ing independent first-order auto-regressive processes for the temporal correlation of the non-zero entries in $\boldsymbol{c}^t$ [35]

$$\begin{aligned} p\left(c_{m,n}^{\tau}|c_{m,n}^{\tau-1}\right) &= \prod_{m,n} p\left(c_{m,n}^{\tau}|c_{m,n}^{\tau-1};\tilde{\varrho}_{m,n},\tilde{\sigma}_{m,n}\right) \\ &= \prod_{m,n} \mathcal{CN}\left(c_{m,n}^{\tau};\tilde{\varrho}_{m,n}\alpha_{m,n}^{t-1},\left(1-\tilde{\varrho}_{m,n}^2\right)\tilde{\sigma}_{m,n}^2\right), \end{aligned} \tag{32}$$

where $p\left(c_{m,n}^1\right) \triangleq p\left(c_{m,n}^1|c_{m,n}^0\right) = \mathcal{CN}\left(c_{m,n}^1;0,\tilde{\sigma}_{m,n}^2\right)$ specifies a steady-state distribution. To this end, the Markov model for the NLOS channel supports can be applicable for encapsulating various channel realizations in practice by subtly tuning parameters thereof. Hence, our analysis takes into account the steadily changing propagation environment between VUEs and BS/RIS, in which case the temporal coefficient $\tilde{\varrho}_{m,n}$ remains (almost) static during the tracking process of interest.

2) Layers II & III: A Pair of Conjugate Distributions: The conditional prior of the precision vector $\boldsymbol{\gamma}^t$ associated with the NLoS channel vector $\mathbf{v}^t$ is given by

$$\begin{aligned} p\left(\boldsymbol{\gamma}^t|\boldsymbol{c}^t\right) &= \prod_{m,n} p\left(\gamma_{m,n}^t|c_{m,n}^t\right) \\ &= \prod_{m,n} \Gamma\left(\gamma_{m,n}^t;\mathfrak{a}_m^N,\mathfrak{b}_m^N\right)^{\mathbb{I}\left(c_{m,n}^t\right)} \Gamma\left(\gamma_{m,n}^t;\bar{\mathfrak{a}}_m^N,\bar{\mathfrak{b}}_m^N\right)^{\mathbb{I}\left(1-c_{m,n}^t\right)}. \end{aligned} \tag{33}$$

The case of $c_{m,n}^t = 1$ indicates that there is an active NLoS path between the RIS and the mth VUE, in which the shape and rate parameters $\mathfrak{a}_m^N$ and $\mathfrak{b}_m^N$ are required to satisfy $\mathfrak{a}_m^N/\mathfrak{b}_m^N = \mathbb{E}\left[\gamma_{m,n}^t\right], \forall m,n$ due to the fact that the variance $1/\gamma_{m,n}^t$ of $z_{m,u}^t$ equals to $\mathbb{E}\left[\gamma_{m,n}^t\right]$ for the active NLoS path. With respect to the opposite case of $c_{m,n}^t = 0$, it holds that $\bar{\mathfrak{a}}_m^N/\bar{\mathfrak{b}}_m^N = \mathbb{E}\left[\gamma_{m,n}^t\right] \gg 1, \forall m,n$ owing to the fact that the variance $1/\gamma_{m,n}^t$ approaches zero for the associated inactive NLoS path.

As regards NLoS channels, a non-stationary Gaussian prior distribution with varying precision $\gamma_{m,n}^t$ is assigned to each entry of the NLoS channel $\mathbf{v}^t$ as follows

$$\begin{aligned} p\left(\mathbf{v}^t|\boldsymbol{\gamma}^t\right) &= \prod_{m,n} p\left(v_{m,n}^t|\gamma_{m,n}^t\right) \\ &= \prod_{m,n} \mathcal{CN}\left(v_{m,n}^t;0,\left(1/\gamma_{m,n}^t\right)\right). \end{aligned} \tag{34}$$

3) Joint Distribution of the Layered Group Sparsity for NLoS Channel: Accordingly, the joint distribution of $\boldsymbol{c}^t$, $\boldsymbol{\gamma}$, and $\mathbf{v}$ is given by

$$p\left(\boldsymbol{c}^{1:t},\boldsymbol{\gamma}^{1:t},\mathbf{v}^{1:t}\right) = \prod_{\tau} \underbrace{p\left(\boldsymbol{c}^{\tau}\right)}_{\text{NLoS channel support}} \underbrace{p\left(\boldsymbol{\gamma}^{\tau}|\boldsymbol{c}^{\tau}\right)}_{\text{Precision}} \underbrace{p\left(\mathbf{v}^{\tau}|\boldsymbol{\gamma}^{\tau}\right)}_{\text{NLoS channel}}. \tag{35}$$

Additionally, recall that the noise in (3) follows a complex Gaussian distribution, and we denote $\kappa^t = \sigma^{-2}$ by the noise precision. The noise precision κ^t is typically unavailable and can be modeled using a Gamma distribution as a hyperprior $p\left(\kappa^t\right) = \Gamma\left(\kappa^t;\mathfrak{a}^{\kappa},\mathfrak{b}^{\kappa}\right)$, where the hyper-parameters $\mathfrak{a}^{\kappa}$ and $\mathfrak{b}^{\kappa}$ can be tuned towards near zero, i.e., $\mathfrak{a}^{\kappa},\mathfrak{b}^{\kappa} \to 0$, for an approximation of a more general hyperprior [36].

Significant angles typically appear in bursts (clusters), with each burst coinciding to a scattering cluster in the propagation environment, as contrasted to the group sparsity associated with the LoS channel vector $\mathbf{z}^t$ previously demonstrated. This is referred to as the *burst sparsity*.

C. Problem Formulation

Our goal is to estimate the off-grid location offsets and AoA offset vectors associated with NLoS path, i.e., $\hat{\boldsymbol{\Omega}}^t = \left\{\Delta\hat{\mathbf{r}}^t,\left(\Delta\hat{\boldsymbol{\omega}}^{\text{AoA}}\right)^t,\left(\Delta\hat{\boldsymbol{\varphi}}^{\text{AoA}}\right)^t,\left(\Delta\hat{\boldsymbol{\vartheta}}^{\text{AoA}}\right)^t\right\}$, given the observations up to the tth slot $\mathbf{y}^{1:t} = \left[\mathbf{y}_1^{1:t};...;\mathbf{y}_M^{1:t}\right] \in \mathbb{C}^{KM\times 1}$ in (3) and $\hat{\boldsymbol{\Omega}}^{t-1}$ up to the $(t-1)$th slot. To be specific, the offset parameters $\boldsymbol{\Omega}^t$ can be obtained by a maximization of the likelihood function as follows

$$\begin{aligned} \hat{\boldsymbol{\Omega}}^t &= \arg\max_{\boldsymbol{\Omega}^t} \ln p\left(\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right) \\ &= \arg\max_{\boldsymbol{\Omega}^t} \int_{\boldsymbol{\alpha}^t} \ln p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right) d\boldsymbol{\alpha}^t, \end{aligned} \tag{36}$$

in which we denote $\boldsymbol{\alpha}^t = \{\mathbf{z}^t,\mathbf{v}^t,\boldsymbol{\rho}^t,\gamma^t,\boldsymbol{q}^t,\boldsymbol{c}^t\}$ by the collection of corresponding variables for brevity. We attempt to determine the off-grid location index q_m^t by calculating marginal posteriors $p\left(q_m^t|\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right)$ for the given estimates of offset parameters $\boldsymbol{\Omega}^t$, after which the estimate $\hat{q}_m^t$ of q_m^t might be provided by the MAP probability estimate

$$\hat{q}_m^t = \arg\max_{q_m^t\in\{1,\ldots,U\}} p\left(q_m^t|\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right), \tag{37}$$

with the conditional marginal posterior $p\left(q_m^t|\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right)$ given by

$$\begin{aligned} &p\left(q_m^t|\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right) \\ &\propto \sum_{\boldsymbol{c}^t} \int_{-q_m^t,\mathbf{z}^t,\mathbf{v}^t,\boldsymbol{\rho}^t,\gamma^t} p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right) \\ &= \sum_{\boldsymbol{c}^t} \int_{-q_m^t,\mathbf{z}^t,\mathbf{v}^t,\boldsymbol{\rho}^t,\gamma^t} p\left(\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right) p\left(\mathbf{y}^{1:t}|\boldsymbol{\alpha}^t\right), \end{aligned} \tag{38}$$

where $-q_m^t$ denotes the vector collection over $\boldsymbol{q}^t$ except for the entry q_m^t, and $\propto$ refers to the equality after scaling. Consequently, the location estimate of the mth VUE in the tth slot can be determined by $\hat{\mathbf{p}}_m^t = \mathbf{r}_{m,\hat{q}}^t + \Delta\hat{\mathbf{r}}_{m,\hat{q}}^t$.

Actually, it is quite challenging to attain the precise estimate of $\hat{\mathbf{p}}_m^t, \forall m$, due to the prohibitive integration in (38). More specifically, problem (38) exposes the time-sequence attribute due to the spatial-temporal correlated off-grid location $\boldsymbol{q}^t$, which implies that the determination of $\boldsymbol{q}^t$ in each slot relies upon the observations $\mathbf{y}^{1:t}$ up to the current slot t. As t increases, the location tracking may become formidable owing to the excessively high computing complexity and memory costs. Furthermore, the complex priors of the RIS/BS-associated latent variables (such as $\boldsymbol{c}^t$, $\boldsymbol{q}^t$, $\boldsymbol{\rho}^t$, $\boldsymbol{\gamma}^t$) and their intricate couplings (correlations) impart intractability to the exact conditional marginal posterior in (38). This also induces the non-convexity in (36) and renders some general optimization techniques fruitless for achieving a stationary solution. Therefore, the section that follows goes on to present beneficial techniques for clearing up the above obstacles, as well as an effective algorithm for the VUE location tracking with high accuracy and great robustness.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

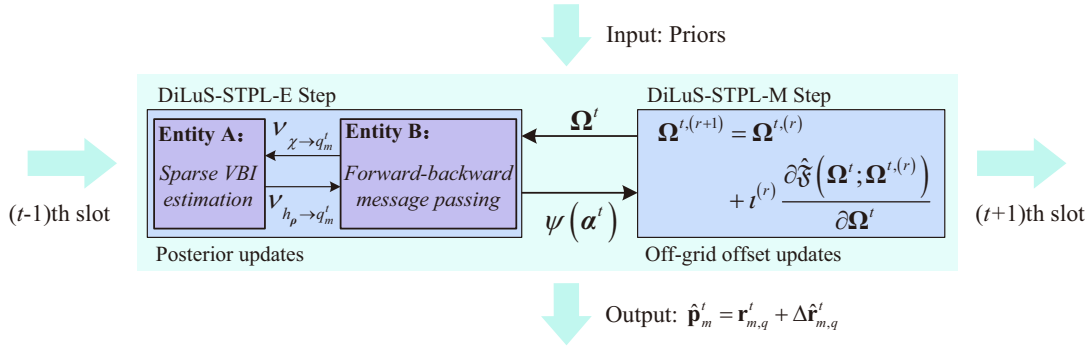


Fig. 2. The proposed DiLuS-STPL algorithm.

IV. DiLuS-STPL Algorithm

The fundamental goal of our examined location tracking is the robust and accurate recovery of the sparse LoS/NLoS channel vectors, e.g., $\mathbf{z}^t$ and $\mathbf{v}^t$, and the offset parameters $\boldsymbol{\Omega}^t$ based on the received signals $\mathbf{y}^{1:t}$ up to the tth slot. Despite this, the computing complexity and memory costs involved should be taken into account if running an algorithm that performs tracking in each slot relying on the observations up to its current slot. To tackle this critical issue, this section aims to propose DiLuS-STPL algorithm in support of the intra-slot VBI estimates of $\mathbf{z}^t$, $\mathbf{v}^t$, and $\boldsymbol{\Omega}^t$, while incorporating the inter-slot message passing at a reduced complexity and overhead.

A. Temporal Decomposition and Problem Approximation

The joint probability distribution in (36) can be decomposed into two lightweight parts. In particular, one part covers the joint distribution in the tth slot which can be harnessed to determine $\mathbf{z}^t$, $\mathbf{v}^t$, and $\boldsymbol{\Omega}^t$. The other part, actually being the message passed from the previous slot, incorporates the information delivered by previous observations $\mathbf{y}^{1:t-1}$. Firstly, we restructure the joint distribution in the tth slot $p\left(\mathbf{y}^t, \boldsymbol{\alpha}^t; \boldsymbol{\Omega}^t\right)$ as

$$\begin{aligned} p\left(\mathbf{y}^t, \boldsymbol{\alpha}^t; \boldsymbol{\Omega}^t\right) =& p\left(\mathbf{y}^t, \mathbf{z}^t, \mathbf{v}^t, \boldsymbol{\rho}^t, \boldsymbol{\gamma}^t, \boldsymbol{q}^t, \boldsymbol{c}^t\right) \\ \overset{(a)}{=}& p\left(\mathbf{y}^t|\mathbf{z}^t, \mathbf{v}^t; \boldsymbol{\Omega}^t\right) p\left(\mathbf{z}^t|\boldsymbol{\rho}^t\right) p\left(\boldsymbol{\rho}^t|\boldsymbol{q}^t\right) \\ & \times p\left(\mathbf{v}^t|\boldsymbol{\gamma}^t\right) p\left(\boldsymbol{\gamma}^t|\boldsymbol{c}^t\right) p\left(\boldsymbol{q}^t\right) p\left(\boldsymbol{c}^t\right), \end{aligned} \tag{39}$$

in which (a) holds due to the proposed DiLuS framework. By taking into account the priors from the previous slot, the joint distribution in (36) can be recast leveraging the sequential Bayesian inference technique [37], [38], i.e.,

$$\begin{aligned} & p\left(\mathbf{y}^{1:t}, \boldsymbol{\alpha}^t; \hat{\boldsymbol{\Omega}}^{1:t-1}, \boldsymbol{\Omega}^t\right) \\ & \overset{(b)}{\propto} p\left(\mathbf{y}^t, \boldsymbol{\alpha}^t; \boldsymbol{\Omega}^t\right) p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right), \end{aligned} \tag{40}$$

where (b) holds due to [37, Eq. (3.2)]. The conditional posterior $p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)$ is given by

$$\begin{aligned} & p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right) \\ & \overset{(c)}{=} \int \underbrace{p\left(\boldsymbol{\alpha}^t|\boldsymbol{\alpha}^{t-1}\right)}_{\text{Transition}} \underbrace{p\left(\boldsymbol{\alpha}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)}_{\text{Prior}} d\boldsymbol{\alpha}^{t-1}, \end{aligned} \tag{41}$$

where (c) holds due to [38, Eq. (3.12)] with $p\left(\boldsymbol{\alpha}^t|\boldsymbol{\alpha}^{t-1}\right)$ given by

$$p\left(\boldsymbol{\alpha}^t|\boldsymbol{\alpha}^{t-1}\right) = p\left(\mathbf{z}^t, \mathbf{v}^t, \boldsymbol{\rho}^t, \boldsymbol{\gamma}^t|\boldsymbol{q}^t, \boldsymbol{c}^t\right) p\left(\boldsymbol{q}^t|\boldsymbol{q}^{t-1}\right) p\left(\boldsymbol{c}^t|\boldsymbol{c}^{t-1}\right). \tag{42}$$

Regarding the conditional posterior $p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)$, we have

$$\begin{aligned} & p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right) \\ & \overset{(d)}{=} p\left(\mathbf{z}^t, \mathbf{v}^t, \boldsymbol{\rho}^t, \gamma^t|\boldsymbol{q}^t, \boldsymbol{c}^t\right) \sum_{\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}} p\left(\boldsymbol{q}^t|\boldsymbol{q}^{t-1}\right) p\left(\boldsymbol{c}^t|\boldsymbol{c}^{t-1}\right) \\ & \quad \times p\left(\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right), \end{aligned} \tag{43}$$

for which (d) holds due to [38, Eq. (3.12)]. Hence, $p\left(\mathbf{y}^{1:t}, \boldsymbol{\alpha}^t; \hat{\boldsymbol{\Omega}}^{1:t-1}, \boldsymbol{\Omega}^t\right)$ in (40) can be further structured as

$$\begin{aligned} & p\left(\mathbf{y}^{1:t}, \boldsymbol{\alpha}^t; \hat{\boldsymbol{\Omega}}^{1:t-1}, \boldsymbol{\Omega}^t\right) \\ & \propto p\left(\mathbf{y}^t, \boldsymbol{\alpha}^t; \boldsymbol{\Omega}^t\right) p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right) \\ & = p\left(\mathbf{y}^t|\mathbf{z}^t, \mathbf{v}^t; \boldsymbol{\Omega}^t\right) p\left(\mathbf{z}^t|\boldsymbol{\rho}^t\right) p\left(\boldsymbol{\rho}^t|\boldsymbol{q}^t\right) p\left(\mathbf{v}^t|\boldsymbol{\gamma}^t\right) p\left(\boldsymbol{\gamma}^t|\boldsymbol{c}^t\right) \\ & \quad \times \sum_{\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}} p\left(\boldsymbol{q}^t|\boldsymbol{q}^{t-1}\right) p\left(\boldsymbol{c}^t|\boldsymbol{c}^{t-1}\right) \\ & \quad \times p\left(\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right). \end{aligned} \tag{44}$$

Owing to the intractability of the exact posterior $p\left(\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)$, one may need a safe approximation in its place, i.e., $p\left(\boldsymbol{q}^{t-1}, \boldsymbol{c}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right) \approx \psi\left(\boldsymbol{q}^{t-1}\right) \psi\left(\boldsymbol{c}^{t-1}\right)$ with $\psi\left(\boldsymbol{q}^{t-1}\right) \approx p\left(\boldsymbol{q}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)$ and $\psi\left(\boldsymbol{c}^{t-1}\right) \approx p\left(\boldsymbol{c}^{t-1}|\mathbf{y}^{1:t-1}; \hat{\boldsymbol{\Omega}}^{1:t-1}\right)$. In this case, the joint prior distribution in (44) can be further recast as (45) shown at the top of the next page, where the approximate priors of $\boldsymbol{q}^t$ and $\boldsymbol{c}^t$ are particularly given by $\hat{p}\left(\boldsymbol{q}^t\right) \triangleq \sum_{\boldsymbol{q}^{t-1}} p\left(\boldsymbol{q}^t|\boldsymbol{q}^{t-1}\right) \psi\left(\boldsymbol{q}^{t-1}\right)$ and $\hat{p}\left(\boldsymbol{c}^t\right) \triangleq \sum_{\boldsymbol{c}^{t-1}} p\left(\boldsymbol{c}^t|\boldsymbol{c}^{t-1}\right) \psi\left(\boldsymbol{c}^{t-1}\right)$. Note that the only discrepancy between the prior in (45) and the original prior lies in that the dynamic hidden Markov model $p\left(\boldsymbol{q}^t\right)$ associated with the off-grid location is replaced by an approximate prior $\hat{p}\left(\boldsymbol{q}^t\right)$ with independent entries, while the Markov model $p\left(\boldsymbol{c}^t\right)$ is approximated by its safe copy

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$$
\begin{aligned}
&p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right)\approx\hat{p}\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1},\boldsymbol{\Omega}^t\right)\\
&=p\left(\mathbf{y}^t|\mathbf{z}^t,\mathbf{v}^t;\boldsymbol{\Omega}^t\right)p\left(\mathbf{z}^t|\boldsymbol{\rho}^t\right)p\left(\boldsymbol{\rho}^t|\boldsymbol{q}^t\right)p\left(\mathbf{v}^t|\boldsymbol{\gamma}^t\right)p\left(\boldsymbol{\gamma}^t|\boldsymbol{c}^t\right)\sum_{\boldsymbol{q}^{t-1},\boldsymbol{c}^{t-1}}p\left(\boldsymbol{q}^t|\boldsymbol{q}^{t-1}\right)p\left(\boldsymbol{c}^t|\boldsymbol{c}^{t-1}\right)\psi\left(\boldsymbol{q}^{t-1}\right)\psi\left(\boldsymbol{c}^{t-1}\right)\\
&=p\left(\mathbf{y}^t|\mathbf{z}^t,\mathbf{v}^t;\boldsymbol{\Omega}^t\right)p\left(\mathbf{z}^t|\boldsymbol{\rho}^t\right)p\left(\boldsymbol{\rho}^t|\boldsymbol{q}^t\right)p\left(\mathbf{v}^t|\boldsymbol{\gamma}^t\right)p\left(\boldsymbol{\gamma}^t|\boldsymbol{c}^t\right)\hat{p}\left(\boldsymbol{q}^t\right)\hat{p}\left(\boldsymbol{c}^t\right),
\end{aligned}\tag{45}
$$

$\hat{p}\left(\boldsymbol{c}^t\right)$. In what follows, we expound the realization of the approximate posteriors $\psi\left(\boldsymbol{q}^{t-1}\right)$ and $\psi\left(\boldsymbol{c}^{t-1}\right)$.

B. Outline of DiLuS-STPL

As a key feature of DiLuS-STPL, the approximate posteriori $\psi\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\boldsymbol{\Omega}^t\right)$ can be attained in the DiLuS-STPL-E Step using the observation at the tth slot and the message passed from the previous slot, while updating $\boldsymbol{\Omega}^t$ by leveraging a gradient-based technique, e.g., majorization-minimization (MM), in the DiLuS-STPL-M Step. Due to myriads of loops in such a dense factor graph, the classical sum-product message passing (SPMP) [39] fails to traverse the factor and variable nodes over the entire factor graph. As a potential resolution, the approximate message passing (AMP)-based algorithms, e.g., Turbo-AMP [40] and Turbo-CS [41], tend to be entrapped into their local optimum and exhibit usually poor performance. In an effort to facilitate the practical implementation, two entities are specified in the DiLuS-STPL-E Step, i.e., Entity A and Entity B as illustrated in Fig. 2, in support of sparse VBI estimates and forward-backward message passing, such that the proposed algorithm may achieve approximate message passing over the entire factor graph with a faithful performance. Concretely, the message passing over the entire factor graph can be interpreted as an iteration of this pair of partitioned entities, one of which undertakes inference on one subgraph that encapsulates the group sparsity in the observation, and the other of which plays inference on the subgraph that captures the spatial correlation among vehicle positions. Furthermore, the interactions between the twin entities may effectively eliminate the self-reinforcement during the update procedure.

1) Entity A: This entity intends to obtain the approximate conditional marginal posteriors of latent variables using sparse VBI estimation, i.e., $\psi\left(\boldsymbol{\alpha}^t\right)=\psi\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\boldsymbol{\Omega}^t\right)$. Let $\nu_{\chi\to q_m^t}\left(q_m^t\right),m=1,...,M$ denote the input of Entity A, which can be interpreted as the message passed from Entity B and incorporates the spatial correlation among different vehicles in the vehicle platoon. The sparse VBI estimates executed in Entity A will be detailed later in Sec. IV-D.

2) Entity B: The message $\nu_{h_{\boldsymbol{\rho}}\to q_m^t}\left(\cdot\right)$ provided by Entity A constructs the input of Entity B. For ease of exposition, the input and the output of Entity B are defined as $\nu_m^{in}\triangleq\nu_{h_{\boldsymbol{\rho}}\to q_m^t}\left(\cdot\right)$ and $\nu_m^{out}\triangleq\nu_{\chi\to q_m^t}\left(\cdot\right),m=1,...,M$, respectively. In each algorithm iteration, the message $\nu_m^{in}\left(q_m^t\right)=\nu_{h_{\boldsymbol{\rho}}\to q_m^t}\left(q_m^t\right)$ can be obtained in accordance with the belief propagation rule [39]

$$
\nu_m^{in}\left(q_m^t\right)\propto\frac{\psi(q_m^t)}{\nu_m^{out}\left(q_m^t\right)},\forall q_m^t.\tag{46}
$$

Next, let us examine the off-grid location q_m^t's corresponding technique as an example of detailing the forward-backward message passing. By leveraging the belief propagation rule, the forward messages $\bar{\nu}_m^{f'}$ and $\bar{\nu}_m^f$ associated with q_m^t are respectively given by

$$
\bar{\nu}_m^{f'}\left(q_m^t\right)\propto\begin{cases}p\left(q_1^t\right), & m=1,\\ \int_{-q_m^t}p\left(q_m^t|q_{m-1}^t\right)\bar{\nu}_m^f\left(q_m^{t-1}\right), & 2\le m\le M,\end{cases}\tag{47}
$$

$$
\bar{\nu}_{m+1}^f\left(q_m^t\right)\propto\bar{\nu}_m^{f'}\left(q_m^t\right)\bar{\nu}_m^{in}\left(q_m^t\right)\bar{\nu}_{h_{T,m}^t\to q_m^t}\left(q_m^t\right).\tag{48}
$$

The message $\bar{\nu}_{h_{T,m}^t\to q_m^t}\left(q_m^t\right)$ characterizes the temporal correlation of the nth VUE, implying that the location in the previous slot provides a priori for the current slot, which, in particular, is given by

$$
\bar{\nu}_{h_{T,m}^t\to q_m^t}\left(q_m^t\right)\propto\int_{-q_m^t}p\left(q_m^t|q_m^{t-1}\right)\bar{\nu}_{q_m^{t-1}\to h_{T,m}^t}\left(q_m^{t-1}\right).\tag{49}
$$

In (49), as demonstrated in Fig. 3, message $\bar{\nu}_{q_m^{t-1}\to h_{T,m}^t}\left(q_m^{t-1}\right)$ corresponds to the approximate posterior $\psi\left(q_m^{t-1}\right)$ associated with the off-grid location q_m^t of the mth VUE in the tth slot, and thus we have

$$
\bar{\nu}_{q_m^{t-1}\to h_{T,m}^t}\left(q_m^{t-1}\right)\propto\bar{\nu}_{h_{S,m}^{t-1}\to q_m^{t-1}}\left(q_m^{t-1}\right)\triangleq\psi\left(q_m^{t-1}\right).\tag{50}
$$

The backward messages $\bar{\nu}_m^{b'}$ and $\bar{\nu}_m^b$ respectively given by

$$
\bar{\nu}_{m+1}^{b'}\left(q_m^t\right)\propto\textstyle\int_{-q_m^t}p\left(q_{m+1}^t|q_m^t\right)\bar{\nu}_{m+1}^b\left(q_{m+1}^t\right),1\le m\le M-1,\tag{51}
$$

$$
\bar{\nu}_m^b\left(q_m^t\right)\propto\begin{cases}\bar{\nu}_{m+1}^{b'}\left(q_m^t\right)\bar{\nu}_m^{in}\left(q_m^t\right)\bar{\nu}_{h_{T,m}^t\to q_m^t}\left(q_m^t\right), & 1\le m\le M-1,\\ \bar{\nu}_M^{in}\left(q_m^t\right)\bar{\nu}_{h_{T,m}^t\to q_M^t}\left(q_m^t\right), & m=M.\end{cases}\tag{52}
$$

We omit here, for the sake of conciseness, the intricacies of the similar procedure that pertains to the message passing associated with the NLoS channel support $c_{m,n}^t$.

Having established the method for carrying out message passing between the twin entities, let us proceed to a thorough investigation of the DiLuS-STPL-M Step and the DiLuS-STPL-E Step in what follows, respectively, based upon which the off-grid location and corresponding parameters can be determined.

C. DiLuS-STPL-M Step: Inexact Majorization-Minimization (MM)

The likelihood function $\ln p\left(\mathbf{y}^{1:t};\boldsymbol{\Omega}^t\right)$ is indeed intractable owing to the absence of the closed-form expressions induced by the prohibitive integrals of the joint distribution with respect

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

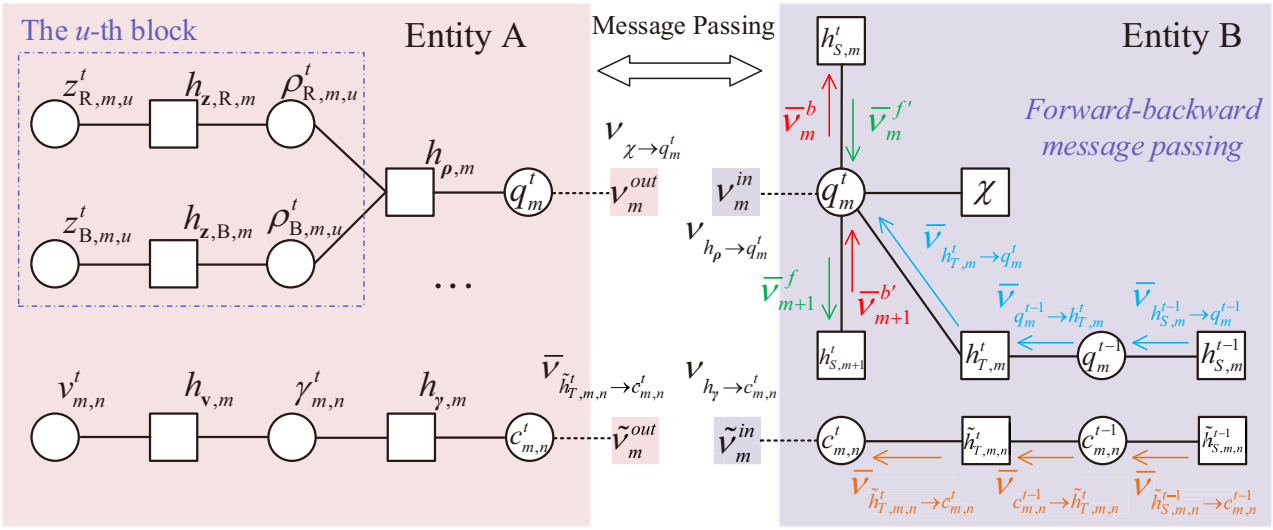


Fig. 3. The partial factor subgraph and the interaction between the two entities. The factors are defined as follows: $h_{\mathbf{z},\mathrm{R},m} \triangleq p\left(z^t_{\mathrm{R},m,u}|\rho^t_{\mathrm{R},m,u}\right)$, $h_{\mathbf{z},\mathrm{B},m} \triangleq p\left(z^t_{\mathrm{B},m,u}|\rho^t_{\mathrm{B},m,u}\right)$, $h_{\boldsymbol{\rho},m} \triangleq p\left(\rho^t_{\mathrm{R},m,u}|q^t_m\right) p\left(\rho^t_{\mathrm{B},m,u}|q^t_m\right)$, $h_{\mathbf{v},m} \triangleq p\left(v^t_{m,n}|\gamma^t_{m,n}\right)$, and $h_{\boldsymbol{\gamma},m} \triangleq p\left(\gamma^t_{m,n}|c^t_{m,n}\right)$.

to $\boldsymbol{\alpha}^t$. An alternative is to construct its surrogate function based on $\boldsymbol{\Omega}^{t,(r)}$ available in the rth iteration in place of the objective in problem (36), providing that the exact posterior $p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\boldsymbol{\Omega}^{t,(r)}\right)$ has been determined. Unfortunately, the exact posterior $p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\boldsymbol{\Omega}^{t,(r)}\right)$ is also highly intractable in our investigated problem because of the intricate loops intrinsic in the factor graph. To circumvent this issue, the VBI and message passing techniques are employed to identify an alternative $\psi\left(\boldsymbol{\alpha}^t\right)$ for a faithful approximation of the posterior $p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\boldsymbol{\Omega}^{t,(r)}\right)$. The to-be-identified posterior $\psi\left(\boldsymbol{\alpha}^t\right)$ has a factorized form, i.e., $\psi\left(\boldsymbol{\alpha}^t\right) = \prod_{l\in\mathcal{L}} \psi\left(\boldsymbol{\alpha}^t_l\right)$, in which $\boldsymbol{\alpha}^t_l$ denotes an individual in $\boldsymbol{\alpha}^t = \{\mathbf{z}^t, \mathbf{v}^t, \boldsymbol{\rho}^t, \boldsymbol{\gamma}^t, \boldsymbol{q}^t, \boldsymbol{c}^t\}$, and we have $\mathcal{L} = \{l|\forall \boldsymbol{\alpha}^t_l \in \boldsymbol{\alpha}^t\}$. Based upon the posterior approximation, an equivalent of the original MLE problem in (36) according to the following proposition.

Proposition 1: Denote $\boldsymbol{\Omega}^{1:t,(r)}$ by the optimal estimates achieved for problem (36) in the rth iteration, the optimality of which is equivalent to the following optimization problem at the tth slot,

$$\max_{\psi(\boldsymbol{\alpha}^t),\boldsymbol{\Omega}^t} \underbrace{\int \psi\left(\boldsymbol{\alpha}^t\right) \ln \frac{p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right)}{\psi\left(\boldsymbol{\alpha}^t\right)} d\boldsymbol{\alpha}^t}_{\mathcal{F}\left(\boldsymbol{\Omega}^t;\boldsymbol{\Omega}^{t,(r)}\right)}. \quad (53)$$

Proof: As regards the optimization problem in (53), we have

$$\begin{aligned}
&\int \psi\left(\boldsymbol{\alpha}^t\right) \ln \frac{p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right)}{\psi\left(\boldsymbol{\alpha}^t\right)} d\boldsymbol{\alpha}^t \\
&\overset{(a)}{\leq} \ln \int \psi\left(\boldsymbol{\alpha}^t\right) \frac{p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right)}{\psi\left(\boldsymbol{\alpha}^t\right)} d\boldsymbol{\alpha}^t \\
&= \ln \int p\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right) d\boldsymbol{\alpha}^t, \quad (54)
\end{aligned}$$

where (a) holds due to the Jensen's inequality. Thus, problem (53) can be maximized when $\psi\left(\boldsymbol{\alpha}^t\right)$ has the following form

$$\psi^*\left(\boldsymbol{\alpha}^t\right) = p\left(\boldsymbol{\alpha}^t|\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right), \quad (55)$$

and the optimization problem is pruned to

$$\max_{\boldsymbol{\Omega}^t} \ln p\left(\mathbf{y}^{1:t};\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right). \quad (56)$$

Therefore, the optimal estimates, i.e., $\boldsymbol{\Omega}^{t,(r)}$, are the optimum to the original problem (53), which thus completes the proof. ■

With the approximation of the joint distribution in (45), the objective function in (53) can be recast as follows

$$\hat{\mathcal{F}}\left(\boldsymbol{\Omega}^t;\boldsymbol{\Omega}^{t,(r)}\right) = \int \psi\left(\boldsymbol{\alpha}^t\right) \ln \frac{\hat{p}\left(\mathbf{y}^{1:t},\boldsymbol{\alpha}^t;\hat{\boldsymbol{\Omega}}^{1:t-1,(r)},\boldsymbol{\Omega}^t\right)}{\psi\left(\boldsymbol{\alpha}^t\right)} d\boldsymbol{\alpha}^t. \quad (57)$$

Then in the $(r+1)$th iteration in the DiLuS-STPL-M Step, we update $\boldsymbol{\Omega}^{t,(r+1)}$ as

$$\boldsymbol{\Omega}^{t,(r+1)} = \arg\max_{\boldsymbol{\Omega}^t} \hat{\mathcal{F}}\left(\boldsymbol{\Omega}^t;\boldsymbol{\Omega}^{t,(r)}\right). \quad (58)$$

The non-concave nature of the objective, nevertheless, impedes the pursuit of the global optimum for (58). We thus exploit the following gradient update in an effort to obtain a stationary solution

$$\boldsymbol{\Omega}^{t,(r+1)} = \boldsymbol{\Omega}^{t,(r)} + \iota^{(r)} \frac{\partial \hat{\mathcal{F}}\left(\boldsymbol{\Omega}^t;\boldsymbol{\Omega}^{t,(r)}\right)}{\partial \boldsymbol{\Omega}^t}, \quad (59)$$

where $\iota^{(r)}$ is the step size determined by the Armijo rule [36].

D. DiLuS-STPL-E Step: Closed-form Updates

1) Outline of Sparse VBI: The approximate conditional marginal posterior can be determined by a minimization of the Kullback-Leibler divergence (KLD) between

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$\hat{p}\left(\mathbf{y}^{1:t}, \boldsymbol{\alpha}^t; \hat{\boldsymbol{\Omega}}^{1:t-1}, \boldsymbol{\Omega}^t\right)$ and $\psi\left(\boldsymbol{\alpha}^t\right)$, subject to the factorized structure on $\psi\left(\boldsymbol{\alpha}^t\right)$, thus yielding the following problem

$$\begin{aligned}\psi^*\left(\boldsymbol{\alpha}^t\right) = &\underset{\psi(\boldsymbol{\alpha}^t), \boldsymbol{\Omega}^t}{\arg\min} \int \psi\left(\boldsymbol{\alpha}^t\right) \ln \frac{\psi\left(\boldsymbol{\alpha}^t\right)}{\hat{p}\left(\boldsymbol{\alpha}^t | \mathbf{y}^{1:t}, ; \boldsymbol{\Omega}^t\right)} d\boldsymbol{\alpha}^t \\ &s.t. \quad \psi\left(\boldsymbol{\alpha}^t\right) = \prod_{l \in \mathcal{L}} \psi\left(\boldsymbol{\alpha}_l^t\right). \end{aligned} \tag{60}$$

The stationary solution $\psi^*\left(\boldsymbol{\alpha}_l^t\right)$ to problem (60) needs to satisfy

$$\begin{aligned}&\psi^*\left(\boldsymbol{\alpha}_l^t\right) \\ &= \underset{\psi\left(\boldsymbol{\alpha}_l^t\right), \boldsymbol{\Omega}^t}{\arg\min} \int \prod_{l \neq j} \psi^*\left(\boldsymbol{\alpha}_l^t\right) \psi\left(\boldsymbol{\alpha}_j^t\right) \ln \frac{\prod_{l \neq j} \psi^*\left(\boldsymbol{\alpha}_l^t\right) \psi\left(\boldsymbol{\alpha}_j^t\right)}{\hat{p}\left(\boldsymbol{\alpha}^t | \mathbf{y}^{1:t}; \boldsymbol{\Omega}^t\right)} d\boldsymbol{\alpha}^t,\end{aligned} \tag{61}$$

which can be attained by triggering alternating updates of $\psi\left(\boldsymbol{\alpha}_l^t\right), \forall l \in \mathcal{L}$. Specifically, for given $\psi\left(\boldsymbol{\alpha}_j^t\right), j \neq l$, a unique solution that facilitates the KLD minimization in problem (60) is given by

$$\psi^*\left(\boldsymbol{\alpha}_l^t\right) \propto \begin{cases} \exp\left(\mathbb{E}_{\prod_{l \neq j} \psi\left(\boldsymbol{\alpha}_j^t\right)} \left[p\left(\boldsymbol{\alpha}^t | \mathbf{y}^{1:t}; \boldsymbol{\Omega}^t\right)\right]\right), t = 1, \\ \exp\left(\mathbb{E}_{\prod_{l \neq j} \psi\left(\boldsymbol{\alpha}_j^t\right)} \left[\hat{p}\left(\boldsymbol{\alpha}^t | \mathbf{y}^{1:t}; \boldsymbol{\Omega}^t\right)\right]\right), t > 1, \end{cases} \tag{62}$$

where $\mathbb{E}_{h(x)}\left[f\left(x\right)\right] = \int f\left(x\right) h\left(x\right) dx$.

2) Initialization of Sparse VBI: The sparse VBI implemented in Entity A is highly dependent on an appropriate initialization, based upon which an alternating algorithm is triggered for the initialization of the corresponding posteriors. In the first outer iteration, the approximate prior of the off-grid location $\boldsymbol{q}^t$ is initialized by

$$\hat{p}\left(\boldsymbol{q}^t\right) = \prod_m \sum_u \pi_{m,u}^{\mathrm{L}} \delta\left(q_m^t - u\right), \tag{63}$$

where $\pi_{m,u}^{\mathrm{L}} = \nu_{\chi \to q_m^t}\left(q_m^t\right)$. Additionally, the approximate prior of the NLoS channel support $\boldsymbol{c}^t$ can be given by

$$\hat{p}\left(\boldsymbol{c}^t\right) = \sum_{\boldsymbol{c}^{t-1}} p\left(\boldsymbol{c}^t | \boldsymbol{c}^{t-1}\right) \psi\left(\boldsymbol{c}^{t-1}\right). \tag{64}$$

3) Update Posterior for LoS Channels $\mathbf{z}^t$: The LoS channel $\psi\left(\mathbf{z}^t\right)$'s can be updated by a complex Gaussian distribution with determined mean and variance

$$\psi\left(\mathbf{z}^t\right) = \mathcal{CN}\left({\mathbf{z}_{\mathrm{R}}}^t; \boldsymbol{\mu}_{\mathrm{R},\mathbf{z}}^t, \boldsymbol{\Sigma}_{\mathrm{R},\mathbf{z}}^t\right) \mathcal{CN}\left({\mathbf{z}_{\mathrm{B}}}^t; \boldsymbol{\mu}_{\mathrm{B},\mathbf{z}}^t, \boldsymbol{\Sigma}_{\mathrm{B},\mathbf{z}}^t\right), \tag{65}$$

where $\boldsymbol{\Sigma}_{i,\mathbf{z}}^t$ and $\boldsymbol{\mu}_{i,\mathbf{z}}^t, i \in \{\mathrm{R}, \mathrm{B}\}$ can be calculated by

$$\begin{aligned}\boldsymbol{\Sigma}_{i,\mathbf{z}}^t = &\left(\mathrm{diag}\left(\rho_{i,m,u}\right) + \left\langle \kappa^t \right\rangle \left(\mathbf{F}_i^t\left(\Delta \mathbf{r}^t\right)\right)^H \mathbf{F}_i^t\left(\Delta \mathbf{r}^t\right)\right)^{-1}, \\ &i \in \{\mathrm{R}, \mathrm{B}\},\end{aligned} \tag{66}$$

$$\begin{aligned}\boldsymbol{\mu}_{i,\mathbf{z}}^t = &\left\langle \kappa^t \right\rangle \boldsymbol{\Sigma}_{i,\mathbf{z}}^t \left(\mathbf{F}_i^t\left(\Delta \mathbf{r}^t\right)\right)^H \\ &\times \left(\mathbf{y}^t - \boldsymbol{\Xi}^t\left(\Delta \boldsymbol{\omega}^{\mathrm{AoA}}, \Delta \boldsymbol{\varphi}^{\mathrm{AoA}}, \Delta \boldsymbol{\vartheta}^{\mathrm{AoA}}\right) \left\langle \mathbf{v}^t \right\rangle\right), i \in \{\mathrm{R}, \mathrm{B}\},\end{aligned} \tag{67}$$

with $\mathbf{F}_{\mathrm{R}}^t\left(\Delta \mathbf{r}^t\right) = \left[\mathbf{H}_{r,b} \boldsymbol{\Theta} \mathbf{A}_N^{\mathrm{LoS}}\left(1\right); ...; \mathbf{H}_{r,b} \boldsymbol{\Theta} \mathbf{A}_N^{\mathrm{LoS}}\left(G\right)\right] \in \mathbb{C}^{KG \times UM}$ and $\mathbf{F}_{\mathrm{B}}^t\left(\Delta \mathbf{r}^t\right) = \left[\mathbf{A}_K^{\mathrm{LoS}}\left(1\right); ...; \mathbf{A}_K^{\mathrm{LoS}}\left(G\right)\right] \in \mathbb{C}^{KG \times UM}$.

4) Update Posterior for NLoS Channels $\mathbf{v}^t$: $\psi\left(\mathbf{v}^t\right)$'s can be updated by

$$\psi\left(\mathbf{v}^t\right) = \mathcal{CN}\left(\mathbf{v}^t; \boldsymbol{\mu}_{\mathbf{v}}^t, \boldsymbol{\Sigma}_{\mathbf{v}}^t\right), \tag{68}$$

where $\boldsymbol{\Sigma}_{\mathbf{v}}^t = \left(\mathrm{diag}\left(\gamma_{m,u}\right) + \left\langle \kappa^t \right\rangle \left(\boldsymbol{\Xi}^t\right)^H \boldsymbol{\Xi}^t\right)^{-1}$ and $\boldsymbol{\mu}_{\mathbf{v}}^t = \left\langle \kappa^t \right\rangle \boldsymbol{\Sigma}_{\mathbf{v}}^t \left(\boldsymbol{\Xi}^t\right)^H \left(\mathbf{y}^t - \mathbf{F}^t\left(\Delta \mathbf{r}_m^t\right) \left\langle \mathbf{z}^t \right\rangle\right)$, with $\left\langle \mathbf{z}^t \right\rangle = \boldsymbol{\mu}_{\mathbf{z}}^t$.

5) Update Posteriors for Precision $\boldsymbol{\rho}^t$ and $\boldsymbol{\gamma}^t$: The posterior of the precision vector $\boldsymbol{\rho}^t$ associated with the LoS channel vector is given by

$$\psi\left(\boldsymbol{\rho}^t\right) = \prod_{i,m,u} \Gamma\left(\rho_{i,m,u}^t; \tilde{\mathfrak{a}}_{i,m,u}^t, \tilde{\mathfrak{b}}_{i,m,u}^t\right), i \in \{\mathrm{R}, \mathrm{B}\}, \tag{69}$$

where $\tilde{\mathfrak{a}}_{i,m,u}^t = \tilde{\pi}_{m,u}^{\mathrm{L}} \mathfrak{a}_{i,m}^t + \left(1 - \tilde{\pi}_{m,u}^{\mathrm{L}}\right) \bar{\mathfrak{a}}_{i,m}^t + 1$ and $\tilde{\mathfrak{b}}_{i,m,u}^t = \tilde{\pi}_{m,u}^{\mathrm{L}} \mathfrak{b}_{i,m}^t + \left(1 - \tilde{\pi}_{m,u}^{\mathrm{L}}\right) \bar{\mathfrak{b}}_{i,m}^t + \left|\left[\mu_{i,\mathbf{z}}^t\right]_{m,u}\right|^2 + \left[\boldsymbol{\Sigma}_{i,\mathbf{z}}^t\right]_{m,u}, i \in \{\mathrm{R}, \mathrm{B}\}$. The posterior associated with the off-grid location $\tilde{\pi}_{m,u}^{\mathrm{L}}$ is given by (70) shown at the top of the next page, in which $\mathfrak{C}_1$ is the normalization constant to let $\sum_{u=1}^U \tilde{\pi}_{m,u}^{\mathrm{L}} = 1$. Additionally, $\psi\left(\boldsymbol{\gamma}^t\right)$ is given by

$$\psi\left(\boldsymbol{\gamma}^t\right) = \prod_{m,n} \Gamma\left(\gamma_{m,n}^t; \widehat{\mathfrak{a}}_{m,n}^N, \widehat{\mathfrak{b}}_{m,n}^N\right), \tag{71}$$

where $\widehat{\mathfrak{a}}_{m,n}^N = \tilde{\pi}_{m,n}^{\mathrm{NL}} \mathfrak{a}_m^N + \left(1 - \tilde{\pi}_{m,n}^{\mathrm{NL}}\right) \bar{\mathfrak{a}}_m^N + 1$ and $\widehat{\mathfrak{b}}_{m,n}^N = \tilde{\pi}_{m,n}^{\mathrm{NL}} \mathfrak{b}_m^N + \left(1 - \tilde{\pi}_{m,n}^{\mathrm{NL}}\right) \bar{\mathfrak{b}}_m^N + \left|\mu_{\mathbf{v},m,n}^t\right|^2 + \boldsymbol{\Sigma}_{\mathbf{v},m,n}^t$. The posterior associated with the off-grid location $\tilde{\pi}_{m,n}^{\mathrm{NL}}$ is given by (72) shown at the top of the next page, with the normalization constant $\mathfrak{C}_2$ given by (73) shown at the top of the next page.

6) Update Posterior for Off-grid Location $\boldsymbol{q}^t$: For the given posterior $\tilde{\pi}_{m,u}^{\mathrm{L}}$, $\boldsymbol{q}^t$'s is updated by

$$\psi\left(\boldsymbol{q}^t\right) = \prod_m \sum_u \tilde{\pi}_{m,u}^{\mathrm{L}} \delta\left(q_m^t - u\right), i \in \{\mathrm{R}, \mathrm{B}\}. \tag{74}$$

7) Update Posterior for NLoS Channel Support $\boldsymbol{c}^t$: Using the posterior $\tilde{\pi}_{m,n}^{\mathrm{NL}}$, the NLoS channel support $\boldsymbol{c}^t$'s is updated by

$$\psi\left(\boldsymbol{c}^t\right) = \prod_{m,n} \tilde{\pi}_{m,n}^{\mathrm{NL}} \delta\left(c_{m,n}^t - 1\right) + \left(1 - \tilde{\pi}_{m,n}^{\mathrm{NL}}\right) \delta\left(c_{m,n}^t\right). \tag{75}$$

E. Message Passing from the $(t-1)$th to the tth Slot

The algorithm running in the tth slot is dependent on the messages passed from the $(t-1)$th slot, e.g., $\psi\left(\boldsymbol{q}^{t-1}\right)$ and $\psi\left(\boldsymbol{c}^{t-1}\right)$. Concerning the fact that the approximate posterior $\psi\left(\boldsymbol{q}^{t-1}\right)$ can be taken from the forward-backward message passing in (50) (as is the same for $\psi\left(\boldsymbol{c}^{t-1}\right)$), the priors $\left(\pi_{m,u}^{\mathrm{L}}\right)^t$ and $\hat{p}\left(\boldsymbol{c}^t\right)$ can be obtained in light of $\pi_{m,u}^{\mathrm{L}} = \nu_{\chi \to q_m^t}\left(q_m^t\right)$ and $\hat{p}\left(\boldsymbol{c}^t\right)$, the former of which incorporates the messages over the forward-backward propagation while the latter of which is derived from (64). Then, the posteriors $\left(\tilde{\pi}_{m,u}^{\mathrm{L}}\right)^t$ and $\left(\tilde{\pi}_{m,n}^{\mathrm{NL}}\right)^t$ can be updated according to (70) and (72). With the closed-form updates in (74) and (75), the message passing across two adjacent slots can thus be achieved. In summary, the proposed DiLuS-STPL algorithm is summarized in Algorithm 1.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

$$\tilde{\pi}^{\mathrm{L}}_{m,u} = \frac{1}{\mathfrak{C}_1}\pi^{\mathrm{L}}_{m,u}\exp\left(\sum_i\left\{\sum_m\left(\mathfrak{a}^t_{i,m}-1\right)\left(\nabla_{\mathfrak{a}^t_{i,m}}\ln\Gamma\left(\tilde{\mathfrak{a}}^t_{i,m,u}\right)-\ln\tilde{\mathfrak{b}}^t_{i,m,u}\right)-\mathfrak{b}^t_{i,m}\frac{\tilde{\mathfrak{a}}^t_{i,m,u}}{\tilde{\mathfrak{b}}^t_{i,m,u}}\right\}\right.$$
$$\left.+\sum_i\left\{\left(\bar{\mathfrak{a}}^t_{i,m}-1\right)\sum_{u\neq u'}\left(\nabla_{\mathfrak{a}^t_{i,m}}\ln\Gamma\left(\tilde{\mathfrak{a}}^t_{i,m,u'}\right)-\ln\tilde{\mathfrak{b}}^t_{i,m,u'}\right)-\bar{\mathfrak{b}}^t_{i,m}\sum_{u\neq u'}\frac{\tilde{\mathfrak{a}}^t_{i,m,u'}}{\tilde{\mathfrak{b}}^t_{i,m,u'}}\right\}\right), i\in\{\mathrm{R},\mathrm{B}\}. \tag{70}$$

$$\tilde{\pi}^{\mathrm{NL}}_{m,n} = \frac{1}{\mathfrak{C}_2}\hat{p}\left(c^t_m\right)\frac{\left(\mathfrak{b}^N_m\right)^{\mathfrak{a}^N_m}}{\Gamma\left(\mathfrak{a}^N_m\right)}\exp\left\{\left(\mathfrak{a}^N_m-1\right)\left(\nabla_{\mathfrak{a}^N_m}\ln\Gamma\left(\mathfrak{a}^N_m\right)-\ln\mathfrak{b}^N_m\right)-\mathfrak{b}^N_m\frac{\widehat{\mathfrak{a}}^t_{m,n}}{\widehat{\mathfrak{b}}^t_{m,n}}\right\}. \tag{72}$$

$$\mathfrak{C}_2 = \hat{p}\left(c^t_m\right)\frac{\left(\mathfrak{b}^N_m\right)^{\mathfrak{a}^N_m}}{\Gamma\left(\mathfrak{a}^N_m\right)}\exp\left\{\left(\mathfrak{a}^N_m-1\right)\left(\nabla_{\mathfrak{a}^N_m}\ln\Gamma\left(\widehat{\mathfrak{a}}^t_{m,n}\right)-\ln\widehat{\mathfrak{b}}^t_{m,n}\right)-\mathfrak{b}^N_m\frac{\widehat{\mathfrak{a}}^t_{m,n}}{\widehat{\mathfrak{b}}^t_{m,n}}\right\}$$
$$+\left(1-\pi^{\mathrm{NL}}_{m,n}\right)\frac{\left(\bar{\mathfrak{b}}^N_m\right)^{\mathfrak{a}^N_m}}{\Gamma\left(\bar{\mathfrak{a}}^N_m\right)}\exp\left\{\left(\bar{\mathfrak{a}}^N_m-1\right)\left(\nabla_{\mathfrak{a}^N_m}\ln\Gamma\left(\widehat{\mathfrak{a}}^t_{m,n}\right)-\ln\widehat{\mathfrak{b}}^t_{m,n}\right)-\bar{\mathfrak{b}}^N_m\frac{\widehat{\mathfrak{a}}^t_{m,n}}{\widehat{\mathfrak{b}}^t_{m,n}}\right\}. \tag{73}$$

F. Computational Complexity and Practical Implementation

This subsection discusses the computational complexity and offers a glimpse into the practical implementation of the proposed DiLuS-STPL algorithm. Following the overall flow of DiLuS-STPL outlined in Algorithm 1, the main computational burden is dominated by the updates of the posteriors $\psi\left(\mathbf{z}^t\right)$, $\psi\left(\mathbf{v}^t\right)$, $\psi\left(\boldsymbol{\rho}^t\right)$, $\psi\left(\boldsymbol{\gamma}^t\right)$, $\psi\left(\boldsymbol{q}^t\right)$, and $\psi\left(\boldsymbol{c}^t\right)$ in Step 8 using (65)-(75), respectively. More precisely, the computational complexity of the matrix inversion operation in (66) is $\mathcal{O}\left(U^3M^3\right)$, along with the total number of multiplications for updating $\psi\left(\mathbf{z}^t\right)$ being $\mathcal{O}\left(K^3G^3+3UM+2KGU^2M^2+2K^2G^2UM+K^2G^2\right)$. Thus, the complexity order can be given by $\mathcal{O}\left(KGU^2M^2\right)$, since it typically holds that $KG \ll UM$. Regarding its practical implementation, the low-latency requirements of the considered RIS-aided vehicle platoons may not be satisfied when running DiLuS-STPL based on a regular desktop computer with a PYTHON implementation. It is just what we use to provide a theoretical guideline. This issue can be effectively addressed by leveraging the powerful and parallel computing power of the cloud, in which the BS first sends the received signals $\mathbf{y}^t$ to the cloud, and then the cloud runs the DiLuS-STPL algorithm to attain the exact locations of M VUEs based on the received signals $\mathbf{y}^t$. Therefore, the low-latency requirements of the practical system can be met due to the large-bandwidth and high-rate fronthaul links between the BS and the cloud [42].

V. Simulation Results

A. Simulation Configuration

A vehicle platoon system consisting of M VUEs is customized according to the generation procedure in [34]. To be specific, the spatial correlation of the M VUEs positioned on a road is modeled using a sampled Gamma distribution as presented in (26) with the shape and scale parameters given by $\varpi = 2$ and $\lambda = 1.5$. Concerning that the inter-VUE distance is a random variable due to the fluctuation in the speed control, it is assumed that the speed of each VUE follows a complex Gaussian distribution $\mathcal{CN}\left(\bar{v},\bar{\sigma}^2\right)$ with $\bar{v} = -18$ m/s as its mean and $\bar{\sigma} = 8$ m/s as its standard deviation. In this case, the temporal probability, i.e., $h^t_{T,m}$, is given by

$$h^t_{T,m} \triangleq p\left(q^t_m|q^{t-1}_m\right) = \left(\frac{\bar{\sigma}\Delta t}{\Delta L}\sqrt{2\pi}\right)^{-1}\exp\left(-\frac{1}{2}\left(\frac{q^t_m-\left(q^{t-1}_m+\frac{\bar{v}\Delta t}{\Delta L}\right)}{\bar{\sigma}\Delta t/\Delta L}\right)^2\right), \tag{76}$$

in which $\Delta L = \left\|\mathbf{r}^t_{m,u}-\mathbf{r}^t_{m,u-1}\right\|$ denotes the grid length and Δt is the slot interval. The BS and the RIS are positioned on each side of the road, both of which are oriented parallel to the road. The three-dimensional coordinates of the BS and the RIS in meters are $[50,100,25]$ and $[150,0,25]$, respectively. The considered system operates at 7 GHz. Furthermore, the exploited primary performance metric for location estimates is the root-mean-square-error (RMSE), i.e., $\sqrt{\frac{1}{TSM}\sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{m=1}^{M}\left\|\hat{\mathbf{r}}^t_m-\mathbf{r}^t_m\right\|^2}$, where S is the number of random channel realizations generated in a slot, and thus the estimate output for each slot t have been averaged over these S samples. The number of observation slots for the system is $T = 100$. Other system parameters are set as follows if not otherwise specified: $M = 4$, $K = 16$, $N = 256$, $\zeta_{m,b} = 3$, $\zeta_{m,r} = 2.5$, $\zeta_{r,b} = 2$ $\Delta L = 1$ m, $\Delta t = 0.1$ s, $S = 200$, $\sigma^2 = -100$ dBm, and $\widetilde{\varrho}_{m,n} = 0.3$.

B. Convergence Behavior

In Fig. 4, we show the convergence behavior of the proposed algorithm by plotting the RMSE curves associated with the estimated location $\hat{\mathbf{r}}^t_m$. Specifically, it is observed in Fig. 4(a) that the RMSE performance decreases rapidly with the iteration number, and ultimately converges to a value within around

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

Algorithm 1 **D**iverse Dynam**i**c **L**ayer Str**u**ctured **S**parsity **S**patial-**T**emporal **P**latoon **L**ocalization (DiLuS-STPL) Algorithm

for $t = 1, 2, ..., T$ **do**
  Initialize $\Delta\mathbf{r}^t = \mathbf{0}$, $\left(\Delta\boldsymbol{\omega}^{\mathrm{AoA}}\right)^t = \mathbf{0}$, $\left(\Delta\boldsymbol{\varphi}^{\mathrm{AoA}}\right)^t = \mathbf{0}$, $\left(\Delta\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)^t = \mathbf{0}$, hyper-parameters $\mathfrak{a}_m^t$, $\mathfrak{b}_m^t$, $\bar{\mathfrak{a}}_m^t$, $\bar{\mathfrak{b}}_m^t$, $\mathfrak{a}_m^N$, $\mathfrak{b}_m^N$, $\bar{\mathfrak{a}}_m^N$, $\bar{\mathfrak{b}}_m^N$, $\mathfrak{a}_m^\kappa$, $\mathfrak{b}_m^\kappa$, the maximum iteration number $r_{\max}$, and the convergence threshold ϵ, ϵ', ε, ε'.
  repeat
    DiLuS-STPL-E Step:
    % Entity A: Sparse VBI Estimation
    Initialize the distribution according to (63) and (64), the message from Entity B passing to Entity A, i.e., $\nu_{\chi\to q_m^t}(q_m^t)$.
    repeat
      Update posteriors $\psi(\mathbf{z}^t)$, $\psi(\mathbf{v}^t)$, $\psi(\boldsymbol{\rho}^t)$, $\psi(\boldsymbol{\gamma}^t)$, $\psi(\boldsymbol{q}^t)$, and $\psi(\boldsymbol{c}^t)$ in compliance with (65)-(75) in Sec. V. D.
    until The convergence criteria of DiLuS-STPL-E Step is met.
    Calculate the message from Entity A, i.e., $\nu_{h_{\boldsymbol{\rho}}\to q_m^t}$, and then send it to Entity B.
    % Entity B: Forward-backward Message Passing
    Update the message $\nu_{\chi\to q_m^t}(q_m^t)$, and then send it to Entity A.
    DiLuS-STPL-M Step:
    Construct surrogate functions $\hat{\mathcal{F}}\left(\boldsymbol{\Omega}^t; \boldsymbol{\Omega}^{t,(r)}\right)$ in (57) using the output of Entity A in DiLuS-STPL-E Step, i.e., $\psi(\boldsymbol{\alpha}^t)$.
    Update $\boldsymbol{\Omega}^{t,(r+1)}$ alternatively according to (59).
  until The convergence criteria $\left\|\boldsymbol{\mu}_{\mathbf{z}}^{t,(r-1)} - \boldsymbol{\mu}_{\mathbf{z}}^{t,(r)}\right\| \leq \epsilon$, $\left\|\boldsymbol{\mu}_{\mathbf{v}}^{t,(r-1)} - \boldsymbol{\mu}_{\mathbf{v}}^{t,(r)}\right\| \leq \epsilon'$, $\left\|\boldsymbol{\Sigma}_{\mathbf{z}}^{t,(r-1)} - \boldsymbol{\Sigma}^{t,(r)}\right\| \leq \varepsilon$, and $\left\|\boldsymbol{\Sigma}_{\mathbf{z}}^{t,(r-1)} - \boldsymbol{\Sigma}_{\mathbf{z}}^{t,(r)}\right\| \leq \varepsilon'$ are met or the maximum iteration number is exceeded.
  Obtain the optimal posteriors $\psi^*(q_m^t)$ and $\psi^*\left(c_{m,n}^t\right)$, as well as the parameters $\boldsymbol{\Omega}^{t,*}$. Then, pass both posteriors to the $(t+1)$th slot.
  Estimate $\hat{q}_m^t$ according to $\hat{q}_m^t = \arg\max_{q_t}\psi^*(q_m^t)$, and the location estimate in the tth slot is given by $\hat{\mathbf{p}}_m^t = \mathbf{r}_{m,\hat{q}}^t + \Delta\hat{\mathbf{r}}_{m,\hat{q}}^t$.
end for

5 iterations in a slot. Fig. 4(b) demonstrates that the RMSE performance tends to stabilize over time, and DiLuS-STPL eventually converges to a better performance. In the first slot, DiLuS-STPL performs poor due primarily to the unavailability of any priori provided. With an increase of the slot, DiLuS-STPL takes advantage of the temporal correlations and the priori delivered by the previous slot for an appropriate parameter adjustment. Upon achieving the performance saturation, the location tracking process progressively transitions into a steady state.

In order to confirm that the surrogate function adopted in (57) typically provides a safe approximation of the likelihood function in (36), Fig. 5 depicts the variation of surrogate

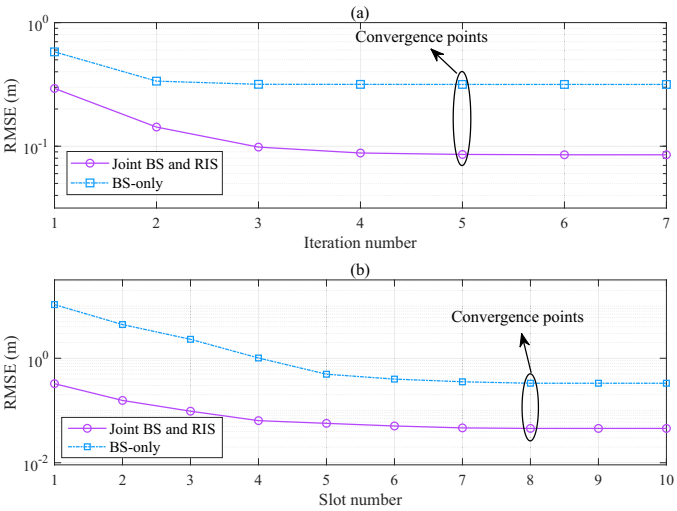


Fig. 4. Convergence behaviors of the proposed DiLuS-STPL algorithm.

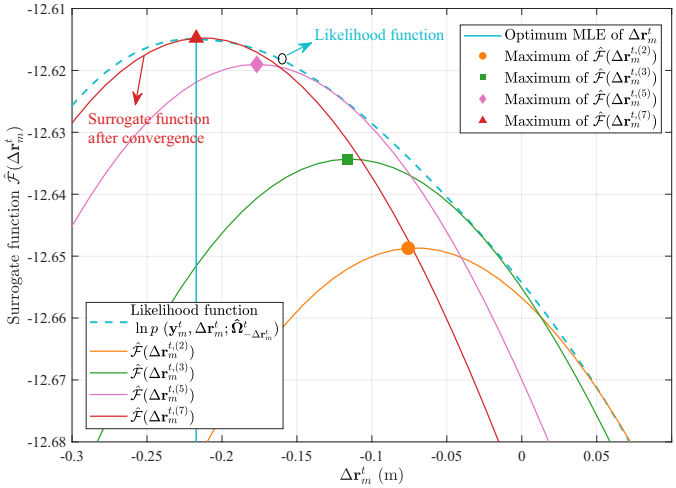


Fig. 5. Surrogate function $\hat{\mathcal{F}}\left(\Delta\mathbf{r}_m^t\right)$ versus the estimated location $\Delta\mathbf{r}_m^t$ during the convergence procedure in a slot.

function $\hat{\mathcal{F}}\left(\Delta\mathbf{r}_m^t\right)$ versus the estimated location $\left(\Delta\mathbf{r}_m^t\right)$ during the convergence procedure in a slot. We are primarily concerned with the impact introduced by the estimated off-grid location, and thus it is assumed here that other angle-associated offsets, e.g., $\Delta\mathbf{r}^t$, $\left(\Delta\boldsymbol{\omega}^{\mathrm{AoA}}\right)^t$, $\left(\Delta\boldsymbol{\varphi}^{\mathrm{AoA}}\right)^t$, and $\left(\Delta\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)^t$, are perfectly available during the estimation. It can be intuitively observed that the surrogate function, as the iteration number increases, approaches the original likelihood more closely. As DiLuS-STPL converges in the seventh iteration, the estimated location $\left(\Delta\mathbf{r}_m^t\right)$ (also the maximum of $\hat{\mathcal{F}}\left(\Delta\mathbf{r}_m^{t,(7)}\right)$) almost coincides with an optimum MLE, indicating that our theoretical analyses are in excellent agreement with simulation results.

C. The Impact of RIS on Localization

1) GDOP Performance Evaluation: Fig. 6 plots the GDOP performance versus the varying VUE location in different points of the road, in which the location tracking relies exclusively upon a single BS positioned at $[50, 100, 25]$ (m) in

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

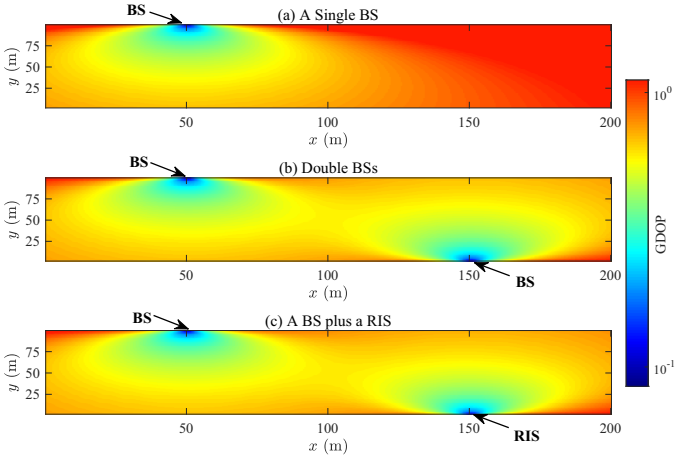


Fig. 6. GDOP performance versus the varying VUE location in different points of the road when $K = 4$, $N = 256$.

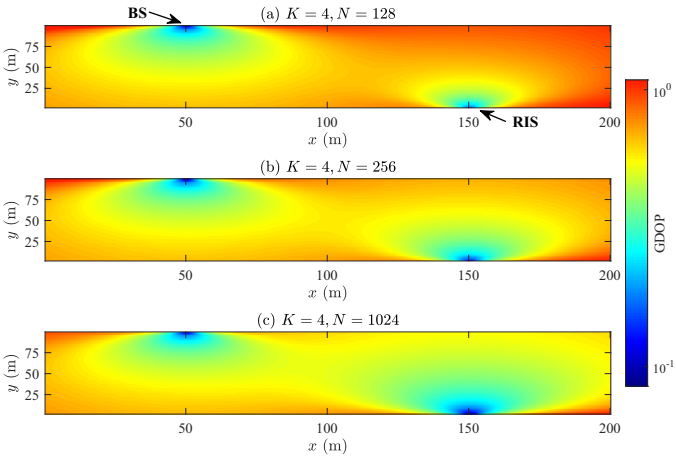


Fig. 7. GDOP performance achieved by a BS plus a RIS for different antenna/element configurations.

(a); a pair of BSs operate cooperatively for location tracking in (b), with their positions given by $[50, 100, 25]\,(\mathrm{m})$ and $[150, 0, 25]\,(\mathrm{m})$, respectively; a BS and a RIS operate in tandem in (c), with their positions given by $[50, 100, 25]\,(\mathrm{m})$ and $[150, 0, 25]\,(\mathrm{m})$, respectively. Firstly, as observed, the achieved GDOP is dependent on the relative VUE location with respect to the BS and RIS. Only when the VUE approaches the BS or RIS will a lower estimation error be achieved, whereas the localization precision deteriorates immediately as soon as the VUE moves further away from them. Given that the uniform grid is fixed, the angular resolution $\left\|\varphi_m^{\mathrm{AoA}}\left(\mathbf{r}_{m,u}^t + \Delta\mathbf{r}_{m,u}^t\right) - \varphi_m^{\mathrm{AoA}}\left(\mathbf{r}_{m,u-1}^t + \Delta\mathbf{r}_{m,u-1}^t\right)\right\|_F$ associated with the off-grid location declines with an increased $\frac{\mathbf{r}_{m,u}^t}{h_v}$ as u grows larger, which leads to the unfavorable column-non-orthogonality in the LoS array response $\mathbf{A}_K^{\mathrm{LoS}}\left(\varphi_m^{\mathrm{AoA}}\left(\mathbf{r}_{m,u}^t + \Delta\mathbf{r}_m^t\right)\right)$, eroding the precision of estimates. Secondly, the dual BSs cooperating for localization provides the best-case bound by virtue of their powerful signal processing capabilities. Intriguingly, we notice that the case of the BS and RIS operating in tandem exhibits the almost comparable localization performance with respect to GDOP, with only marginal precision loss spotted on the side of the RIS. This can be attributed to the following facts. The passive nature of the RIS regulates its reflective behavior in response to an incident signal. The localization signal, in this context, may entail attenuation due to the double fading effect [32], leading to a drop in propagation signal-to-noise ratio (SNR) and an increase in estimate error. Despite this, no noticeable performance loss is spotted, which reveals RIS's great potentials for enabling high-accuracy localization owing to its cost-effective implementations in practice.

2) Localization Capabilities: BS vs. RIS: One may argue that the double-BS configuration could supplant the combination of a BS and a RIS, albeit their comparable localization performance demonstrated in Fig. 6. To further justify this, we would like to go deeper into the rationale of employing a BS and RIS for localization based upon the simulation results in Fig. 7. Specifically, the GDOP ratio $\mathrm{GDOP}_{\mathrm{BS}}/\mathrm{GDOP}_{\mathrm{RIS}}$ presented in (23) motivates us to explore just their individual localization capabilities in a neat manner, instead of more complicated GDOP forms, e.g., $\mathrm{GDOP}_{\mathrm{BS\&RIS}}$ and $\mathrm{GDOP}_{\mathrm{BS\&BS}}$. Thus, we are interested in the case of a BS and a RIS functioning in tandem with different antenna/element configurations. The BS and the RIS are placed at $[50, 100, 25]\,(\mathrm{m})$ and $[150, 0, 25]\,(\mathrm{m})$, respectively. It is intuitive in Fig. 7 that an increase in N boosts RIS's localization capability in terms of the GDOP performance. This occurrence is in good agreement with the fact drawn in (23), since the ratio $\mathrm{GDOP}_{\mathrm{BS}}/\mathrm{GDOP}_{\mathrm{RIS}}$ is proportionate to N_h and N_v. A larger N has the potential to compensate for the multiplicative fading nature inherent to the cascaded channels, thus achieving favorable localization performance. Therefore, we would prefer to implement a BS plus a RIS rather than the double-BS configuration, due to the extremely low implementation cost of a RIS (about $ 7000 [43]) in comparison to the high cost of a 5G BS (about $ 23,000 [44]). Additionally, concerning the passive nature of the RIS, additional interference cannot be introduced to the users in a single cell.

D. Performance of DiLuS-STPL

For comparison, we employ the following several benchmark schemes:

- Naive VBI [27]: It is assumed that the naive VBI algorithm employs i.i.d. priors for both LoS and NLoS channels without considering the diverse sparsity structures presented in DiLuS.
- MAP [38]: The fundamental of MAP is a Bayesian filter employed to approximate the posteriors of the corresponding parameters, i.e., $\mathbf{r}^t$, $\left(\boldsymbol{\omega}^{\mathrm{AoA}}\right)^t$, $\left(\boldsymbol{\varphi}^{\mathrm{AoA}}\right)^t$, and $\left(\boldsymbol{\vartheta}^{\mathrm{AoA}}\right)^t$, whose extremum points can be determined by a brute-force-kind approach.
- LASSO [45]: The LASSO algorithm is a widely used CS-based sparse signal recovery technique that directly filters the non-zero entries in a sparse vector by a soft or hard threshold.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

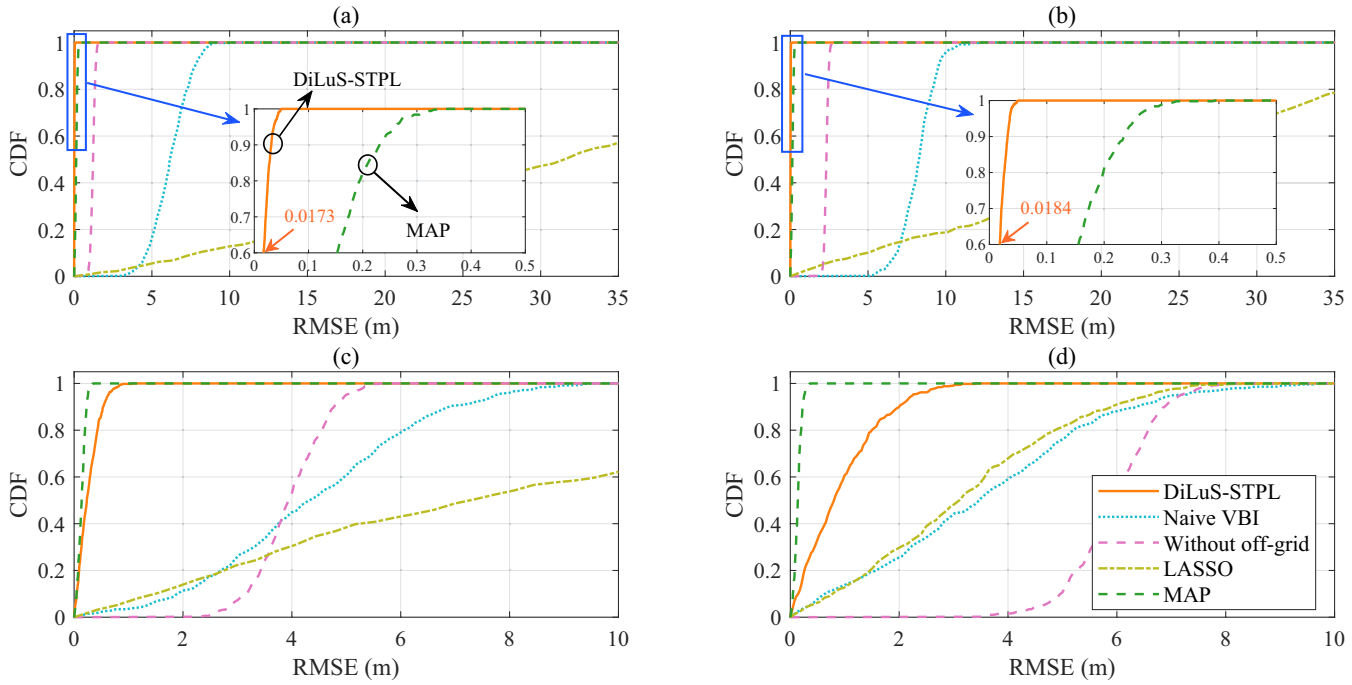


Fig. 8. CDF of the RMSE for various grid resolutions. (a) The grid length is $\Delta L = 0.5$ m, and the number of grids is $U = 200$. (b) The grid length is $\Delta L = 1$ m, and the number of grids is $U = 100$. (c) The grid length is $\Delta L = 5$ m, and the number of grids is $U = 40$. (d) The grid length is $\Delta L = 20$ m, and the number of grids is $U = 10$.

- Without off-grid: The proposed DiLuS-STPL is utilized for location tracking by employing the sparse priors presented in DiLuS without factoring the off-grid basis, i.e., the offsets $\Delta\mathbf{r}^t$, $\left(\Delta\boldsymbol{\omega}^{\text{AoA}}\right)^t$, $\left(\Delta\boldsymbol{\varphi}^{\text{AoA}}\right)^t$, and $\left(\Delta\boldsymbol{\vartheta}^{\text{AoA}}\right)^t$ are set to be zero.

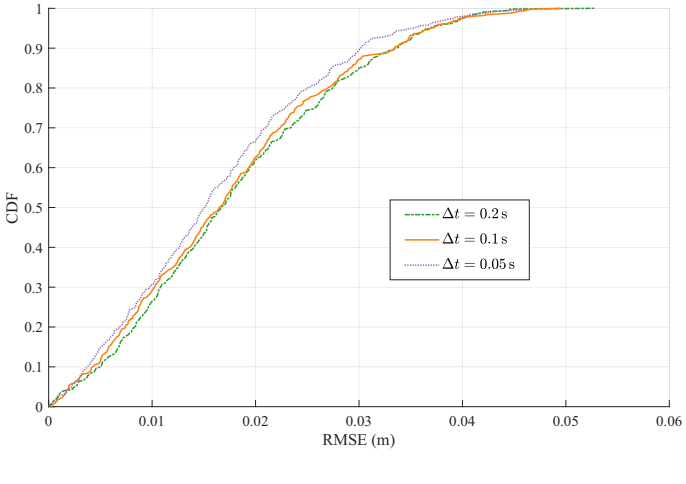
1) Impact of Grid Resolution: We present in Fig. 8 the CDF of RMSE under several schemes for different grid resolutions. It can be observed from Fig. 8 that our proposed DiLuS-STPL achieves highly desirable RMSE performance for all grid resolution cases and even outperforms that of MAP for the instances illustrated in Figs. 8(a)&(b), i.e., the grid length $\Delta L = \left\|\mathbf{r}_{m,u}^t - \mathbf{r}_{m,u-1}^t\right\|$ being both 0.5 m and 1 m, as well as the number of grids being both $U = 200$ and $U = 100$. This is attributed to the smoothness and convexity of the surrogate function used in DiluS-STPL, as well as the beneficial sparse priors provided by DiLuS, significantly facilitating the posterior approximation and parameter updates. Furthermore, as transpired in Figs. 8(a)&(b), a finer-grained gird resolution delivers fairly slight improvement in RMSE performance but adversely doubles the search space, which might make no sense in the practical algorithm design since we are always striving for a critical compromise between the precision and computational complexity. By comparison, due to the absence of these properties in MAP, its performance is bounded by the grid-by-grid search, in which the grid resolution becomes the bottleneck for achieving a higher level of localization precision. Furthermore, concerning that the approximate posteriors in MAP are independent of the grid length, RMSE somehow does not vary with the grid length using MAP approaches. MAP's use of a brute-force-kind approach allows it still to perform excellently even at coarse grid resolutions, but at the sacrifice of an extremely high computational complexity. Additionally, a somewhat counterintuitive occurrence in Fig. 8 is that the RMSE performance tends to be improved as the grid resolution lowers for the LASSO scheme. This is because LASSO is highly dependent on the column-orthogonality in the sensing matrix [36]. A refined grid length results in marginal differences between the neighboring columns in the AoA array response, i.e., $\mathbf{A}_K^{\text{LoS}}\left(\varphi_m^{\text{AoA}}\left(\mathbf{r}_{m,u}^t + \Delta\mathbf{r}_m^t\right)\right)$, hence rendering the undesirable correlations that may be detrimental to the accuracy of grid estimates.

2) Impact of Temporal Resolution: Fig. 9 shows the CDF of the RMSE performance achieved by DiLuS-STPL with different temporal resolutions, i.e., $\Delta t = 0.05\,\text{s}$, $\Delta t = 0.1\,\text{s}$, and $\Delta t = 0.2\,\text{s}$. As transpired in Fig. 9(a), the RMSE performance might be slightly improved if the slot interval is shortened. Despite this, their differences in regard to the three temporal resolutions are marginal at best and, in practice, may be ignored. Furthermore, Fig. 9(b) depicts an analysis of the possible causes for this slight improvement. In particular, the samples for each temporal resolution follow a truncated Gaussian distribution whose variance is independent on Δt. A finer-grained Δt induces a smaller variance, thus allowing for a more subtle vibration of RMSE throughout the whole observation procedure.

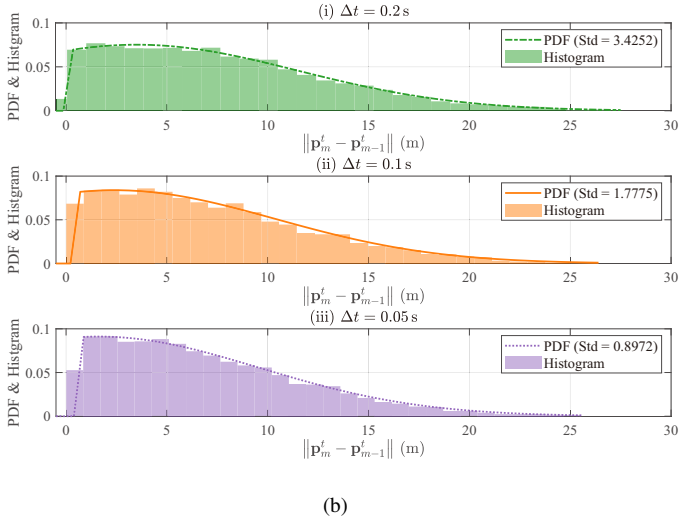
3) Impact of the Number of RIS Elements: In Fig. 10, the RMSE performance is plotted for different numbers of RIS elements N, when the number of BS antennas is fixed with $K = 16$. It is evident that RMSE performance obtained by all schemes improves with an increased N due to the enhancement in angular resolution. As N grows to a certain magnitude, neither the naive VBI nor the MAP

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262



(a)



(b)

Fig. 9. (a) CDF of the RMSE for various temporal resolutions. (b) Truncated Gaussian distributions with various variances associated with different Δt.

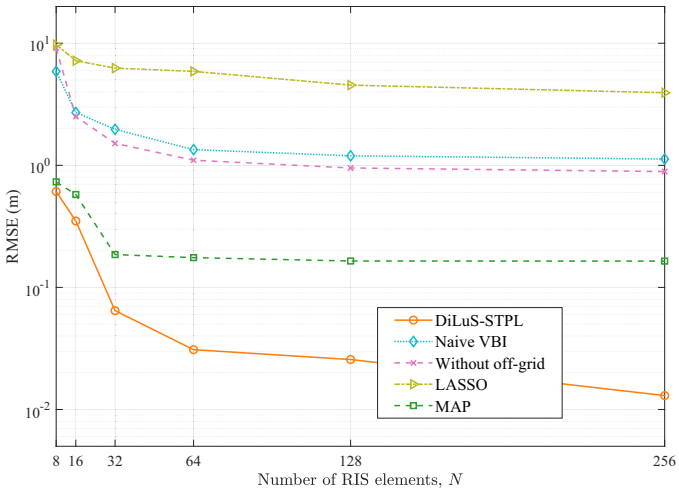


Fig. 10. RMSE performance versus the number of RIS elements.

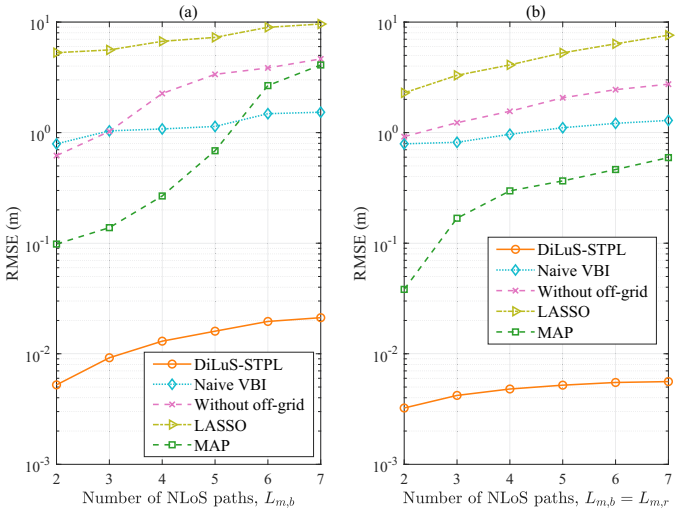


Fig. 11. RMSE performance versus the number of NLoS paths. (a) Location tracking relies exclusively upon the BS. (b) The BS and the RIS jointly track VUE's location.

scheme demonstrates additional precision gains with respect to RMSE. This is due primarily to the absence of a tailored design for capturing additional NLoS priors brought by the cross-slot temporal correlations, as well as the inability to capture the location priors. Furthermore, we note that the RMSE performance of DiLuS-STPL without off-grid levels off as N grows because the estimation error is in general bounded by the grid resolution as portrayed in Fig. 8. This can also be attributed to the fact that DiLuS-STPL without off-grid tends to converge around the grid length. Therefore, we evince that the proposed DiLuS-STPL is capable of harnessing the potential benefits brought by the improved angular resolution, since it encapsulates various sparsity structures associated with the LoS/NLoS channels and the spatial-temporal correlations inherent to the vehicle platoon, in an attempt to recursively identify exact priors for high-accuracy location tracking.

4) Impact of the Number of NLoS Paths: Fig. 11 investigates the RMSE performance versus the varying number of NLoS paths in cases of only the BS being active as well as the BS and RIS functioning in tandem. It is assumed for simplicity that the direct and the cascaded channels hold the same number of NLoS paths, i.e., $L_{m,b} = L_{m,r}$. Firstly, RMSE is intuitively observed to somewhat deteriorate with an increasing number of NLoS paths, particularly when only the BS is responsible for the localization. The increased number of NLoS paths makes it prohibitive to distinguish the LoS paths that carries useful information for location tracking from the NLoS paths, referred to as the NLoS-induced misleading effect, thus eroding the accuracy of localization. Fortunately, the favorable LoS paths constructed by the RIS enriches the localization-associated information, along with the significantly improved accuracy by limiting the NLoS-induced misleading effect towards the LoS paths. Secondly, despite the fluctuating number of NLoS paths, the proposed DiLuS-STPL algorithm levels off even in the presence of an increase in number of NLoS paths, showing its great robustness. Besides, several benchmarks are able to mitigate the NLoS-induced misleading effect while identifying the location-associated LoS channels with the aid of the RIS. This facilitates the precise

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

and robust localization despite the intricate dynamics of the vehicular environment.

VI. Conclusion

In this work, we have examined the RIS as an extension of the BS's localization capability as it provides additional location priors and location-associated LoS paths. By fully exploiting the diverse sparsities that are available in the LoS/NLoS channels, a DiLuS framework is customized for encapsulating the sparse priors of corresponding channels, followed by a MAP-form location tracking problem that is highly intractable due to the uncertain sensing matrix and coupled latent variables associated with the BS and RIS. To resolve this hindrance, the DiLuS-STPL algorithm is developed by recursively achieving updates of the parameters, i.e., sparse LoS/NLoS channel vectors, off-grid location and angle offsets. Our simulation results have verified the significant potential of the RIS to enable high-accuracy localization. We evince through the GDOP metric that location tracking using a BS and a RIS may arrive at the comparable precision performance obtained by the two individual BSs, which is an exciting discovery due to the cost-effective attributes of the RIS in practice. Additionally, the proposed DiLuS-STPL algorithm has the potential to mitigate the NLoS-induced misleading effect while attempting to identify the LoS channels in the presence of a fluctuating number of NLoS paths, thus revealing the critical robustness for high-accuracy localization in vehicle platoons.

Appendix
Derivation in (19)

Derivation is likewise performed by first achieving the outcome for the ULA case, and then extending it to the UPA case. As for the cascaded channel $\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{h}_{m,r}^{t}$, the entry $\mu_{\mathrm{B},k}$ in (13) can be replaced by $\mu_{\mathrm{R},k}$ that represents the (k,n)th entry of $\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{h}_{m,r}^{t}$. With the AoA φ_m^{AoA}, the FIM can be written as

$$\begin{aligned}&\mathrm{FIM}\left(\varphi_m^{\mathrm{AoA}}\right)\\&=\frac{2}{\sigma^2}\,\mathrm{Re}\left\{\nabla_{\varphi_m^{\mathrm{AoA}}}\left(\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{h}_{m,r}^{t}\right)^{H}\nabla_{\varphi_m^{\mathrm{AoA}}}\left(\mathbf{H}_{r,b}\boldsymbol{\Theta}\mathbf{h}_{m,r}^{t}\right)\right\}\\&=\frac{2}{\sigma^2}\,\mathrm{Re}\left\{\nabla_{\varphi_m^{\mathrm{AoA}}}\left(\mathbf{h}_{m,r}^{t}\right)^{H}\left(\boldsymbol{\Theta}^{H}\mathbf{H}_{r,b}^{H}\mathbf{H}_{r,b}\boldsymbol{\Theta}\right)\nabla_{\varphi_m^{\mathrm{AoA}}}\mathbf{h}_{m,r}^{t}\right\}.\end{aligned}\tag{77}$$

By letting $\widetilde{\mathbf{H}}_{r,b}=\mathbf{H}_{r,b}\boldsymbol{\Theta}$, the auto-correlation matrix can be given by $\widetilde{\mathbf{H}}_{r,b}^{H}\widetilde{\mathbf{H}}_{r,b}$ whose diagonal entries tend to zero due to its sparsity in the angular-domain representation. The (n,n)th entry of $\widetilde{\mathbf{H}}_{r,b}^{H}\widetilde{\mathbf{H}}_{r,b}$ can thus be approximated as $\left[\widetilde{\mathbf{H}}_{r,b}^{H}\widetilde{\mathbf{H}}_{r,b}\right]_{n,n}\approx|\mathrm{PL}_{r,b}|^2K$. Furthermore, (77) can be recast to

$$\mathrm{FIM}\left(\varphi_m^{\mathrm{AoA}}\right)=|\mathrm{PL}_{r,b}|^2K\cdot\frac{2}{\sigma^2}\,\mathrm{Re}\left\{\sum_{n=1}^{N}\frac{\partial\left(\mathbf{h}_{m,r}^{t}\right)_n^{*}}{\partial\varphi_m^{\mathrm{AoA}}}\frac{\partial\left(\mathbf{h}_{m,r}^{t}\right)_n}{\partial\varphi_m^{\mathrm{AoA}}}\right\}.\tag{78}$$

We observe that the structural distinction between (13) and (78) resides in a cascaded channel-induced extra term. Thus, by following the flow from (15) to (18), we can obtain $\mathrm{GDOP}_{\mathrm{RIS}}$ in (19).

References

[1] A.-A. A. Boulogeorgos, J. M. Jornet, and A. Alexiou, "Directional terahertz communication systems for 6G: Fact check," *IEEE Veh. Technol. Mag.*, vol. 16, no. 4, pp. 68–77, Dec. 2021.

[2] H. Viswanathan and P. E. Mogensen, "Communications in the 6G era," *IEEE Access*, vol. 8, pp. 57 063–57 074, Mar. 2020.

[3] Y. Chen, Y. Wang, J. Zhang, P. Zhang, and L. Hanzo, "Reconfigurable intelligent surface (RIS)-aided vehicular networks: Their protocols, resource allocation, and performance," *IEEE Veh. Technol. Mag.*, vol. 17, no. 2, pp. 26–36, Jun. 2022.

[4] "Technical specification group services and system aspects; Release 17 description; summary of Rel-17 work items," 3rd Generation Partnership Project, Tech. Rep. 3GPP TR 21.917 V0.7.0, Jun. 2022.

[5] M. H. C. Garcia, A. Molina-Galan, M. Boban, J. Gozalvez, B. Coll-Perales, T. Şahin, and A. Kousaridas, "A tutorial on 5G NR V2X communications," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1972–2026, Third-quarter, 2021.

[6] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Machine learning methods for RSS-based user positioning in distributed massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8402–8417, Dec. 2018.

[7] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu, "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 4, pp. 2634–2657, Fourth-quarter, 2020.

[8] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.

[9] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.

[10] A. Fascista, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "Downlink single-snapshot localization and mapping with a single-antenna receiver," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4672–4684, Jul. 2021.

[11] A. Liu, L. Lian, V. Lau, G. Liu, and M.-J. Zhao, "Cloud-assisted cooperative localization for vehicle platoons: A turbo approach," *IEEE Trans. Signal Process.*, vol. 68, pp. 605–620, Jan. 2020.

[12] H. Kim, S. H. Lee, and S. Kim, "Cooperative localization with constraint satisfaction problem in 5G vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 4, pp. 3180–3189, Apr. 2022.

[13] Y. Tian, S. Liu, W. Liu, H. Chen, and Z. Dong, "Vehicle positioning with deep-learning-based direction-of-arrival estimation of incoherently distributed sources," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20 083–20 095, Oct. 2022.

[14] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[15] H. Zhang, S. Zeng, B. Di, Y. Tan, M. Di Renzo, M. Debbah, Z. Han, H. V. Poor, and L. Song, "Intelligent omni-surfaces for full-dimensional wireless communications: Principles, technology, and implementation," *IEEE Commun. Mag.*, vol. 60, no. 2, pp. 39–45, Feb. 2022.

[16] E. C. Strinati *et al.*, "Reconfigurable, intelligent, and sustainable wireless environments for 6G smart connectivity," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 99–105, Oct. 2021.

[17] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Metalocalization: Reconfigurable intelligent surface aided multi-user wireless indoor localization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7743–7757, Dec. 2021.

[18] A. Fascista, M. F. Keskin, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "RIS-aided joint localization and synchronization with a single-antenna receiver: Beamforming design and low-complexity estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1141–1156, Aug. 2022.

[19] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, P. Popovski, and H. Wymeersch, "RIS-enabled SISO localization under user mobility and spatial-wideband effects," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1125–1140, Aug. 2022.

[20] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, "Indoor localization with reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 161–165, Jan. 2021.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

[21] G. C. Alexandropoulos, I. Vinieratou, and H. Wymeersch, "Localization via multiple reconfigurable intelligent surfaces equipped with single receive RF chains," *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 1072–1076, May 2022.

[22] B. Teng, X. Yuan, R. Wang, and S. Jin, "Bayesian user localization and tracking for reconfigurable intelligent surface aided MIMO systems," *IEEE J. Sel. Topics Signal Process.*, to appear, 2022.

[23] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, "LOS/NLOS near-field localization with a large reconfigurable intelligent surface," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4282–4294, Jun. 2022.

[24] W. Wang and W. Zhang, "Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6282–6297, Oct. 2021.

[25] H. Zhang, B. Di, K. Bian, Z. Han, H. V. Poor, and L. Song, "Toward ubiquitous sensing and localization with reconfigurable intelligent surfaces," *Proc. IEEE*, to appear, 2022.

[26] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Trans. Signal Process.*, vol. 69, pp. 5386–5402, Aug. 2021.

[27] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.

[28] J. Tan, Y. Ma, and D. Baron, "Compressive imaging via approximate message passing with image denoising," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2085–2092, Apr. 2015.

[29] Y. Chen, Y. Wang, and L. Jiao, "Robust transmission for reconfigurable intelligent surface aided millimeter wave vehicular communications with statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 928–944, Feb. 2022.

[30] F. Zhu, A. Liu, and V. K. N. Lau, "Channel estimation and localization for mmwave systems: A sparse bayesian learning approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019.

[31] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[32] Y. Chen, Y. Wang, Z. Wang, and P. Zhang, "Robust beamforming for active reconfigurable intelligent omni-surface in vehicular communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 3086–3103, Oct. 2022.

[33] I. Sharp, K. Yu, and Y. J. Guo, "GDOP analysis for positioning system design," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3371–3382, Sep. 2009.

[34] B. Li, "Stochastic modeling for vehicle platoons (I): Dynamic grouping behavior and online platoon recognition," *Transp. Res. Part B Methodological*, vol. 95, pp. 364–377, Jan. 2017.

[35] X. Liu, W. Wang, X. Song, X. Gao, and G. Fettweis, "Sparse channel estimation via hierarchical hybrid message passing for massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7118–7134, Nov. 2021.

[36] A. Liu, G. Liu, L. Lian, V. K. N. Lau, and M.-J. Zhao, "Robust recovery of structured sparse signals with uncertain sensing matrix: A turbo-VBI approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3185–3198, May 2020.

[37] P. Munezero, "Bayesian sequential inference for dynamic regression models," Ph.D. dissertation, Department of Statistics, Stockholm University, 2020.

[38] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, *Bayesian multiple target tracking*. Artech House, 2013.

[39] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[40] 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.

[41] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3439–3448, Jul. 2012.

[42] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Commun. Surv. Tutor.*, vol. 17, no. 1, pp. 405–426, First-quarter, 2015.

[43] [Online]. Available: actenna.cn.

[44] [Online]. Available: https://cntechpost.com/2020/04/07/how-much-does-it-cost-to-build-a-5g-base-station.

[45] C. Thrampoulidis, A. Panahi, and B. Hassibi, "Asymptotically exact error analysis for the generalized l_2^2-LASSO," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, China, Jun. 2015, pp. 2021–2025.

Yuanbin Chen received the B.S. degree in communications engineering from the Beijing Jiaotong University, Beijing, China, in 2019. He is currently pursuing the Ph.D. degree in information and communication systems with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications. His current research interests are in the area of reconfigurable intelligent surface (RIS), vehicle-to-everything (V2X), and radio resource management (RRM) in future wireless networks. He was the recipient of the National Scholarship in 2020 and 2022.

Ying Wang (IEEE Member) received the Ph.D. degree in circuits and systems from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2003. In 2004, she was invited to work as a Visiting Researcher with the Communications Research Laboratory (renamed NiCT from 2004), Yokosuka, Japan. She was a Research Associate with the University of Hong Kong, Hong Kong, in 2005. She is currently a Professor with BUPT and the Director of the Radio Resource Management Laboratory, Wireless Technology Innovation Institute, BUPT. Her research interests are in the area of the cooperative and cognitive systems, radio resource management, and mobility management in 5G systems. She is active in standardization activities of 3GPP and ITU. She took part in performance evaluation work of the Chinese Evaluation Group, as a Representative of BUPT. She was a recipient of first prizes of the Scientific and Technological Progress Award by the China Institute of Communications in 2006 and 2009, respectively, and a second prize of the National Scientific and Technological Progress Award in 2008. She was also selected in the New Star Program of Beijing Science and Technology Committee and the New Century Excellent Talents in University, Ministry of Education, in 2007 and 2009, respectively. She has authored over 100 papers in international journals and conferences proceedings.

Xufeng Guo received the B.E. degree in communications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2021. He is currently pursuing the Ph.D. degree in information and communication systems with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications. His current research interests are in the area of reconfigurable intelligent surface (RIS), extremely large-scale multiple-input-multiple-output (XL-MIMO), and signal processing techniques in future wireless networks.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3288262

Zhu Han (S'01–M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a John and Rebecca Moores Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. Dr. Han's main research targets on the novel game-theory related concepts critical to enabling efficient and distributive use of wireless networks with limited resources. His other research interests include wireless resource allocation and management, wireless communications and networking, quantum computing, data science, smart grid, security and privacy. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Dr. Han was an IEEE Communications Society Distinguished Lecturer from 2015-2018, AAAS fellow since 2019, and ACM distinguished Member since 2019. Dr. Han is a 1% highly cited researcher since 2017 according to Web of Science. Dr. Han is also the winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."

Ping Zhang (IEEE Fellow) is a Professor with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, the Director of the State Key Laboratory of Networking and Switching Technology, a member of IMT-2020 (5G) Experts Panel, and a member of Experts Panel for China's 6G Development. He served as a Chief Scientist of National Basic Research Program (973 Program), an Expert in information technology division of National High-Tech Research and Development Program (863 Program), and a member of Consultant Committee on International Cooperation of National Natural Science Foundation of China. His research interests mainly focus on wireless communications.

Authorized licensed use limited to: University of Houston. Downloaded on July 02,2023 at 14:23:13 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.