RIS-Assisted High Resolution Radar Sensing

Citation for the original published paper (version of record):
Vejling, M., Kim, H., Biscio, C. et al (2025). RIS-Assisted High Resolution Radar Sensing. IEEE Transactions on Signal Processing, 73: 2940-2955. http://dx.doi.org/10.1109/TSP.2025.3586551

RIS-Assisted High Resolution Radar Sensing

Martin V. Vejling, Hyowon Kim *Member, IEEE*, Christophe A. N. Biscio,
Henk Wymeersch *Fellow, IEEE*, Petar Popovski *Fellow, IEEE*

***Abstract*—This paper analyzes monostatic sensing by a user equipment (UE) for a setting in which the UE is unable to resolve multiple targets due to their interference within a single resolution bin. It is shown how sensing accuracy, in terms of both detection rate and localization accuracy, can be boosted by a reconfigurable intelligent surface (RIS), which can be advantageously used to provide signal diversity and aid in resolving the targets. Specifically, assuming prior information on the presence of a cluster of targets, a RIS beam sweep procedure is used to facilitate the high resolution sensing. This setting requires a tailored Fisher analysis, as well as introduction of two new coherence concepts that are central to the derived theoretical bounds, namely the Cramér-Rao lower bound and a new upper bound on the detection probability. Next, we propose an orthogonal matching pursuit channel estimation algorithm combined with data association to fuse the information of the non-RIS signal and the RIS signal and perform sensing. Finally, we provide numerical results to verify the potential of RIS for improving sensor resolution, and to demonstrate that the proposed methods can realize this potential for RIS-assisted high resolution sensing.**

***Index Terms*—reconfigurable intelligent surface, high resolution sensing, orthogonal matching pursuit, Cramér-Rao lower bound, detection probability**

I. INTRODUCTION

Integrated sensing and communication (ISAC) is anticipated to be a significant technological advancement in Beyond-5G and 6G communication systems [1]. It is motivated by the limited availability of spectrum caused by increasing demands on key performance indicators in sensing and communication, and the potential mutual benefits between the signals of the two systems [2]. Reconfigurable Intelligent Surfaces (RISs) can be seen as one enabling technology for ISAC, offering the possibility to advantageously control the wireless propagation environment [3], [4].

Sensing should not degrade communication performance and therefore rely on minimal requirements on energy, spectrum usage, and delay due to pilot transmissions [1]. Still, sensing needs to be accurate and implementable by low-cost devices with limited hardware specifications to make it ubiquitous [5]. It can be observed that, either the feasible sensing performance is dictated by the limited sensing resources or the sensing performance requirements dictate the necessary sensing resources [1].

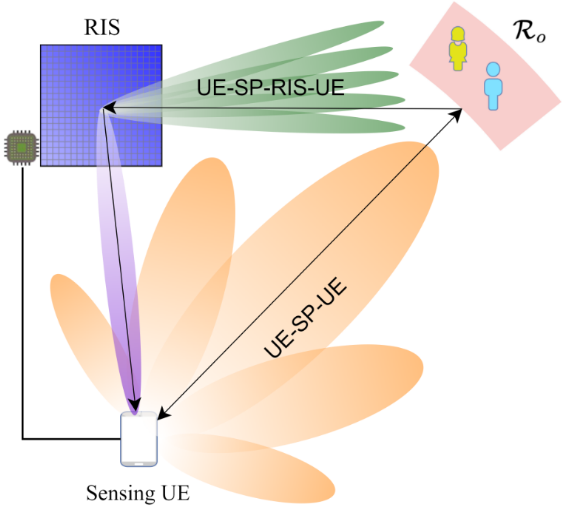


Fig. 1. Illustration of the RIS-aided sensing scenario with a prior (not drawn to scale).

We consider a setup where an RIS is available and devise a full sensing protocol for the environment. The RIS can operate in two modes: (i) exclusive mode where the user equipment (UE) gets associated with the RIS and can send signals via a dedicated control channel, or (ii) shared mode where the RIS changes configurations according to an announced schedule that the UE has to use. In both cases, the monostatic sensing UE attempts to sense the environment and decides whether to use the RIS in shared or exclusive mode.

The scenario is depicted in Fig. 1: here a sensor with limited resolution, i.e., a low bandwidth and a small antenna array, attempts to resolve a scattering point (SP) cluster which we define as a collection of SPs within a resolution bin of the UE. Such sensors can commonly be encountered in applications with WiFi sensors [6], internet of things sensors [7], and reduced capability sensors [8]. Without the RIS, only a single SP is detected which can be critical for some sensing applications. Consider a vehicular network scenario with a human close to a car, as in [9]: we risk causing injury if the human is not detected, however, with the support of an RIS, we can improve the resolution of the sensor [10], thereby avoiding such failures. The improvement in resolution is due to two properties: (i) the RIS adds spatial diversity allowing a view of the SP cluster from a different direction, and (ii) the RIS can have a high angular resolution [10]. Examples of application scenarios are smart environments [11], herein indoor monitoring [12], autonomous robots in industry 4.0 [13], and sensing-assisted device-to-device communication [14].

M. V. Vejling (mvv@{math,es}.aau.dk, corresponding author) is with the Dept. of Mathematical Sciences and the Dept. of Electronic Systems, Aalborg University, Denmark. C. A. N. Biscio (christophe@math.aau.dk) is with the Dept. of Mathematical Sciences, Aalborg University, Denmark. P. Popovski (petarp@es.aau.dk) is with the Dept. of Electronic Systems, Aalborg University, Denmark. H. Wymeersch (henkw@chalmers.se) is with the Dept. of Electrical Engineering, Chalmers University of Technology, Sweden. H. Kim (hyowon.kim@cnu.ac.kr) is with the Dept. of Electronics Engineering, Chungnam National University, Daejeon, South Korea. The work of M. V. Vejling and P. Popovski was partly funded by the Villum Investigator Grant “WATER” financed by the Villum Foundation, Denmark. The work of H. Wymeersch has been supported, in part, by the SNS JU project 6G-DISAC under the EU’s Horizon Europe research and innovation program under Grant Agreement No 101139130.

A. Related work

Whilst sensing with RIS has significant potential [10], there are also limitations, particularly due to the severe attenuation of double-bounce paths [15]. For this reason, most works focus on RIS-enabled scenarios where: RIS enables a UE to sense SPs in non-line-of-sight (NLOS) through the indirect line-of-sight (LOS) path via the RIS [16], [17]; RIS provides an anchor in place of a base station (BS) which can be used for self-localization [18]; cooperative sidelink positioning is enabled by using multiple RISs as anchor points [19], [20]; multiple RISs with a priori unknown locations are used for simultaneous localization and mapping (SLAM) [21]; and more [22]. All these scenarios differ substantially from the monostatic sensing problem considered in this paper as the problems cannot be solved without the RIS and since they do not consider clusters of SPs.

Still, some works explore passive RIS-assisted scenarios [23]–[26]. A two-step approach to SLAM based on weighted least squares is presented in [23], exploiting the NLOS paths from active and passive scatterers to assist in localization. However, they assume perfectly known complex channel coefficients such that the RIS phase profiles can be optimized to maximize the radar cross section for sensing. This is not an assumption that can be realized in fast fading environments. In [24], RIS phase profiles are optimized according to the signal-to-noise-ratio (SNR), and the target detection of RIS-aided sensing is considered. The analysis confirms the intuition that an RIS is best when positioned close to the radar (transmitter or receiver) such that the path loss of the additional bounce is limited. A monostatic sensing scenario considering both single bounce (SB) and double bounce (DB) multipaths is considered in [25], showing significant performance gains by exploiting the RIS, however, with the assumption that the sensing UE is only a few wavelengths away from the RIS, which essentially requires that the RIS is attached to the UE. A similar scenario is considered in [26], however, without collocated UE and RIS. The main takeaway is that the DB measurements provide no real performance gains. Hence, a commonality in these works on sensing assisted by a passive RIS is that the RIS should essentially be attached to the monostatic radar.

A potential direction to alleviate the limitations caused by the DB attenuation is that of active RIS [27]. With an active RIS, the RIS amplifies the signal instead of just passively reflecting it. With such an amplification, a satisfactory SNR of the DB path can be realized, and the potential was demonstrated in [28]. However, the cost of the active RIS comes in terms of the power consumed at the RIS, and for this reason we consider only the case of passive RIS in this work.

Thus, there is somewhat a consensus in the prior work that the assistance of passive RIS for sensing has limited capability due to the DB attenuation. Despite this, our work starts from the following observation: when the sensing device has low resources, for instance a small antenna array, then the high directionality provided by the RIS may still provide a substantial benefit. The potential of passive RIS for sensing with low resource devices has been explored in [29]–[31]: [29] considers sensing in an indoor scenario using commodity WiFi signals and RIS, while [30], [31] investigates high-resolution capabilities of RIS in the realm of extended targets. The motivation of [29] is that the low resolution properties in angle and delay of WiFi makes sensing challenging but this can be alleviated with an RIS. Considering then a dynamic indoor scenario with static interfering reflectors, they optimize the RIS phases to locate the moving targets. In [30], a cross-entropy loss function is optimized for the RIS phase profiles and the parameters of a neural network that maps the received signal into the estimate for the target shape. The optimization relies on a Markov decision process framework, and notably such an approach requires an extensive training procedure to calibrate the model, thereby making it incapable of adapting to non-stationary scenarios. In a related work [31], the RIS focuses beams towards each point in a grid, and an optimization algorithm is proposed to estimate the target shape. This method relies on the use of an impractical number of RIS phase profiles, and the complexity of the optimization problem scales cubically with the number of grid points.

B. Contributions and organization

We position our work in extension to [26] by considering a scenario not previously considered with a low resource device attempting to sense multiple SPs within a single resolution bin, and show that in such a case using the DB signal can significantly improve sensing performance. Specifically, we consider the scenario depicted in Fig. 1, where a sensor with limited resolution attempts to resolve an SP cluster, and devise a full sensing protocol in which the UE, following an initial sensing step, can ask for RIS assistance in shared or exclusive mode. We study how an RIS can improve the effective sensing resolution, presenting a novel upper bound on the detection probability and the Cramér-Rao lower bound (CRLB) on positioning. In specific cases, we draw new insights from analytical expressions of these bounds and introduce a coherence and a generalized coherence concept. Using orthogonal matching pursuit (OMP), the SB and DB parameters are estimated in parallel, followed by data association in the domain of the angle of arrival (AOA) at the RIS, and joint position estimation, realized via weighted non-linear least squares.

Our main contributions can be summarized as follows:

1) We propose an upper bound on the detection probability of each target which is derived from an oracle estimator relying on a greedy approach similar to OMP. For the specific case of three targets, analytical expressions are provided. We define two new concepts of coherence and show how they limit the detection probability. The first coherence concept captures pair-wise interactions between targets, while the generalized coherence concept summarizes the interactions between three targets.
2) A Fisher analysis of channel parameter estimation and localization is presented, deriving delay error bounds (DEBs), angle error bounds (AEBs), and position error bounds (PEBs) for both SB and DB signals, as well as a joint PEB when combining SB/DB information. For a simplified case with two targets, we give an analytical expression for the equivalent Fisher information of the

azimuth angle parameter, yielding the novel discovery that the information loss of the first target due to the second target depends on the new coherence concepts.

3) We use OMP for estimation of the SB and DB channel parameters. To fuse non-RIS and RIS information, we devise a data association method, and develop a joint sensing procedure for the position. Through numerical analysis, we display the potential performance gains by using an RIS for resolving an SP cluster. Particularly, for an SP cluster of three targets, we are able to improve the area under the curve (AUC) for the third target from 0.55 without the RIS to 0.95 with the RIS, and improve the generalized optimal sub-pattern assignment (GOSPA) metric by up to 60 %.

The paper is organized as follows. Section II presents the scenario herein the assumptions on the prior knowledge, the design of precoder and RIS phase profiles, and the resulting signal model. The theoretical bounds on detection, channel estimation, and localization are presented in Section III. In Section IV, the OMP channel estimation algorithm is recalled, followed by the proposed data association method, and the joint position estimation algorithm. Numerical results are presented in Section V, and Section VI concludes the paper and presents some future research directions.

Supplementary resources: The code used for simulations can be found at https://github.com/Martin497/RIS-Assisted-High-Resolution-Radar-Sensing.

Notation: The notations $(\cdot)^\top$, $(\cdot)^H$, $\|\cdot\|$, $\langle\cdot,\cdot\rangle$, $(\cdot)^\dagger$, and $\mathrm{Tr}(\cdot)$ denotes the transpose, Hermitian transpose, 2-norm, inner product, Moore-Penrose pseudo-inverse, and trace operators, respectively. $\mathrm{diag}(a_1,\dots,a_L)$ denotes the diagonal matrix with scalars a_i, $i=1,\dots,L$ on the diagonal. We use $\otimes$ and $\odot$ for the Kronecker and Hadamard products, respectively. $\mathbb{C}^{N\times M}$ denotes the space of $(N\times M)$-dimensional complex matrices. The notation $\Re\{\cdot\}$ and $\Im\{\cdot\}$ refers to the real part and imaginary part of a complex number, respectively, and we use j to denote the complex unit. For a matrix $\boldsymbol{A}$, the notation $\boldsymbol{A}_{a:b}$ refers to the submatrix of $\boldsymbol{A}$ from the a-th column to the b-th column. Moreover, for a matrix $\boldsymbol{A}$, the notation $\exp(j\boldsymbol{A})$ refers to element-wise evaluation of the complex exponential function. Finally, $\mathbb{E}$ and Var denotes the expectation and variance operators, respectively.

II. SYSTEM MODEL

A. Considered scenario

Fig. 1 depicts a scenario where a UE attempts to sense the environment consisting of an RIS and L targets also called SPs. We assume a static scenario[1], so that the UE, RIS, and SPs have fixed positions, in a domain $\mathcal{D}\subset\mathbb{R}^3$. Practically, this assumption can apply when mobility is sufficiently low, for instance in the case of indoor monitoring. The UE's state is known and given as $\boldsymbol{s}=[\boldsymbol{p}^\top,\boldsymbol{o}^\top]^\top$ where $\boldsymbol{p}\in\mathcal{D}$ is the spatial position and $\boldsymbol{o}=[o_1,o_2,o_3]$ is the three-dimensional orientation vector specifying the Euler angles. Similarly, the RIS has state $\boldsymbol{s}_\mathrm{r}=[\boldsymbol{p}_\mathrm{r}^\top,\boldsymbol{o}_\mathrm{r}^\top]^\top$ for spatial position $\boldsymbol{p}_\mathrm{r}\in\mathcal{D}$ and orientation vector $\boldsymbol{o}_\mathrm{r}$. The SPs constitute a point pattern $\Phi=\{\boldsymbol{c}_1,\dots,\boldsymbol{c}_L\}$ where $\boldsymbol{c}_l\in\mathcal{D}$ for $l=1,\dots,L$ and L is the number of SPs which is unknown a priori.

We assume that the UE is full-duplex operating in orthogonal frequency division multiplexing (OFDM) at carrier frequency f_c, such that the wavelength is $\lambda=c/f_c$ where c is the speed of light, and subcarrier frequencies $f_n=f_c+n\Delta_f$, for $n=0,\dots,N-1$ where N is the number of subcarriers and Δ_f is the subcarrier spacing. Then, the bandwidth of the sensing UE is $W=N\Delta_f$. The UE is equipped with a uniform planar array (UPA) consisting of $N_\mathrm{u}=N_\mathrm{u}^\mathrm{az}\times N_\mathrm{u}^\mathrm{el}$ antennas that are spaced with distance $\lambda/2$. We use a precoder $\boldsymbol{f}\in\mathbb{C}^{N_\mathrm{u}}$ $\big(\|\boldsymbol{f}\|=1\big)$ on T time resource blocks. We assume that the RIS has $N_\mathrm{r}=N_\mathrm{r}^\mathrm{az}\times N_\mathrm{r}^\mathrm{el}$ elements with distance $\lambda/4$, and that $N_\mathrm{r}\gg N_\mathrm{u}$. For time $t=1,\dots,T$, $\boldsymbol{\omega}_t\in\mathbb{C}^{N_\mathrm{r}}$ are the RIS phase profiles given by $\boldsymbol{\omega}_t=[\exp(j\varphi_{t,1}),\dots,\exp\,(j\varphi_{t,N_\mathrm{r}})]^\top$ for phase-shift of the i-th element $\varphi_{t,i}$, $i=1,\dots,N_\mathrm{r}$. To allow synchronization and control of RIS, we assume the UE has an out-of-band channel to the RIS controller [32], [33]. We consider far-field propagation, noting that accounting for near-field propagation is outside the scope of this paper.

B. Constructing and using the prior

We assume that in an initial sensing step, a prior map of the environment is made: (i) the SP cluster of interest has been detected and lies in resolution region $\mathcal{R}_o=\{\boldsymbol{c}\in\mathcal{D}\mid\boldsymbol{h}_\mathrm{n}(\boldsymbol{c})\in\bar{\mathcal{R}}_o\}$ where $\boldsymbol{h}_\mathrm{n}(\boldsymbol{c})$ is the mapping from the Euclidean domain into the non-RIS channel parameter space (see Appendix A for details) and $\bar{\mathcal{R}}_o=\bar{\mathcal{R}}_{o,\tau}\times\bar{\mathcal{R}}_{o,\theta^\mathrm{az}}\times\bar{\mathcal{R}}_{o,\theta^\mathrm{el}}$ for resolution intervals in the delay, azimuth, and elevation angles; and (ii) the UE and RIS positions and orientations are perfectly estimated [18].

In practice, we can construct the resolution intervals through the width of the peak in the ambiguity function [34]: we form a bounding box around points that satisfy $(\mathrm{ref}-\mathrm{ref}_\mathrm{th})$ dB, where ref is the value of the ambiguity function in the detection peak and $\mathrm{ref}_\mathrm{th}>0$ is a parameter, and find the largest and smallest delay, azimuth, and elevation angles within this bounding box. Moreover, we can estimate the UE and RIS positions and orientations using existing techniques [18].

After the initial sensing step, the UE requests assistance from the RIS in exclusive mode. Alternatively, if resolving the SP cluster is not time critical or if the RIS control does not allow for real-time reservation and control, the UE can wait for the RIS beam sweep in shared mode to cover the resolution region.

C. Signal model

To separate the RIS and non-RIS parts of the signal, we design the RIS phase profiles as a time-orthogonal sequence, i.e., $\boldsymbol{\omega}_{2\tilde{t}-1}=-\boldsymbol{\omega}_{2\tilde{t}}=\tilde{\boldsymbol{\omega}}_{\tilde{t}}$, $\tilde{t}=1,\dots,\tilde{T}$, $\tilde{T}=T/2$, for design specific RIS phase profiles $\tilde{\boldsymbol{\omega}}_{\tilde{t}}$. Using the prior information about the SP cluster and the RIS position, we construct the precoder, $\boldsymbol{f}$, towards the SP cluster and with a null towards the

---

[1] The RIS-assisted sensing in this work relies on a beam sweep taking up to 0.4 ms to complete, see Section II-D. The width of the beams in the numerical experiments are on a sub-meter scale, hence, in low mobility scenarios the positions can be considered constant.

RIS. This setup allows us to separately observe the UE-SP-UE paths and the UE-SP-RIS-UE paths [26]. [2]

The received non-RIS signal is modelled as

$$\boldsymbol{y}^{\mathrm{n}} = \sum_{l=1}^{L} \alpha_l \boldsymbol{g}^{\mathrm{n}}(\boldsymbol{\eta}_l^{\mathrm{n}}) + \boldsymbol{\varepsilon} = \boldsymbol{G}^{\mathrm{n}}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}^{\mathrm{n}} \tag{1a}$$

where $\boldsymbol{\varepsilon}^{\mathrm{n}}$ is circularly-symmetric complex additive white Gaussian noise (AWGN) with zero mean and covariance $\frac{\sigma^2}{2}\boldsymbol{I}$, $\sigma > 0$, $\boldsymbol{g}_l^{\mathrm{n}} := \boldsymbol{g}^{\mathrm{n}}(\boldsymbol{\eta}_l^{\mathrm{n}}) = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), \boldsymbol{f}\rangle \boldsymbol{1}_{\tilde{T}} \otimes \boldsymbol{d}(\tau_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)$, $\boldsymbol{1}_{\tilde{T}}$ is a column vector of ones with length $\tilde{T}$, $\boldsymbol{G}^{\mathrm{n}} := [\boldsymbol{g}_1^{\mathrm{n}}, \dots, \boldsymbol{g}_L^{\mathrm{n}}]$, $\boldsymbol{\eta}_l^{\mathrm{n}} := [\tau_l, \boldsymbol{\theta}_l^\top]^\top$, and $\boldsymbol{\alpha} := [\alpha_1, \dots, \alpha_L]^\top$. In this model, $\alpha_l \in \mathbb{C}$ are the complex path coefficients modeling path loss due to distance and scattering as well as stochastic phase shifts and radar cross section (RCS) fluctuation loss due to small movements or rotations. $\boldsymbol{\theta}_l = [\theta_l^{\mathrm{az}}, \theta_l^{\mathrm{el}}]^\top$, $l = 1, \dots, L$, is the azimuth and elevation AOA at the UE from the l-th SP, τ_l are the time of arrivals (TOAs) of the paths, $d_n(\tau) = \exp(-j2\pi n\tau\Delta_f)$ is the delay response, and $\boldsymbol{d}(\tau) := [d_0(\tau), \dots, d_{N-1}(\tau)]^\top$ is the delay response vector. The array matrix is given by $\boldsymbol{A}_{\mathrm{u}}(\boldsymbol{\theta}_l, \boldsymbol{\theta}_k) = \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)\boldsymbol{a}_{\mathrm{u}}^\top(\boldsymbol{\theta}_k)$. The array response vector is defined by $\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}) = \exp\big(j\boldsymbol{P}_{\mathrm{u}}^\top \boldsymbol{k}(\boldsymbol{\theta})\big)$ where $\boldsymbol{P}_{\mathrm{u}} = [\boldsymbol{p}_{\mathrm{u},1}, \dots, \boldsymbol{p}_{\mathrm{u},N_{\mathrm{u}}}]$ with $\boldsymbol{p}_{\mathrm{u},i} = [x_{\mathrm{u},i}, y_{\mathrm{u},i}, z_{\mathrm{u},i}]^\top$ are the positions of the antennas in local coordinates, and $\boldsymbol{k}(\boldsymbol{\theta}) = \frac{2\pi}{\lambda}[\cos(\theta^{\mathrm{az}})\sin\big(\theta^{\mathrm{el}}\big), \sin(\theta^{\mathrm{az}})\sin\big(\theta^{\mathrm{el}}\big), \cos\big(\theta^{\mathrm{el}}\big)]^\top$ is the wavenumber vector [35].

The received RIS signal is modelled as [26]

$$\boldsymbol{y}^{\mathrm{r}} = \sum_{l=1}^{L} \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}(\boldsymbol{\eta}_l^{\mathrm{r}}) + \boldsymbol{\varepsilon} = \boldsymbol{G}^{\mathrm{r}}\bar{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}^{\mathrm{r}} \tag{1b}$$

where $\boldsymbol{\varepsilon}^{\mathrm{r}}$ is again zero mean with covariance $\frac{\sigma^2}{2}\boldsymbol{I}$, $\boldsymbol{g}_l^{\mathrm{r}} := \boldsymbol{g}^{\mathrm{r}}(\boldsymbol{\eta}_l^{\mathrm{r}}) = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), \boldsymbol{f}\rangle \boldsymbol{\nu}(\boldsymbol{\phi}_l) \otimes \boldsymbol{d}(\bar{\tau}_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0)$, $\boldsymbol{G}^{\mathrm{r}} := [\boldsymbol{g}_1^{\mathrm{r}}, \dots, \boldsymbol{g}_L^{\mathrm{r}}]$, $\boldsymbol{\eta}_l^{\mathrm{r}} := [\boldsymbol{\theta}_l^\top, \bar{\tau}_l, \boldsymbol{\phi}_l^\top]^\top$, and $\bar{\boldsymbol{\alpha}} := [\bar{\alpha}_1, \dots, \bar{\alpha}_L]^\top$. The RIS response is $\nu(\boldsymbol{\phi}_l, \boldsymbol{\phi}_0; \boldsymbol{\omega}_t) = \boldsymbol{\omega}_t^\top\big(\boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_l) \odot \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_0)\big)$, with RIS response vector $\boldsymbol{\nu}(\boldsymbol{\phi}) := [\nu(\boldsymbol{\phi}, \boldsymbol{\phi}_0; \tilde{\boldsymbol{\omega}}_1), \dots, \nu(\boldsymbol{\phi}, \boldsymbol{\phi}_0; \tilde{\boldsymbol{\omega}}_{\tilde{T}})]^\top$. The parameter $\boldsymbol{\phi}_l = [\phi_l^{\mathrm{az}}, \phi_l^{\mathrm{el}}]^\top$, $l = 1, \dots, L$, is the AOA at the RIS from the l-th SP, and $\boldsymbol{\phi}_0$ is the AOA at the RIS from the UE. Similarly, $\boldsymbol{\theta}_0$ is the AOA at the UE from the RIS. As before, the array response vector is defined as $\boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}) = \exp\big(j\boldsymbol{P}_{\mathrm{r}}^\top \boldsymbol{k}(\boldsymbol{\phi})\big)$ where $\boldsymbol{P}_{\mathrm{r}} = [\boldsymbol{p}_{\mathrm{r},1}, \dots, \boldsymbol{p}_{\mathrm{r},N_{\mathrm{r}}}]$ with $\boldsymbol{p}_{\mathrm{r},i} = [x_{\mathrm{r},i}, y_{\mathrm{r},i}, z_{\mathrm{r},i}]^\top$ are the positions of the RIS elements in local coordinates. The complex path coefficients are α_0 and $\bar{\alpha}_l \in \mathbb{C}$. Finally, $\tau_0, \bar{\tau}_l$ models the TOAs of the paths. The specifications of all the introduced channel parameters is elaborated in Appendix A.

D. RIS phase profile design

We design the RIS phase profiles $\tilde{\omega}_{\tilde{t}}$ to direct beams towards sub-regions of $\mathcal{R}_{\tilde{t}} \subset \mathcal{R}_o$, as illustrated in Fig. 1. Implicitly, $\mathcal{R}_o$ specifies bounds in the AOA at the RIS, denoted $\phi_{\min}^{\mathrm{az}}$, $\phi_{\max}^{\mathrm{az}}$, $\phi_{\min}^{\mathrm{el}}$, and $\phi_{\max}^{\mathrm{el}}$. Given this, let $\phi_i^{\mathrm{az}} = \phi_{\min}^{\mathrm{az}} + \zeta_{\mathrm{az}}/2 + i\zeta_{\mathrm{az}}$ for $i = 0, \dots, D_{\mathrm{az}} - 1$ and $\phi_i^{\mathrm{el}} = \phi_{\min}^{\mathrm{el}} + \zeta_{\mathrm{el}}/2 + i\zeta_{\mathrm{el}}$ for $i = 0, \dots, D_{\mathrm{el}} - 1$ be sample points where $\zeta_{\mathrm{az}} = (\phi_{\max}^{\mathrm{az}} - \phi_{\min}^{\mathrm{az}})/D_{\mathrm{az}}$ and $\zeta_{\mathrm{el}} = (\phi_{\max}^{\mathrm{el}} - \phi_{\min}^{\mathrm{el}})/D_{\mathrm{el}}$ are the spacing between sample points, and D_{az}, D_{el} satisfy $\tilde{T} = D_{\mathrm{az}}D_{\mathrm{el}}$. Then, we specify the sequence of central angles at the RIS for the directional beamforming as

$$\boldsymbol{\phi}_{\tilde{t}} = \left[\phi_{\tilde{t} - D_{\mathrm{el}}(\lceil \tilde{t}/D_{\mathrm{el}}\rceil - 1)}^{\mathrm{az}}, \phi_{\lceil \tilde{t}/D_{\mathrm{az}}\rceil}^{\mathrm{el}}\right]^\top$$

for $\tilde{t} = 1, \dots, \tilde{T}$, and we can construct focused beams as $\tilde{\boldsymbol{\omega}}_{\tilde{t}} = \big(\boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_{\tilde{t}}) \odot \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_0)\big)^*$. When the RIS is large or the number of channel uses is low, these beams can fail to illuminate the entire sub-region $\mathcal{R}_{\tilde{t}}$. To overcome this issue, we construct $\tilde{\boldsymbol{\omega}}_{\tilde{t}}$ as uniform beams on the region $\mathcal{R}_{\tilde{t}}$ [36].

We have assumed that the angle $\boldsymbol{\phi}_0$ is perfectly known based on the initial sensing step which could rely on methods as in [18], [37]. However, when $\boldsymbol{\phi}_0$ is imperfectly estimated, the focused beams used here may fail to direct the beam back to the UE. In such a case, the RIS beams should be designed to illuminate uniformly an uncertainty region. This comes with a degradation of the SNR that is proportional to the size of this uncertainty region.

III. THEORETICAL PERFORMANCE BOUNDS

We derive performance bounds for channel parameter estimation, position estimation, and detection. First, we consider the problem of detection, conditioning on the positions of the targets, from which we develop a novel upper bound on the detection probability when following a greedy approach. This result can be seen in relation to the lower bound on the probability of correct detection for the OMP algorithm introduced in [38], although we have found this lower bound to be pessimistic relative to what is practically achieveable. Second, conditioning on detecting the correct number of targets, we derive the CRLB on channel parameters as well as the position of the targets. These two theoretical results can be seen as considering individual parts of a complete sensing system.

A. Detection probability

1) General case: In this section, we present an upper bound on the detection probability that share similarities with that of [39] but is generalized to simultaneous detection of multiple targets in proximity. The fundamental idea is to formulate an oracle estimator relying on a greedy approach, similar to OMP, while perfectly estimating the geometrical channel parameters.

Proposition 1. *Let the received signal be on the form* $\boldsymbol{y} = \sum_{l=1}^{L} \alpha_l \boldsymbol{g}_l + \boldsymbol{\varepsilon} = \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ *where* $\boldsymbol{\varepsilon} \sim \mathcal{CN}(\boldsymbol{0}, \frac{\sigma^2}{2}\boldsymbol{I})$*, and* $\boldsymbol{G} = [\boldsymbol{g}_1, \dots, \boldsymbol{g}_L]$*, and* $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_L]^\top$*. Then, following optimal processing with a greedy algorithm yields that the detection probability of the* l*-th target and conditioned on the channel coefficients* $\boldsymbol{\alpha}_{l:L}$*, is*

$$p_{\mathrm{d},l}(\boldsymbol{\alpha}_{l:L}) = Q_1(\sqrt{\mu_l}, \sqrt{\gamma_{\mathrm{th}}}), \tag{2}$$

where $Q_1(\cdot,\cdot)$ *is the Marcum Q-function,* $\gamma_{\mathrm{th}} := -2\log(p_{\mathrm{fa}})$ *for a given false alarm probability,* p_{fa}*, and non-centrality parameter* $\mu_l := A_l\beta_l$ *for* $\beta_l := |\boldsymbol{g}_l^H(\boldsymbol{I} - \boldsymbol{P}_{l-1})\boldsymbol{G}_{l:L}\boldsymbol{\alpha}_{l:L}|^2$, $A_l := \frac{4}{\sigma^2\boldsymbol{g}_l^H(\boldsymbol{I} - \boldsymbol{P}_{l-1})\boldsymbol{g}_l}$*, with* $\boldsymbol{P}_{l-1} := \boldsymbol{G}_{1:l-1}\boldsymbol{G}_{1:l-1}^\dagger$ *for* $l > 1$*, and* $\boldsymbol{P}_0 := \boldsymbol{0}$*.*

---

[2] Note that the UE-RIS-SP-UE paths vanish due to the precoder's null towards the RIS. We opt to use only one of the DB signals because observing both, as in [26], requires dividing the time resources that halves the number of different RIS phase profiles and thereby negatively impacting sensing resolution. Moreover, we use the UE-SP-RIS-UE path as this is amplified with a precoder towards the SP same as the UE-SP-UE path, while for the UE-RIS-SP-UE path a precoder towards the RIS is needed.

Proof. The proof is deferred to Appendix B-A. □

Remark 1. *In Proposition 1 we assume optimal processing in the sense that the geometric parameters are assumed to be known, thereby allowing to analyze the detection problem in isolation. In case the geometric parameters were estimated, one would not know* $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_L$*, but instead have estimates* $\hat{\boldsymbol{g}}_1, \ldots, \hat{\boldsymbol{g}}_L$*. This results in the non-centrality parameter for the* l*-th target instead being*

$$\mu_l^{\text{imperfect}} = \frac{4\left|\hat{\boldsymbol{g}}_l^H(\boldsymbol{I}-\hat{\boldsymbol{P}}_{l-1})\boldsymbol{G}\boldsymbol{\alpha}\right|^2}{\sigma^2\hat{\boldsymbol{g}}_l^H(\boldsymbol{I}-\hat{\boldsymbol{P}}_{l-1})\hat{\boldsymbol{g}}_l},$$

where $\hat{\boldsymbol{P}}_{l-1}$ *is the projection matrix using the estimated channel vectors. Take note that here* $\boldsymbol{\alpha}$*, and not just* $\boldsymbol{\alpha}_{l:L}$*, appears as a direct result of the imperfect geometric parameter estimation.*

To summarize the detection performance across all false alarm probabilities, we will also use the AUC defined as $\mathrm{AUC}_l(\boldsymbol{\alpha}_{l:L}) := \int_0^1 p_{\mathrm{d},l}(\boldsymbol{\alpha}_{l:L})\mathrm{d}p_{\mathrm{fa}}$, considering that the detection probability is a function of the false alarm probability. The AUC quantifies the detection performance across the range of false alarm probability, and it is lower bounded by 0.5, corresponding to the case $p_{\mathrm{d}} = p_{\mathrm{fa}}$, and upper bounded by 1, corresponding to the case $p_{\mathrm{d}} = 1$.

Proposition 1 provides a way to compute the detection probability for both the non-RIS and RIS signals conditioned on the realization of the channel coefficients, denoted respectively by $p_{\mathrm{d},l}^{\mathrm{n}}(\boldsymbol{\alpha}_{l:L})$ and $p_{\mathrm{d},l}^{\mathrm{r}}(\bar{\boldsymbol{\alpha}}_{l:L})$, by considering the signal models (1a) and (1b)

Taking the expectation over the channel coefficients with the Gaussian distribution assumption (cf. Appendix A), yields the result of the following corollary.

Corollary 1. *With the same assumptions of Proposition 1, assuming further that* $\mathcal{R}_o$ *is fixed, and* $|\alpha_l|$ *is Rayleigh distributed with scale parameter* ς_{α_l}*, then the marginal distribution for detection, i.e., the expected detection probability, is given by*

$$p_{\mathrm{d},l} := \mathbb{E}_{\beta_l}[p_{\mathrm{d},l}(\boldsymbol{\alpha}_{l:L})] = \exp\Big(\frac{\log(p_{\mathrm{fa}})}{A_l\zeta_l + 1}\Big), \tag{3}$$

$\zeta_l := \frac{1}{2}\boldsymbol{g}_l^H(\boldsymbol{I}-\boldsymbol{P}_{l-1})\boldsymbol{G}_{l:L}\mathrm{diag}(\varsigma_{\alpha_l}^2, \ldots, \varsigma_{\alpha_L}^2)\boldsymbol{G}_{l:L}^H(\boldsymbol{I}-\boldsymbol{P}_{l-1})\boldsymbol{g}_l$. *Moreover, the AUC is*

$$\mathrm{AUC}_l := \int_0^1 p_{\mathrm{d},l}\mathrm{d}p_{\mathrm{fa}} = \frac{1 + A_l\zeta_l}{2 + A_l\zeta_l}. \tag{4}$$

Proof. The proof is given in Appendix B-B, and is based on evaluating a well-studied integral [40]. □

As before, this result applies straightforwardly to both the non-RIS and RIS signals, with detection probability denoted as $p_{\mathrm{d},l}^{\mathrm{n}}$ and $p_{\mathrm{d},l}^{\mathrm{r}}$, respectively, and AUC denoted respectively by $\mathrm{AUC}_l^{\mathrm{n}}$ and $\mathrm{AUC}_l^{\mathrm{r}}$.

By the probability of the union of not necessarily mutually exclusive events, and assuming that the stochastic channel coefficients $\boldsymbol{\alpha}_{l:L}$ and $\bar{\boldsymbol{\alpha}}_{l:L}$ are independent, the expected probability of detecting the l-th target by either the non-RIS or RIS signals is

$$p_{\mathrm{d},l}^{\mathrm{joint}} = p_{\mathrm{d},l}^{\mathrm{n}} + p_{\mathrm{d},l}^{\mathrm{r}} - p_{\mathrm{d},l}^{\mathrm{n}}p_{\mathrm{d},l}^{\mathrm{r}}, \tag{5}$$

with false alarm probability $p_{\mathrm{fa}}^{\mathrm{joint}} = 2p_{\mathrm{fa}} - p_{\mathrm{fa}}^2$.

For the joint detection to be more effective than individual detections, $p_{\mathrm{d},l}^{\mathrm{joint}}$ as a function of $p_{\mathrm{fa}}^{\mathrm{joint}}$, must exceed $p_{\mathrm{d},l}^{\mathrm{n}}$ and $p_{\mathrm{d},l}^{\mathrm{r}}$ as functions of p_{fa}. This scenario occurs when the detection probabilities for RIS and non-RIS are similar. Conversely, if the RIS detection probability significantly surpasses the non-RIS detection probability, it is preferable to rely solely on the RIS signal, and vice versa.

In general, the projection matrices $\boldsymbol{P}_{l-1}$, and in turn the non-centrality parameters, μ_l, and expected detection probabilities, $p_{\mathrm{d},l}$, are not easy to quantify in terms of the signal and channel parameters, which motivates us to look into a representative case of a cluster with 3 targets.

2) Case study with three targets: To give some insight into the derived detection probability, we consider the case of $L = 3$, i.e., $\boldsymbol{y} = \sum_{l=1}^3 \boldsymbol{g}_l\alpha_l + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon} \sim \mathcal{CN}(\boldsymbol{0}, \frac{\sigma^2}{2}\boldsymbol{I})$ and $|\alpha_l|$ is Rayleigh distributed with scale parameter ς_{α_l}. This provides a simplification with sufficient generality to understand the full problem.[3]

Before proceeding, we define two coherence concepts which will be central to the subsequent analysis.

Definition 1 (Coherence)**.** *The coherence* $C(\boldsymbol{g}_l, \boldsymbol{g}_k) \in [0, 1]$ *of two complex vectors* $\boldsymbol{g}_l, \boldsymbol{g}_k \in \mathbb{C}^N$ *is*

$$C(\boldsymbol{g}_l, \boldsymbol{g}_k) = \frac{|\langle \boldsymbol{g}_l, \boldsymbol{g}_k\rangle|^2}{\|\boldsymbol{g}_l\|^2\|\boldsymbol{g}_k\|^2}. \tag{6}$$

Remark 2. *By the Cauchy-Schwarz inequality, the coherence equals one if and only if* $\boldsymbol{g}_l$ *and* $\boldsymbol{g}_k$ *are linearly dependent, and the coherence is zero if and only if* $\boldsymbol{g}_l$ *and* $\boldsymbol{g}_k$ *are orthogonal.*

Definition 2 (Generalized Coherence)**.** *The generalized concept of coherence* $\mathscr{C}(\boldsymbol{g}_l, \boldsymbol{g}_k, \boldsymbol{g}_i) \in [0, 1]$ *of three complex vectors* $\boldsymbol{g}_l, \boldsymbol{g}_k, \boldsymbol{g}_i \in \mathbb{C}^N$ *is*

$$\mathscr{C}(\boldsymbol{g}_l, \boldsymbol{g}_k, \boldsymbol{g}_i) := \frac{\|\boldsymbol{g}_l\langle \boldsymbol{g}_i, \boldsymbol{g}_k\rangle - \boldsymbol{g}_k\langle \boldsymbol{g}_i, \boldsymbol{g}_l\rangle\|^2}{\|\boldsymbol{g}_l\|^2\|\boldsymbol{g}_k\|^2\|\boldsymbol{g}_i\|^2(1 - C(\boldsymbol{g}_l, \boldsymbol{g}_k))}). \tag{7}$$

For notational convenience, we define $C_{l,k} := C(\boldsymbol{g}_l, \boldsymbol{g}_k)$ as the coherence between the l-th and k-th targets, and $\mathscr{C} := \mathscr{C}(\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3)$ as the generalized coherence between targets 1, 2, and 3.

The non-centrality parameters, cf. (2), are

$$\mu_1 = \frac{4\|\boldsymbol{g}_1\|^2}{\sigma^2}\Big|\alpha_1 + \alpha_2\frac{\langle \boldsymbol{g}_1, \boldsymbol{g}_2\rangle}{\|\boldsymbol{g}_1\|^2} + \alpha_3\frac{\langle \boldsymbol{g}_1, \boldsymbol{g}_3\rangle}{\|\boldsymbol{g}_1\|^2}\Big|^2, \tag{8a}$$

$$\begin{aligned}\mu_2 = {}& \frac{4\|\boldsymbol{g}_2\|^2(1 - C_{1,2})}{\sigma^2} \\ & \times \Big|\alpha_2 + \alpha_3\frac{\|\boldsymbol{g}_1\|^2\langle \boldsymbol{g}_2, \boldsymbol{g}_3\rangle - \langle \boldsymbol{g}_2, \boldsymbol{g}_1\rangle\langle \boldsymbol{g}_1, \boldsymbol{g}_3\rangle}{\|\boldsymbol{g}_1\|^2\|\boldsymbol{g}_2\|^2(1 - C_{1,2})}\Big|^2,\end{aligned} \tag{8b}$$

$$\mu_3 = \frac{4\|\boldsymbol{g}_3\|^2(1 - \mathscr{C})}{\sigma^2}|\alpha_3|^2, \tag{8c}$$

where we notice that the generalized concept of coherence $\mathscr{C}$ has the same influence for the third target as the coherence $C_{1,2}$ has for the second target. Accordingly, the expected detection probabilities are

$$p_{\mathrm{d},1} = \exp\Big(\frac{\sigma^2\log(p_{\mathrm{fa}})}{4\sum_{l=1}^3 \varsigma_{\alpha_l}^2\|\boldsymbol{g}_l\|^2 C_{1,l} + \sigma^2}\Big), \tag{9a}$$

---

[3] Considering the case of only two targets is not sufficient here: the detection probability of the second target in case of three targets depends both on the error made when subtracting the contribution of the first target, and the contribution of the third target to the signal. Such a case does not occur when considering only two targets.

$$p_{\mathrm{d},2} = \exp\Big(\frac{\sigma^2 \log(p_{\mathrm{fa}})}{4\big(\varsigma_{\alpha_2}^2 \|\boldsymbol{g}_2\|^2(1 - C_{1,2}) + \varsigma_{\alpha_3}^2 \|\boldsymbol{g}_3\|^2(\mathscr{C} - C_{1,3})\big) + \sigma^2}\Big), \tag{9b}$$

$$p_{\mathrm{d},3} = \exp\Big(\frac{\sigma^2 \log(p_{\mathrm{fa}})}{4\varsigma_{\alpha_3}^2 \|\boldsymbol{g}_3\|^2(1 - \mathscr{C}) + \sigma^2}\Big). \tag{9c}$$

The non-centrality parameters and the expected detection probabilities applies to both the non-RIS and RIS signals using the signal models (1a) and (1b), respectively. We can interpret the expected detection probabilities in relation to the expected receive SNR which for the l-th target is $2\varsigma_{\alpha_l}^2 \|\boldsymbol{g}_l\|^2/\sigma^2$. This term appears in the expected detection probabilities (9a) to (9c) in the expected way (the same way as when there is only a single target [39]), however, each time weighted by a coherence factor.

Additional interpretation follows by considering some special cases. Consider first the case that $\boldsymbol{g}_2$ and $\boldsymbol{g}_3$ are linearly dependent (a special case is when targets two and three coincide), then (9) collapses to the two target case, which we deduce by observing that $\mathscr{C} = 1$. Hence, the expected detection performance of the first two targets is not degraded, however, the probability of detecting the third target is equal to the false alarm probability, which is the worst case scenario for detection. Another notable interpretation is that the non-centrality parameter for the second target, see (8b), depends linearly on the complement of the coherence between the first and second targets, $1 - C_{1,2}$.

Considering alternatively a scenario where $\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3$ are pair-wise orthogonal. Then, the expected detection probabilities, cf. (9), simplify to $p_{\mathrm{d},l} = \exp\big(\frac{\sigma^2 \log(p_{\mathrm{fa}})}{4\varsigma_{\alpha_l}^2 \|\boldsymbol{g}_l\|^2 + \sigma^2}\big)$, i.e., the detection probabilities depends only on the expected receive SNR of the l-th path and the false alarm probability.

Overall, the expected detection probabilities depend mainly on the coherence factors, with good detection performance for all targets if the coherence factors are close to zero, and degraded detection performance for the second and third targets, if the coherence factors are close to one. The coherence for non-RIS and RIS signals are

$$C(\boldsymbol{g}_l^{\mathrm{n}}, \boldsymbol{g}_k^{\mathrm{n}}) = \frac{|\langle \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l), \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_k)\rangle|^2}{N_{\mathrm{u}}^2} \frac{|\langle \boldsymbol{d}(\tau_l), \boldsymbol{d}(\tau_k)\rangle|^2}{N^2}, \tag{10a}$$

$$C(\boldsymbol{g}_l^{\mathrm{r}}, \boldsymbol{g}_k^{\mathrm{r}}) = \frac{|\langle \boldsymbol{\nu}(\boldsymbol{\phi}_l), \boldsymbol{\nu}(\boldsymbol{\phi}_k)\rangle|^2}{\|\boldsymbol{\nu}(\boldsymbol{\phi}_l)\|^2 \|\boldsymbol{\nu}(\boldsymbol{\phi}_k)\|^2} \frac{|\langle \boldsymbol{d}(\bar{\tau}_l), \boldsymbol{d}(\bar{\tau}_k)\rangle|^2}{N^2}. \tag{10b}$$

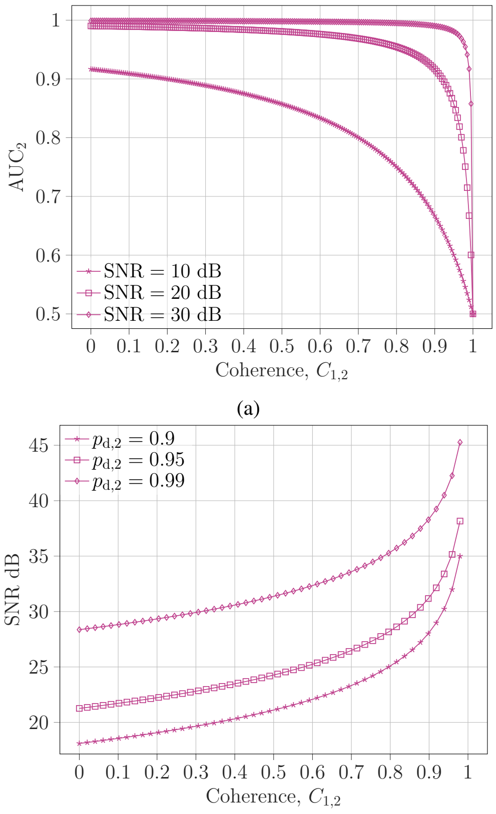
For the non-RIS signal, the coherence depends on the coherence between the delay response vectors and the coherence between the angle response vectors. On the other hand, for the RIS signal, the coherence depends on the coherence between the delay response vectors and the coherence between the RIS response vectors. This emphasizes the resolution capabilities of the RIS signal compared to the non-RIS signal: if the RIS response vectors decohere more quickly as the angle separation $\|\boldsymbol{\phi}_l - \boldsymbol{\phi}_k\|$ increases, than the array response vectors does as $\|\boldsymbol{\theta}_l - \boldsymbol{\theta}_k\|$ increases, then the RIS signal has improved angular resolution over the non-RIS signal.

The generalized coherence can be bounded by the coherences as

$$\mathscr{C} \geq \max\Big\{0, \frac{C_{1,3} + C_{2,3} - 2\sqrt{C_{1,3}C_{2,3}C_{1,2}}}{1 - C_{1,2}}\Big\}, \tag{11a}$$

$$\mathscr{C} \leq \min\Big\{1, \frac{C_{1,3} + C_{2,3} + 2\sqrt{C_{1,3}C_{2,3}C_{1,2}}}{1 - C_{1,2}}\Big\}. \tag{11b}$$

From these bounds, we observe that if all the coherences $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ are small, the generalized coherence will also be small, and vice versa. In this sense, the physical parameters influence the coherence and the generalized coherence in a similar way.



(a)

(b)

Fig. 2. Plot of (a) AUC against the coherence for varying receive SNR levels, and (b) the required receive SNR against coherence to realize false alarm and detection probability requirements, $p_{\mathrm{fa}} = 0.001$, $p_{\mathrm{d},2} = 0.9, 0.95, 0.99$.

In the case of two targets, the non-centrality parameter μ_2, and in turn AUC_2, depends only on the coherence and the receive SNR. We show this relation in Fig. 2a and note that given a receive SNR of 30 dB we can achieve an AUC that is nearly 1 if the coherence is less than 0.9, while for a receive SNR of 20 dB the coherence should nearly be 0 to achieve an AUC that is nearly 1. We can also shed light on the required receive SNR to realize requirements of false alarm probability and detection probability as a function of the coherence, see Fig. 2b: having a coherence of 0.1 rather than 0.9, we can realize the requirements that $p_{\mathrm{fa}} = 0.001$ and $p_{\mathrm{d},2} = 0.99$, with 10 dB less SNR. This clarifies the trade-off between non-RIS and RIS, considering that the RIS signal can have a lower coherence, but with a lower SNR.

We will see in the results that how quickly the RIS response vectors decohere depends both on the size of the RIS, N_{r}, and the number of RIS phase profiles, $\tilde{T}$. Additionally, we note that $\|\boldsymbol{g}_l^{\mathrm{r}}\|^2$, and thereby the receive SNR for the l-th path, scales proportionally with the number of RIS elements, but inverse proportionally to the size of the resolution region in the RIS angle parameter, i.e., $(\phi_{\mathrm{max}}^{\mathrm{az}} - \phi_{\mathrm{min}}^{\mathrm{az}})(\phi_{\mathrm{max}}^{\mathrm{el}} - \phi_{\mathrm{min}}^{\mathrm{el}})$.

B. Fisher analysis

Using Fisher analysis [35], we will characterize the upper bound on the performance of the channel parameter estimator and position estimator. This is also referred to as the CRLB and gives us the lower bound of the covariance among unbiased estimators. Specifically, the statement asserts that for an unbiased estimator $\hat{\boldsymbol{z}} = \boldsymbol{b}(\boldsymbol{y})$ of the parameter $\boldsymbol{z}$ defined by a function $\boldsymbol{b}$ of the data $\boldsymbol{y}$, the variance is lower-bounded as $\mathrm{Var}[\hat{\boldsymbol{z}}] \geq \boldsymbol{F}^{-1}(\boldsymbol{z})$, where $\boldsymbol{F}(\boldsymbol{z})$ is the Fisher information matrix (FIM) and the inequality is in the positive semi-definite sense. Estimators achieving this lower bound are called unbiased and efficient estimators. Such a property can be found, e.g., in the maximum likelihood estimator which is asymptotically unbiased and efficient. [41]

1) General case: We begin by deriving the FIM for the non-RIS and RIS signals. Conditioning on the fading parameters $\boldsymbol{\alpha}$ the channel is AWGN, and hence the conditional FIM for the non-RIS signal is [35]

$$[\boldsymbol{F}^{\mathrm{n}}]_{i,j} = \frac{4}{\sigma^2}\mathrm{Re}\left\{\left(\frac{\partial\boldsymbol{\mu}^{\mathrm{n}}}{\partial[\tilde{\boldsymbol{\eta}}^{\mathrm{n}}]_i}\right)^H \frac{\partial\boldsymbol{\mu}^{\mathrm{n}}}{\partial[\tilde{\boldsymbol{\eta}}^{\mathrm{n}}]_j}\right\} \tag{12a}$$

where $\tilde{\boldsymbol{\eta}}^{\mathrm{n}} = [\Re\{\boldsymbol{\alpha}\}^\top, \Im\{\boldsymbol{\alpha}\}^\top, (\boldsymbol{\eta}_1^{\mathrm{n}})^\top, \dots, (\boldsymbol{\eta}_L^{\mathrm{n}})^\top]^\top$, and $\boldsymbol{\mu}^{\mathrm{n}} = \boldsymbol{G}^{\mathrm{n}}\boldsymbol{\alpha}$. Conditioning on the fading parameters $\bar{\boldsymbol{\alpha}} = [\bar{\alpha}_1, \dots, \bar{\alpha}_L]^\top$, the conditional FIM for the RIS signal is

$$[\boldsymbol{F}^{\mathrm{r}}]_{i,j} = \frac{4}{\sigma^2}\mathrm{Re}\left\{\left(\frac{\partial\boldsymbol{\mu}^{\mathrm{r}}}{\partial[\tilde{\boldsymbol{\eta}}^{\mathrm{r}}]_i}\right)^H \frac{\partial\boldsymbol{\mu}^{\mathrm{r}}}{\partial[\tilde{\boldsymbol{\eta}}^{\mathrm{r}}]_j}\right\} \tag{12b}$$

where $\tilde{\boldsymbol{\eta}}^{\mathrm{r}} = [\Re\{\bar{\boldsymbol{\alpha}}\}^\top, \Im\{\bar{\boldsymbol{\alpha}}\}^\top, (\boldsymbol{\eta}_1^{\mathrm{r}})^\top, \dots, (\boldsymbol{\eta}_L^{\mathrm{r}})^\top]^\top$, and $\boldsymbol{\mu}^{\mathrm{r}} = \boldsymbol{G}^{\mathrm{r}}\bar{\boldsymbol{\alpha}}$. The first order derivatives required in the FIMs can be computed straightforwardly, see Appendix C.

With the FIM at hand, we can define the DEB and AEB for the relevant parameters:

$$\mathrm{DEB}_l^{\mathrm{n}} = \sqrt{[(\boldsymbol{F}^{\mathrm{n}})^{-1}]_{2L+l,2L+l}}, \tag{13a}$$

$$\mathrm{DEB}_l^{\mathrm{r}} = \sqrt{[(\boldsymbol{F}^{\mathrm{r}})^{-1}]_{4L+l,4L+l}}, \tag{13b}$$

$$\mathrm{AEB}_l^{\mathrm{n}} = \sqrt{[(\boldsymbol{F}^{\mathrm{n}})^{-1}]_{3L+l,3L+l} + [(\boldsymbol{F}^{\mathrm{n}})^{-1}]_{4L+l,4L+l}}, \tag{13c}$$

$$\mathrm{AEB}_l^{\mathrm{r}} = \sqrt{[(\boldsymbol{F}^{\mathrm{r}})^{-1}]_{5L+l,5L+l} + [(\boldsymbol{F}^{\mathrm{r}})^{-1}]_{6L+l,6L+l}}. \tag{13d}$$

We can also characterize the position estimation capabilities with the SB and DB signals through a coordinate system transformation. We define $[\boldsymbol{T}^{\mathrm{n}}]_{i,j} = \frac{\partial[\boldsymbol{\eta}^{\mathrm{n}}]_i}{\partial\boldsymbol{x}_j}$ and $[\boldsymbol{T}^{\mathrm{r}}]_{i,j} = \frac{\partial[\boldsymbol{\eta}^{\mathrm{r}}]_i}{\partial\boldsymbol{x}_j}$ as the Jacobian matrices of the transformation from position to channel parameters. The partial derivatives are provided in (30). Then, the FIMs in the Euclidean coordinates are $\boldsymbol{F}_{\mathrm{euc}}^{\mathrm{n}} = (\boldsymbol{T}^{\mathrm{n}})^\top([(\boldsymbol{F}^{\mathrm{n}})^{-1}]_{2L:,2L:})^{-1}\boldsymbol{T}^{\mathrm{n}}$ and $\boldsymbol{F}_{\mathrm{euc}}^{\mathrm{r}} = (\boldsymbol{T}^{\mathrm{r}})^\top([(\boldsymbol{F}^{\mathrm{r}})^{-1}]_{4L:,4L:})^{-1}\boldsymbol{T}^{\mathrm{r}}$.[4] We can define the PEB as

$$\mathrm{PEB}_l^{\mathrm{n}} = \sqrt{\mathrm{Tr}([(\boldsymbol{F}_{\mathrm{euc}}^{\mathrm{n}})^{-1}]_{1+3(l-1):3l,1+3(l-1):3l})}, \tag{13e}$$

$$\mathrm{PEB}_l^{\mathrm{r}} = \sqrt{\mathrm{Tr}([(\boldsymbol{F}_{\mathrm{euc}}^{\mathrm{r}})^{-1}]_{1+3(l-1):3l,1+3(l-1):3l})}. \tag{13f}$$

If we assume that the data association between the SB and DB estimator results is known, then the joint FIM can be formulated as $\boldsymbol{F} = \mathrm{bdiag}(\boldsymbol{F}^{\mathrm{n}}, \boldsymbol{F}^{\mathrm{r}})$ where bdiag denotes block-diagonal matrix, and as a consequence, in Euclidean coordinates the non-RIS and RIS Fisher information is simply aggregated: $\boldsymbol{F}_{\mathrm{euc}} = \boldsymbol{F}_{\mathrm{euc}}^{\mathrm{n}} + \boldsymbol{F}_{\mathrm{euc}}^{\mathrm{r}}$. The PEB for the joint sensing is

$$\mathrm{PEB}_l = \sqrt{\mathrm{Tr}([\boldsymbol{F}_{\mathrm{euc}}^{-1}]_{1+3(l-1):3l,1+3(l-1):3l})}. \tag{13g}$$

We can compute explicit expressions for the equivalent Fisher information for individual parameters. Full derivations are cumbersome, however, we can gain insight into the full problem by considering a simplified but sufficiently complicated special case.

2) Case study with two targets: Consider the special case of two targets, i.e., $L = 2$. Beginning with the non-RIS signal, we assume for simplicity that θ_1^{el}, θ_2^{el}, τ_1, and τ_2 are known, such that the unknowns are $\Re\{\alpha_1\}$, $\Re\{\alpha_2\}$, $\Im\{\alpha_1\}$, $\Im\{\alpha_2\}$, θ_1^{az}, and θ_2^{az}. We can now focus on the equivalent Fisher information for θ_1^{az} which provides the necessary understanding to interpret the information loss due to the unknown channel coefficients and the interfering second target:

$$F(\mathrm{n}, \theta_1^{\mathrm{az}}) := \frac{4}{\sigma^2}\left(|\alpha_1|^2\|\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\|^2\left(1 - \mathrm{IL}_{\boldsymbol{\alpha}}^{\mathrm{n}}\right) - \mathrm{IL}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\right), \tag{14}$$

where $\mathrm{IL}_{\boldsymbol{\alpha}}^{\mathrm{n}}$ is the information loss due to the lack of knowledge about the channel coefficients

$$\mathrm{IL}_{\boldsymbol{\alpha}}^{\mathrm{n}} := \frac{\|\boldsymbol{x}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\|^2}{\|\boldsymbol{g}_1^{\mathrm{n}}\|^2\|\boldsymbol{g}_2^{\mathrm{n}}\|^2\|\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\|^2(1 - C_{1,2}^{\mathrm{n}})} = \mathscr{C}(\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_2^{\mathrm{n}}, \boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}),$$

and $\mathrm{IL}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}$ is the information loss due to the lack of knowledge of the azimuth angle of the second target

$$\mathrm{IL}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}} := \frac{\Re\left\{\alpha_1^*\alpha_2\left(\langle\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\rangle - \frac{\langle\boldsymbol{x}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}, \boldsymbol{x}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\rangle}{\|\boldsymbol{g}_1^{\mathrm{n}}\|^2\|\boldsymbol{g}_2^{\mathrm{n}}\|^2(1-C_{1,2}^{\mathrm{n}})}\right)\right\}^2}{|\alpha_2|^2\|\boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\|^2\left(1 - \mathscr{C}(\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_2^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}})\right)},$$

where we have defined $\boldsymbol{x}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}} := \boldsymbol{g}_2^{\mathrm{n}}\langle\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\rangle - \boldsymbol{g}_1^{\mathrm{n}}\langle\boldsymbol{g}_2^{\mathrm{n}}, \boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\rangle$, $\boldsymbol{x}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}} := \boldsymbol{g}_2^{\mathrm{n}}\langle\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\rangle - \boldsymbol{g}_1^{\mathrm{n}}\langle\boldsymbol{g}_2^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}}\rangle$, $\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}} := \frac{\partial\boldsymbol{g}_1^{\mathrm{n}}}{\partial\theta_1^{\mathrm{az}}}$, and $\boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}} := \frac{\partial\boldsymbol{g}_2^{\mathrm{n}}}{\partial\theta_2^{\mathrm{az}}}$ (see Appendix C).

We observe that if $\theta_1^{\mathrm{az}} = \theta_2^{\mathrm{az}}$, then $F(\mathrm{n}, \theta_1^{\mathrm{az}}) = 0$. On the other hand, if $C(\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_2^{\mathrm{n}})$, $C(\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}})$, $C(\boldsymbol{g}_2^{\mathrm{n}}, \boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}})$, and $C(\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}, \boldsymbol{g}_{\theta_2^{\mathrm{az}}}^{\mathrm{n}})$ are all equal to zero, then the equivalent Fisher information is maximized as $F(\mathrm{n}, \theta_1^{\mathrm{az}}) = \frac{4|\alpha_1|^2}{\sigma^2}\|\boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}}\|^2\big(1 - C(\boldsymbol{g}_1^{\mathrm{n}}, \boldsymbol{g}_{\theta_1^{\mathrm{az}}}^{\mathrm{n}})\big)$, i.e., there is no loss of information of the first target due to the second target. Inspecting further these inner products, it is derived that this condition comes down to $C(\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_1), \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_2))$, $C(\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_1), \frac{\partial\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_2)}{\partial\theta_2^{\mathrm{az}}})$, $C(\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_2), \frac{\partial\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_1)}{\partial\theta_1^{\mathrm{az}}})$, and $C(\frac{\partial\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_1)}{\partial\theta_1^{\mathrm{az}}}, \frac{\partial\boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_2)}{\partial\theta_2^{\mathrm{az}}})$ being equal to zero. Generally speaking these coherences decrease as $|\theta_2^{\mathrm{az}} - \theta_1^{\mathrm{az}}|$ increases, and the rate at which this decrease occurs depends on $N_{\mathrm{u}}^{\mathrm{az}}$ [38].

Considering now the RIS signal and assuming that θ_1^{az}, θ_2^{az}, θ_1^{el}, θ_2^{el}, τ_1, τ_2, ϕ_1^{el}, and ϕ_2^{el} are known, and the unknowns are $\Re\{\alpha_1\}$, $\Re\{\alpha_2\}$, $\Im\{\alpha_1\}$, $\Im\{\alpha_2\}$, ϕ_1^{az}, and ϕ_2^{az}. Reusing again (14), we find that the equivalent Fisher information for ϕ_1^{az} with the RIS signal is $F(\mathrm{r}, \phi_1^{\mathrm{az}})$, replacing in (14) all superscript n by r, all θ_1^{az} by ϕ_1^{az}, and all α_1, α_2 by $\bar{\alpha}_1, \bar{\alpha}_2$. As before, we observe that when $\phi_1^{\mathrm{az}} = \phi_2^{\mathrm{az}}$, then $F(\mathrm{r}, \phi_1^{\mathrm{az}}) = 0$, but if $C(\boldsymbol{\nu}(\boldsymbol{\phi}_1), \boldsymbol{\nu}(\boldsymbol{\phi}_2))$, $C(\boldsymbol{\nu}(\boldsymbol{\phi}_1), \frac{\partial\boldsymbol{\nu}(\boldsymbol{\phi}_2)}{\partial\phi_2^{\mathrm{az}}})$, $C(\boldsymbol{\nu}(\boldsymbol{\phi}_2), \frac{\partial\boldsymbol{\nu}(\boldsymbol{\phi}_1)}{\partial\phi_1^{\mathrm{az}}})$, and $C(\frac{\partial\boldsymbol{\nu}(\boldsymbol{\phi}_1)}{\partial\phi_1^{\mathrm{az}}}, \frac{\partial\boldsymbol{\nu}(\boldsymbol{\phi}_2)}{\partial\phi_2^{\mathrm{az}}})$ are all equal to zero, then the Fisher information is maximized.

---

[4] We treat the $\boldsymbol{\theta}_l$ angles as nuisance parameters with the RIS signal. This is justified by observing that using this signal to estimate the angles of departure at the UE when we use just a single precoder will result in a highly inaccurate estimate. Further justification follows from observing through numerical evaluations that the Fisher information of the $\boldsymbol{\theta}_l$ parameters is vanishingly small.

Remark 3. *We have seen that the coherence plays a central role for both the CRLB and the detection probability. For the detection probability, it is the coherences between vectors* $\boldsymbol{g}_l$*,* $l = 1, \ldots, L$*, which appear, while for the CRLB it is also the coherences between vectors* $\boldsymbol{g}_l$ *and their partial derivatives.*

IV. CHANNEL AND POSITION ESTIMATION

With the formulations (1a) and (1b), we can in parallel apply channel estimation techniques such as expectation-maximization (EM) maximum likelihood (ML) [42], compressive sensing [43], [44], or beamspace multiple signal classification (MUSIC) [45], followed by a data association step, and subsequent position estimation. We begin by briefly recalling the well-studied channel parameter estimation technique OMP

A. Channel parameter estimation

OMP: Consider an observed signal $\boldsymbol{y} = \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, which can be either the non-RIS or RIS signal (1a) and (1b). To apply the OMP method, we construct a redundant dictionary $\boldsymbol{\Psi} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_M]$ where M is a redundancy parameter, and $\boldsymbol{g}_i = \boldsymbol{g}^{\mathrm{n}}(\boldsymbol{\eta}_i^{\mathrm{n}})$ with $\boldsymbol{\eta}_i^{\mathrm{n}}$ in the resolution region for the non-RIS signal, while $\boldsymbol{g}_i = \boldsymbol{g}^{\mathrm{r}}(\boldsymbol{\eta}_i^{\mathrm{r}})$ with $\boldsymbol{\eta}_i^{\mathrm{r}}$ in the resolution region for the RIS signal. Then, we search for a sparse representation $\boldsymbol{v} \in \mathbb{C}^M$ such that $\|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{v}\|$ is minimized by initializing $\boldsymbol{v}$ as the zero vector and then iteratively updating the support set. Specifically, the iterative update is found by maximizing the OMP objective $|\langle \boldsymbol{g}_i, \boldsymbol{r} \rangle|$ with respect to index i, where $\boldsymbol{r}$ is the residual. The algorithm is summarized below [44]:

- Let $\boldsymbol{r} = \boldsymbol{y}$ denote the residual, initialize $\boldsymbol{v} = \boldsymbol{0}$, and let $\Lambda = \emptyset$ denote the support set of $\boldsymbol{v}$.
- While $\|\boldsymbol{r}\| \geq \mathrm{residual}_{\mathrm{th}}$ do:
    1) $I \leftarrow \arg\max_{i=1,\ldots,M} |\langle \boldsymbol{g}_i, \boldsymbol{r} \rangle|$.
    2) $\Lambda \leftarrow \Lambda \cup \{I\}$.
    3) $[\boldsymbol{v}]_\Lambda \leftarrow [\boldsymbol{\Psi}]_\Lambda^\dagger \boldsymbol{y}$. [5]
    4) $\boldsymbol{r} \leftarrow \boldsymbol{y} - [\boldsymbol{\Psi}]_\Lambda [\boldsymbol{v}]_\Lambda$.

The support set Λ specifies a set of estimated channel parameters $\{\hat{\boldsymbol{\eta}}_i^{\mathrm{n}}\}_{i\in\Lambda}$ or $\{\hat{\boldsymbol{\eta}}_i^{\mathrm{r}}\}_{i\in\Lambda}$ in case of non-RIS or RIS, respectively. The threshold parameter $\mathrm{residual}_{\mathrm{th}} > 0$ specifies the trade-off between detection probability and false alarm probability, and can be chosen based on the noise variance, specifically $\mathrm{residual}_{\mathrm{th}}^2$ should be larger than $\sigma^2 N\tilde{T}N_{\mathrm{u}}/2$ such that the algorithm terminates when the residual only contains noise. We let $\mathrm{residual}_{\mathrm{th}}^{\mathrm{n}}$ and $\mathrm{residual}_{\mathrm{th}}^{\mathrm{r}}$ denote the thresholds for the non-RIS and RIS signals, respectively.

B. Data association

Once we have estimated the non-RIS and RIS channel parameters, the data association task is to find pairs of non-RIS and RIS channel parameters which associate with the same target. We can compare the measurements in the domain of the AOA at the RIS, by noting that the non-RIS parameters uniquely identify a position which subsequently uniquely identifies the RIS channel parameters. Specifically, from the non-RIS parameter estimate $\hat{\boldsymbol{\eta}}_i^{\mathrm{n}}$ we can uniquely find the estimate of the position of the target $\hat{\boldsymbol{c}}_i^{\mathrm{n}} = \boldsymbol{h}_{\mathrm{n}}^{-1}(\hat{\boldsymbol{\eta}}_i^{\mathrm{n}})$ from which we can compute the AOA at the RIS as $\hat{\boldsymbol{\phi}}_i^{\mathrm{n}} = \boldsymbol{\phi}(\hat{\boldsymbol{c}}_i^{\mathrm{n}})$, cf. Appendix A.

With this in mind, to solve the data association problem, we define the cost matrix

$$\begin{aligned}\boldsymbol{C}_{i,k} =& \min\left(([\hat{\boldsymbol{\phi}}_i^{\mathrm{n}}]_1 - [\hat{\boldsymbol{\phi}}_k]_1)^2, (|[\hat{\boldsymbol{\phi}}_i^{\mathrm{n}}]_1 - [\hat{\boldsymbol{\phi}}_k]_1| - 2\pi)^2\right) \\ &+ \min\left(([\hat{\boldsymbol{\phi}}_i^{\mathrm{n}}]_2 - [\hat{\boldsymbol{\phi}}_k]_2)^2, (|[\hat{\boldsymbol{\phi}}_i^{\mathrm{n}}]_2 - [\hat{\boldsymbol{\phi}}_k]_2| - \pi)^2\right),\end{aligned} \tag{15}$$

where $i = 1, \ldots, \hat{L}_{\mathrm{n}}$ and $k = 1, \ldots, \hat{L}_{\mathrm{r}}$, $\hat{L}_{\mathrm{n}}$ and $\hat{L}_{\mathrm{r}}$ are the estimated number of targets with the non-RIS and RIS signals, respectively, and $\hat{\boldsymbol{\phi}}_k$ are the estimated AOAs at the RIS with the RIS signal. Let $\boldsymbol{X} \in \{0,1\}^{\hat{L}_{\mathrm{n}} \times \hat{L}_{\mathrm{r}}}$ be the boolean assignment matrix such that $\boldsymbol{X}_{i,k} = 1$ if the i-th non-RIS estimate is associated with the k-th RIS estimate and $\boldsymbol{X}_{i,k} = 0$ otherwise. When $\hat{L}_{\mathrm{r}} \geq \hat{L}_{\mathrm{n}}$, to solve the data association problem is to solve the optimization problem

$$\begin{aligned}\text{minimize} \quad & \mathrm{Tr}(\boldsymbol{X}^\top \boldsymbol{C}) \\ \text{subject to} \quad & \boldsymbol{X}_{i,k} \in \{0,1\},\ \forall i,k, \\ & \sum_i \boldsymbol{X}_{i,k} \leq 1,\ \forall k, \quad \sum_k \boldsymbol{X}_{i,k} = 1,\ \forall i.\end{aligned} \tag{16}$$

The constraints mean that each RIS detection is associated with only one non-RIS detection and vice versa. Now, if more detection are made with the RIS signal (resp. non-RIS signal), the targets that are not associated with a non-RIS detection (resp. RIS detection), are kept and can be used for subsequent positioning, although with less precision. The estimated number of targets is $\hat{L} = \max(\hat{L}_{\mathrm{n}}, \hat{L}_{\mathrm{r}})$.

C. Position estimation

By the definition of the channel parameters, as detailed in Appendix A, the mapping $\boldsymbol{h}$ from Euclidean space to the channel parameter space is known. We formulate the weighted non-linear least squares optimization problem as

$$\hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{c}\in\mathcal{D}} \quad \frac{1}{2}(\boldsymbol{h}(\boldsymbol{c}) - \hat{\boldsymbol{\eta}})^\top \boldsymbol{F}_{\mathrm{efim}}(\boldsymbol{h}(\boldsymbol{c}) - \hat{\boldsymbol{\eta}}), \tag{17}$$

where $\boldsymbol{F}_{\mathrm{efim}} = \mathrm{bdiag}(([(\boldsymbol{F}^{\mathrm{n}})^{-1}]_{2L:,2L:})^{-1}, ([(\boldsymbol{F}^{\mathrm{r}})^{-1}]_{4L:,4L:})^{-1})$ is the equivalent Fisher information matrix (EFIM) evaluated at $\hat{\boldsymbol{\eta}} = [\hat{\tau}, \hat{\bar{\tau}}, \hat{\boldsymbol{\theta}}^\top, \hat{\boldsymbol{\phi}}^\top]^\top$. The partial derivatives required in the Jacobian is provided in (30), and any gradient-based minimization method can solve this numerically.

V. NUMERICAL EXPERIMENTS

In this section, the high resolution sensing capabilities of RIS is evaluated through numerical experiments. First we specify the setting, followed by a presentation of the key performance indicators. Then, the working principle is illustrated for an exemplary scenario, followed by a study of the coherence and the theoretical bounds using the results from Section III. Finally, the performance of the sensing algorithm outlined in Section IV with and without the RIS is analyzed.

A. Settings

With the numerical experiments we want to understand when using the RIS can improve the sensing performance,

---

[5] The notation $[\boldsymbol{\Psi}]_\Lambda$ refers to the sub-matrix of $\boldsymbol{\Psi}$ containing the columns with indices in the index set Λ. Similarly, $[\boldsymbol{v}]_\Lambda$ is the sub-vector of $\boldsymbol{v}$ containing the elements with indices in the index set Λ.

and how this can be attributed to a beneficial trade-off between coherence and SNR achieved by the RIS signal, as compared to the non-RIS signal. Hence, the settings for the numerical experiments are chosen to uncover a scenario where a UE has insufficient resolution to separate targets in an SP cluster without RIS-assistance. Meanwhile, we are not interested in opportunistically positioning the RIS as was done in [24], [25], nor are we interested in the worst case where the RIS is positioned behind the SP cluster. Instead, we consider that the RIS is at an intermediate point between the UE and the SP [26].

Accordingly, we let the bandwidth be only $W = 9$ MHz and the number of UE antennas be just $N_\mathrm{u} = 2 \times 2$. The transmission energy is set to $E_\mathrm{s} W T N_\mathrm{u} = 65$ dBm, E_s is the energy per subcarrier, cf. Appendix A, with a noise power spectral density of $N_0 = -166$ dBm, $\sigma^2 = N_0 W$, consistent with the specifications used in [26]. The size of the RIS is set to $N_\mathrm{r} = 35 \times 35$ with number of OFDM symbols set to $T = 50$ which results in $\tilde{T} = 25$ different RIS configurations in the beam sweep, unless otherwise specified. The carrier frequency is chosen to lie in frequency range 3, specifically $f_c = 15$ GHz, resulting in a wavelength of $\lambda = 2$ cm. The UE is placed at $\boldsymbol{p} = [0, 0, 0]^\top$ m with the RIS at $\boldsymbol{p}_\mathrm{r} = [3, 5, 6]^\top$ m, i.e., 8.4 m away from the UE. Settings for the numerical experiments are summarized in Table I.

We will consider an SP cluster of up to three targets with channel parameters $\boldsymbol{\eta}_1^\mathrm{n} = [\tau_1, \theta_1^\mathrm{az}, \theta_1^\mathrm{el}] = [120, 0.7, 0.8]$, $\boldsymbol{\eta}_2^\mathrm{n} = [\tau_1, \theta_1^\mathrm{az} + \Delta, \theta_1^\mathrm{el} - \Delta]$, and $\boldsymbol{\eta}_3^\mathrm{n} = [\tau_1, \theta_1^\mathrm{az} - \Delta, \theta_1^\mathrm{el} + \Delta]$ where the delay is expressed in nanoseconds (ns), the angles are expressed in radians (rad), and $\Delta > 0$ rad is the spacing between the targets in the AOA at the UE. This setup is used to evaluate numerically how performance varies depending on the distance between the targets. The expected radar cross sections are $\sigma^2_{\mathrm{rcs},1} = 50$ m^2, $\sigma^2_{\mathrm{rcs},2} = 5.0$ m^2, and $\sigma^2_{\mathrm{rcs},3} = 0.5$ m^2, cf. Appendix A. Hence, on average, the signal received from target one is the strongest with significantly weaker signals from targets two and three.

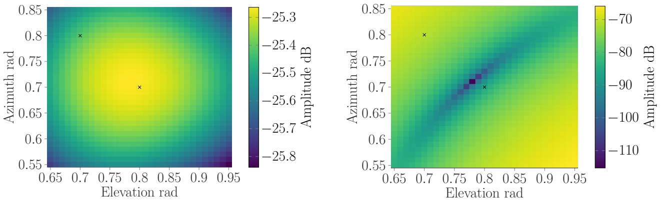
B. Performance metrics and baseline

We evaluate the performance with the empirical expected AUC and the GOSPA metric, see [46] for details. For the GOSPA metric, we use the 2-norm as the pairwise distance, let the maximum allowable localization error be fixed as 5 m, and choose the cardinality penalty normalization of 2 as recommended [46]. We use GOSPA$^\mathrm{n}$, GOSPA$^\mathrm{r}$, and GOSPA$^\mathrm{joint}$ to denote the GOSPA metric using the non-RIS signal, the RIS signal, and the combination of the two, respectively.

We consider, the probability of detecting $l \leq L$ targets conditioned on the channel parameters defined as $p_{\mathrm{d},l} = \mathbb{E}[\mathbb{1}[\hat{L} \geq l]]$, with corresponding probability of false alarm $p_\mathrm{fa} = \mathbb{E}[\mathbb{1}[\hat{L} > L]]$, where $\mathbb{1}[x]$ is the indicator function equal to one when the condition x is true and equal to zero when x is false. The empirical detection probability, denoted $\hat{p}_{\mathrm{d},l}$, and empirical false alarm probability, denoted $\hat{p}_\mathrm{fa}$, replaces the expectation operators with averages over Monte Carlo simulations. The empirical expected AUC, denoted $\widehat{\mathrm{AUC}}_l$ is found by integrating $\hat{p}_{\mathrm{d},l}$ considered a function of $\hat{p}_\mathrm{fa}$. We add superscripts n, r, and joint to denote the empirical results using the non-RIS, RIS, and the combination of the two, respectively.

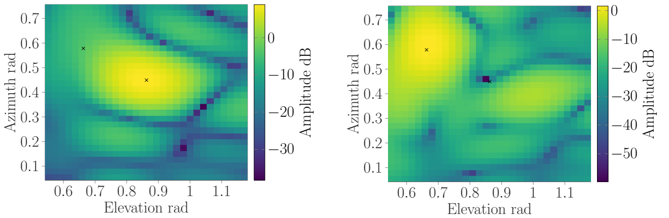
TABLE I
Scenario & hyperparameter settings

Parameter	Value	Description
$\boldsymbol{s}$	[0, 0, 0, 0, 0, 0]	UE state
$\boldsymbol{s}_\mathrm{r}$	[3, 5, 6, 0, 0, 0]	RIS state
f_c	15 GHz	Carrier frequency
λ	2 cm	Wavelength
W	9 MHz	Bandwidth
Δ_f	120 kHz	Subcarrier spacing
N	75	Number of subcarriers
$E_\mathrm{s} W T N_\mathrm{u}$	65 dBm	Total transmission energy
N_0	-166 dBm	Noise power spectral density
N_r	1225 (35×35)	RIS array size
N_u	4 (2×2)	UE array size
T	50	OFDM symbols
ref_th	0.3	Reference threshold



(a) Non-RIS iteration 1.

(b) Non-RIS iteration 2.



(c) RIS iteration 1.

(d) RIS iteration 2.

Fig. 3. Non-RIS and RIS signal OMP objective, $|\langle \boldsymbol{g}_i, \boldsymbol{r} \rangle|$. The black crosses indicate the true positions of the SPs.

As a baseline method, we will consider the well-known super-resolution method MUSIC with the non-RIS signal [47], [48]. The number of targets is estimated based on a likelihood ratio test on the eigenvalues of the data covariance matrix: we test the hypotheses that the $N - i$ smallest eigenvalues are equal for $i > 0$ and where N is the number of eigenvalues, and estimate the number of targets as the smallest i such that we accept the hypothesis under a specified confidence level set to 10^{-5} [49]. To estimate the data covariance matrix, we average across the OFDM symbols and also perform smoothing along the frequency axis using a frequency sub-array of size 45 [50].

C. Working principle

To illustrate the working principle, we include only the first two targets with $\Delta = 0.1$ rad. This corresponds to SPs located at $[9.87, 8.31, 12.53]$ m and $[8.07, 8.31, 13.76]$ m meaning that the distance between the two SPs is 2.17 m, and that the distances to the UE and RIS are approx. 18 m and 10 m, respectively. Meanwhile, the Fraunhofer distance of the RIS when $N_r = 35 \times 35$ is 6.1 m so that we may consider the SPs to be in the far-field. In this situation, the channel parameters in the view of the UE varies only slightly along the azimuth

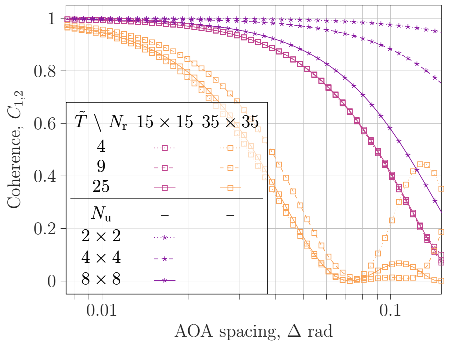


Fig. 4. Plot of coherence against the spacing of the targets in the AOA at the UE for varying RIS specifications, $\tilde{T}$ and N_{r}, and UE specifications, N_{u}.

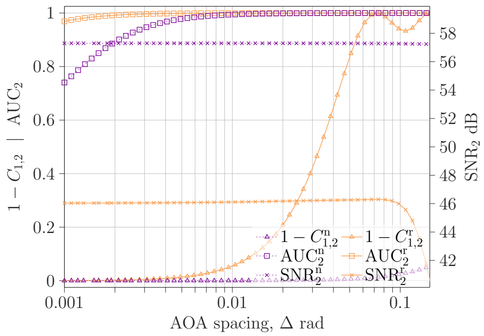


Fig. 5. Trade-off analysis between SNR and coherence of RIS and non-RIS signals, plotted together with the AUC. The left y-axis is used for the coherence and the AUC, while the right y-axis is used for the SNR.

and elevation angles, and it will be difficult for the UE alone to detect both SPs.

In Fig. 3, the OMP objective, $|\langle \boldsymbol{g}_i, \boldsymbol{r} \rangle|$, for the non-RIS and RIS signals are shown[6]. From Fig. 3b we see that with the non-RIS signal we cannot clearly resolve the second SP as we cannot see a distinguishable peaks in the second iteration, however, using the RIS signal we see that the angles are well-separated in the view of the RIS, and we notice a distinguishable peak in Fig. 3d.

D. Coherence and detection bound

To use the results on the expected detection probability Corollary 1, we require that $\mathcal{R}_o$ is fixed. Let $\mathcal{R}_o = [112, 128] \times [0.55, 0.85] \times [0.65, 0.95]$ ns $\cdot$ rad^2 be the fixed resolution region. We consider in this section a variety of physical parameter settings. For the RIS we use either a small aperture of $N_{\mathrm{r}} = 15 \times 15$ or the larger aperture of $N_{\mathrm{r}} = 35 \times 35$, and moreover we consider different numbers of RIS phase profiles $\tilde{T} \in \{4, 9, 25\}$, in all cases using $N_{\mathrm{u}} = 2 \times 2$. Without the RIS we consider having the small array of size $N_{\mathrm{u}} = 2 \times 2$ as well as two larger arrays of sizes $N_{\mathrm{u}} = 4 \times 4$ and $N_{\mathrm{u}} = 8 \times 8$.

1) Two targets: Consider initially that only targets one and two are present. As we observed earlier, in this case the detection probability depends only on the false alarm probability, the receive SNR, and the coherence, $C_{1,2}$. In Fig. 4, we show the coherence of the RIS and non-RIS signals, as a function of Δ with varying number of RIS elements, RIS phase profiles, and number of UE antennas: we see that the size of the RIS is crucial for fast decoherence while the importance of the number of RIS phase profiles is less pronounced. Moreover, we generally observe a lower coherence with the RIS signal than with the non-RIS signal for all tested specifications, revealing the feasible detection performance gains by using the RIS. The trade-off between SNR and coherence between RIS and non-RIS signals, is presented in Fig. 5 when $N_{\mathrm{u}} = 2 \times 2$, $N_{\mathrm{r}} = 35 \times 35$, and $\tilde{T} = 25$. This result reveals that the gains in terms coherence of the RIS signal indeed results in improved detection performance even though the RIS signal has a lower SNR due to double-bounce attenuation.

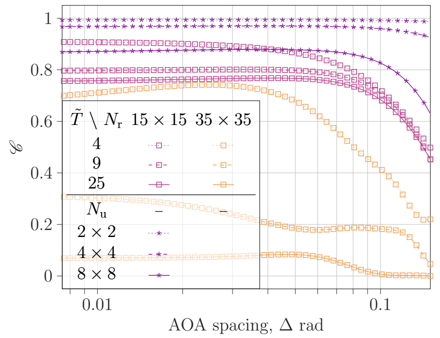


Fig. 6. Plot of $\mathscr{C}$ against the spacing between the targets in the AOA at the UE.

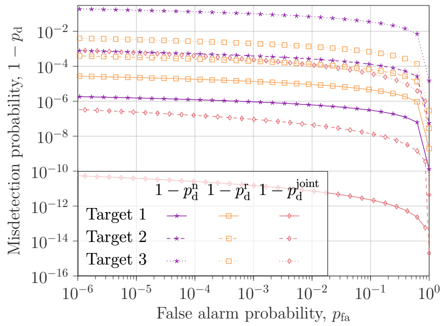


Fig. 7. Expected misdetection probability against the false alarm probability on a log-log scale.

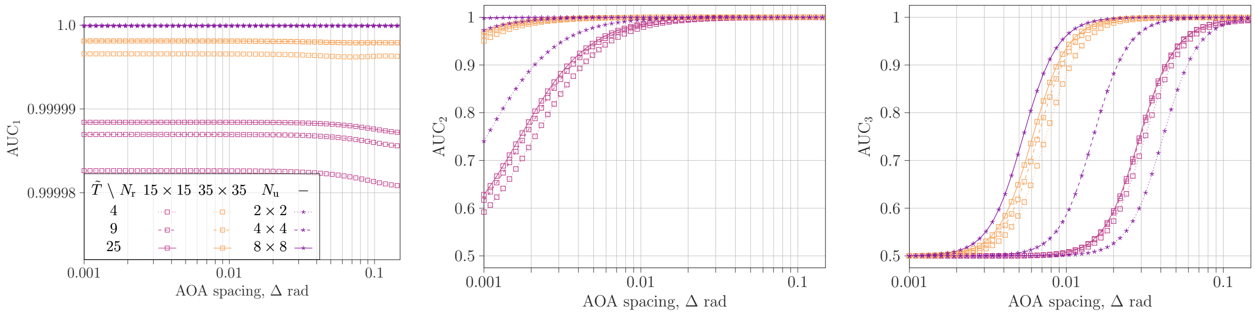
2) Three targets: Consider now the SP cluster of three targets. The detection probability in this case not only depends on the coherences between the targets, but also $\mathscr{C}$, cf. Definition 2. We plot in Fig. 6 $\mathscr{C}$ as a function of Δ: as was the case for $C_{1,2}$ in Fig. 4, the RIS signal with $N_{\mathrm{r}} = 35 \times 35$ shows significant improvements in decoherence as compared to the non-RIS signal, and this is particularly pronounced when $\tilde{T} = 25$.

Using the results for the expected detection probability (9a) to (9c), including (5), we show in Fig. 7 the expected misdetection probability against the false alarm probability on a log-log scale for each of the three targets when $\Delta = 0.1$ rad. This shows that for these settings, the RIS signal has improved detection capabilities compared to the non-RIS signal for both the second and third targets, although the non-RIS signal has better detection capabilities for the first target, due to the higher SNR. Moreover, we see that joint detection outperforms individual detection using either RIS signal or non-RIS signal for all three targets.

Comparing the detection performance depending on the spacing between the targets, Δ, we show in Fig. 8 the

---

[6] For visualization purposes, we evaluate the OMP objective in the true delay, and display the OMP objective only as a function of azimuth and elevation.

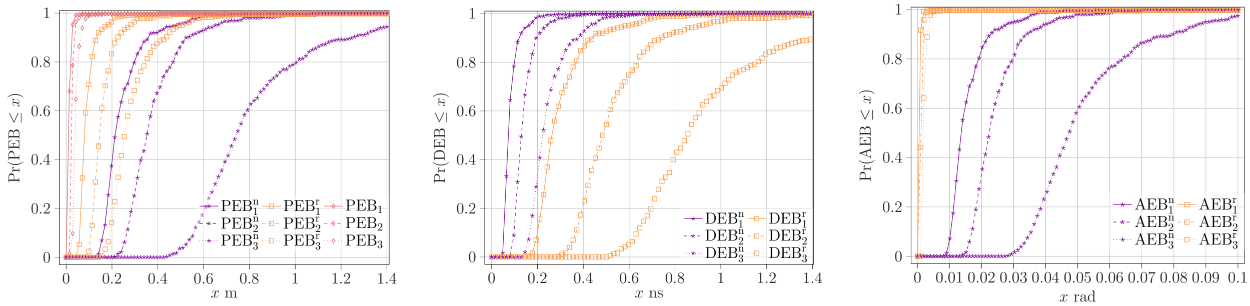


(a) Target 1.

(b) Target 2.

(c) Target 3.

Fig. 8. Plot of the expected AUC dependent on the spacing between the targets in the AOA at the UE for varying number of RIS elements, N_r, RIS phase profiles, $\tilde{T}$, and number of UE antennas, N_u. The legend is shared for all the plots.



(a) PEB.

(b) DEB.

(c) AEB.

Fig. 9. CDFs of the PEB, DEB, and AEB for the three targets with either the non-RIS, RIS, or combined. The spacing between the targets is $\Delta = 0.1$.

expected AUC for varying specifications regarding number of RIS elements, N_r, and the number of RIS phase profiles, $\tilde{T}$, as well as the UE specifications, herein the number of UE antennas. It is observed that we can achieve improved detection performance with the RIS signal over the non-RIS signal, in cases where N_u is small while N_r is relatively large: with the RIS signal, when $\tilde{T} = 25$ and $N_\mathrm{r} = 35 \times 35$, an AUC of 1 is approached at $\Delta = 0.02$ rad, while for the non-RIS signal, with $N_\mathrm{u} = 2 \times 2$, this occurs at $\Delta = 0.1$ rad. Additionally, we note that even with the smaller RIS, a higher AUC is achieved in Fig. 8c compared to the non-RIS case with $N_\mathrm{u} = 2 \times 2$.

E. Positioning bound

With the same setup as in the previous section, we show in Fig. 9 the empirical CDFs of the PEB, DEB, and AEB, derived in Section III-B, for each of the three targets when $\Delta = 0.1$ rad and considering the non-RIS, RIS, and combined signals. From Fig. 9a we see that combining the non-RIS and RIS signals in all cases outperforms the individual ones in terms of PEB, and we also see that the RIS signal performs better than the non-RIS signal. Going to the DEBs and AEBs in Figs. 9b and 9c, we notice that the non-RIS delay is better estimated than the RIS delay which is due to better SNR conditions for the non-RIS signal, but on the other hand the angle is better estimated by the RIS than the non-RIS signal due to the higher angular resolution provided by the RIS.

F. Sensing performance

In Figs. 10a and 10b, we show the expected AUC for the sensing algorithm with the OMP channel estimation method, as well as the MUSIC algorithm using the non-RIS signal, in comparison to the derived bound as a function of the AOA spacing, Δ rad. We observe that while we do not achieve the bound, the relation between the RIS, non-RIS, and joint methods follow a similar trend as with the bound. Moreover, as in the bound, we observe that the AUC with the RIS signal dominates that of the non-RIS signal, while the combination of the two signals has similar performance to only using the RIS signal. Focusing on $\Delta = 0.1$ rad, we are able to improve the AUC for the third target from 0.55 without the RIS to 0.95 with the RIS.

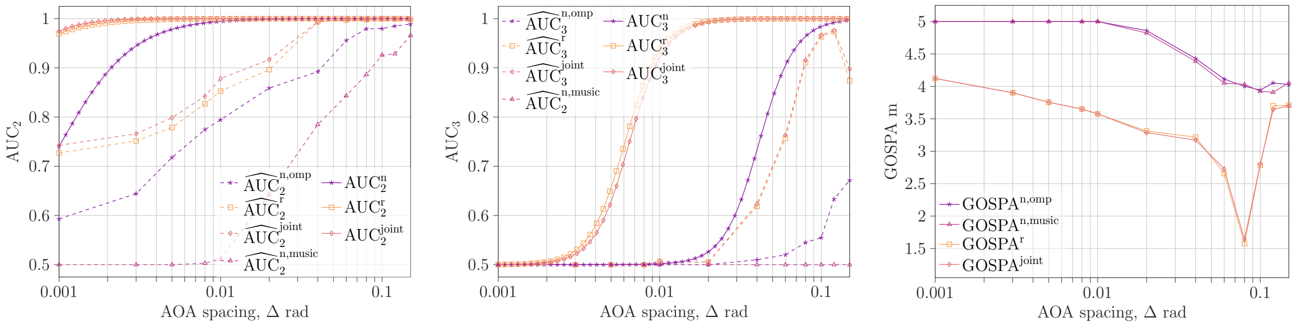
Based on preliminary calibration[7], we fix the threshold parameters $\mathrm{residual}_\mathrm{th}^\mathrm{n} = 3 \cdot 10^{-5}$, and $\mathrm{residual}_\mathrm{th}^\mathrm{r} = 3.5 \cdot 10^{-5}$, and show the GOSPA for varying AOA spacing, Δ rad, in Fig. 10c. Here, we observe that the RIS signal has a lower GOSPA than the non-RIS signal for small Δ with largest improvement of 2.5 m at $\Delta = 0.8$, however, when $\Delta > 0.1$ the RIS performance degrades. This happens due to two reasons: (i) as Δ increases the resolution region construction can fail to cover all the SPs, and (ii) as Δ increases the resolution region tends to grow causing less concentration of the RIS beams resulting in a lower receive SNR.

VI. CONCLUSION

In this work, we have assessed the potential of RIS for sensing targets in close proximity in the presence of a LOS link. With a theoretical analysis of the sensing system, considering detection probability and Fisher analysis, we have uncovered

---

[7] We ran the algorithm for a fine grid of thresholds and checked the probability of correct detection when $L = 1, 2, 3$. Then, we chose the smallest threshold which gave almost no false alarms when $L = 1, 2$.



(a) AUC target 2. (b) AUC target 3. (c) GOSPA.

Fig. 10. Results using the channel estimation algorithm for varying AOA spacing: (a,b) comparison of empirical expected AUC and the theoretical upper bound on the AUC, and (c) the GOSPA metric.

the importance of two novel coherence concepts. We can conclude that an RIS can improve the sensing resolution if the RIS-assisted signal decoheres more rapidly than the non-RIS signal, while maintaining a sufficiently high receive SNR. The results show the capability to achieve a lower coherence with the RIS-assisted signal than the non-RIS signal when the sensor has a small antenna array relative to the size of the RIS and the number of different RIS phase profiles. This improvement in terms of coherence explains the observed improvements in sensing performance in terms of both detection and localization accuracy with the RIS-assisted signal as opposed to the non-RIS signal. Finally, we find that combining the two signals can lead to sensing improvements if the signals are of similar quality. In future work we will use the coherence concept to guide optimization of the RIS phase profiles, and investigate coherent processing of the non-RIS and RIS signals, which may further improve capabilities of RIS-assisted high-resolution sensing.

APPENDIX A
CHANNEL PARAMETERS

The complex path gains α_0, $\bar{\alpha}_l$, and α_l are modelled as $\bar{\alpha}_l = P_{\bar{\alpha}_l}\bar{\xi}_l$ and $\alpha_l = P_{\alpha_l}\xi_l$ where $\bar{\xi}_l, \xi_l \sim \mathcal{CN}(0, \frac{4\sigma^2_{\mathrm{rcs},l}}{\pi})$ are independent and simulates a random phase shift and a Rayleigh distributed square root of RCS with mean σ_{rcs}. This also means that $|\alpha_l|$ (resp. $|\bar{\alpha}_l|$) is Rayleigh distributed with scale parameter $\varsigma_{\alpha_l} := \sqrt{\frac{2}{\pi}}P_{\alpha_l}\sigma_{\mathrm{rcs}}$ (resp. $\varsigma_{\bar{\alpha}_l} := \sqrt{\frac{2}{\pi}}P_{\bar{\alpha}_l}\sigma_{\mathrm{rcs}}$). The amplitudes are modelled as

$$P^2_{\alpha_0} = \frac{E_s\lambda^4(g^{\mathrm{ur}})^{4q_0}}{(4\pi)^3\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\|^4}, \tag{18a}$$

$$P^2_{\alpha_l} = \frac{E_s\lambda^2}{(4\pi)^3\|\boldsymbol{c}_l-\boldsymbol{p}\|^4}, \tag{18b}$$

$$P^2_{\bar{\alpha}_l} = \frac{E_s\lambda^4(g^{\mathrm{ur}})^{2q_0}(g^{\mathrm{sr}}_l)^{2q_0}}{(4\pi)^4\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\|^2\|\boldsymbol{p}_{\mathrm{r}}-\boldsymbol{c}_l\|^2\|\boldsymbol{c}_l-\boldsymbol{p}\|^2}, \tag{18c}$$

where E_{s} is the energy per subcarrier, $q_0 = 0.285$, $g^{\mathrm{ur}} = \frac{(\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}})^\top\boldsymbol{n}_{\mathrm{r}}}{\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\|}$, $g^{\mathrm{sr}}_l = \frac{(\boldsymbol{c}_l-\boldsymbol{p}_{\mathrm{r}})^\top\boldsymbol{n}_{\mathrm{r}}}{\|\boldsymbol{c}_l-\boldsymbol{p}_{\mathrm{r}}\|}$, and $\boldsymbol{n}_{\mathrm{r}}$ is the normal vector of the RIS [26], [35]. The TOAs τ_0, $\bar{\tau}_l$, and τ_l and the AOAs $\boldsymbol{\phi}_0$, $\boldsymbol{\phi}_l$, $\boldsymbol{\theta}_0$, and $\boldsymbol{\theta}_l$ are defined as

$$\tau_0 = \frac{2\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\|}{c}, \tag{19a}$$

$$\boldsymbol{\theta}_0 = \big(\mathrm{atan2}(y^{\mathrm{ru}}, x^{\mathrm{ru}}), \arccos(z^{\mathrm{ru}})\big)^\top, \tag{19b}$$

$$\boldsymbol{\phi}_0 = \big(\mathrm{atan2}(y^{\mathrm{ur}}, x^{\mathrm{ur}}), \arccos(z^{\mathrm{ur}})\big)^\top, \tag{19c}$$

$$\tau_l = \tau(\boldsymbol{c}_l) = \frac{2\|\boldsymbol{c}_l-\boldsymbol{p}\|}{c}, \tag{19d}$$

$$\bar{\tau}_l = \bar{\tau}(\boldsymbol{c}_l) = \frac{\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\| + \|\boldsymbol{p}_{\mathrm{r}}-\boldsymbol{c}_l\| + \|\boldsymbol{c}_l-\boldsymbol{p}\|}{c}, \tag{19e}$$

$$\boldsymbol{\theta}_l = \boldsymbol{\theta}(\boldsymbol{c}_l) = \big(\mathrm{atan2}(y^{\mathrm{su}}_l, x^{\mathrm{su}}_l), \arccos(z^{\mathrm{su}}_l)\big)^\top, \tag{19f}$$

$$\boldsymbol{\phi}_l = \boldsymbol{\phi}(\boldsymbol{c}_l) = \big(\mathrm{atan2}(y^{\mathrm{sr}}_l, x^{\mathrm{sr}}_l), \arccos(z^{\mathrm{sr}}_l)\big)^\top, \tag{19g}$$

where $\boldsymbol{x}^{\mathrm{ru}} = (x^{\mathrm{ru}}, y^{\mathrm{ru}}, z^{\mathrm{ru}})^\top$, $\boldsymbol{x}^{\mathrm{su}}_l = (x^{\mathrm{su}}_l, y^{\mathrm{su}}_l, z^{\mathrm{su}}_l)^\top$, $\boldsymbol{x}^{\mathrm{ur}} = (x^{\mathrm{ur}}, y^{\mathrm{ur}}, z^{\mathrm{ur}})^\top$, and $\boldsymbol{x}^{\mathrm{sr}}_l = (x^{\mathrm{sr}}_l, y^{\mathrm{sr}}_l, z^{\mathrm{sr}}_l)^\top$ are the direction vectors of the RIS to the UE, the lth SP to the UE, the UE to the RIS, and the lth SP to the RIS in the respective local coordinate system. These vectors can be expressed using global positions $\boldsymbol{c}_l$, $\boldsymbol{p}$, $\boldsymbol{p}_{\mathrm{r}}$ and rotation matrices $\boldsymbol{Q}_{\mathrm{u}}$ and $\boldsymbol{Q}_{\mathrm{r}}$ as $\boldsymbol{x}^{\mathrm{ru}} = \boldsymbol{Q}^\top_{\mathrm{u}}\frac{\boldsymbol{p}_{\mathrm{r}}-\boldsymbol{p}}{\|\boldsymbol{p}_{\mathrm{r}}-\boldsymbol{p}\|}$, $\boldsymbol{x}^{\mathrm{su}}_l = \boldsymbol{Q}^\top_{\mathrm{u}}\frac{\boldsymbol{c}_l-\boldsymbol{p}}{\|\boldsymbol{c}_l-\boldsymbol{p}\|}$, $\boldsymbol{x}^{\mathrm{sr}}_l = \boldsymbol{Q}^\top_{\mathrm{r}}\frac{\boldsymbol{c}_l-\boldsymbol{p}_{\mathrm{r}}}{\|\boldsymbol{c}_l-\boldsymbol{p}_{\mathrm{r}}\|}$, and $\boldsymbol{x}^{\mathrm{ur}} = \boldsymbol{Q}^\top_{\mathrm{r}}\frac{\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}}{\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\|}$ [35]. The function $\mathrm{atan2}(x, y)$ returns the principal value of the argument of $x + jy$.

The expressions (19d) to (19g) define a mapping from the SP position $\boldsymbol{c}_l$ into the channel parameter space $(\tau_l, \bar{\tau}_l, \boldsymbol{\theta}_l, \boldsymbol{\phi}_l)$, specifically, we define $\boldsymbol{h}(\boldsymbol{c}) = [\tau(\boldsymbol{c}), \boldsymbol{\theta}^\top(\boldsymbol{c}), \bar{\tau}(\boldsymbol{c}), \boldsymbol{\phi}^\top(\boldsymbol{c})]^\top$.

APPENDIX B
PROOF OF RESULTS IN SECTION III.

A. *Proof of Proposition 1*

Conditioned on the channel coefficients we have that $\boldsymbol{y}|\boldsymbol{\alpha} \sim \mathcal{CN}(\boldsymbol{G}\boldsymbol{\alpha}, \frac{\sigma^2}{2}\boldsymbol{I})$. In the l-th iteration for $l = 2, \dots, L$, given that we know $\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_{l-1}$, the maximum likelihood estimate of $\boldsymbol{\alpha}_{1:l-1}$ is

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_{l,1:l-1} &= \big(\boldsymbol{G}^H_{1:l-1}\boldsymbol{G}_{1:l-1}\big)^{-1}\boldsymbol{G}^H_{1:l-1}\boldsymbol{y} \\ &= \boldsymbol{\alpha}_{1:l-1} + \boldsymbol{G}^\dagger_{1:l-1}\big(\boldsymbol{G}_{l:L}\boldsymbol{\alpha}_{l:L} + \boldsymbol{\varepsilon}\big)\end{aligned} \tag{20}$$

where $\boldsymbol{G}^\dagger_{1:l-1} := \big(\boldsymbol{G}^H_{1:l-1}\boldsymbol{G}_{1:l-1}\big)^{-1}\boldsymbol{G}^H_{1:l-1}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{G}_{1:l-1}$. Following the greedy approach of OMP, we define a residual as

$$\boldsymbol{r}_l := \boldsymbol{y} - \boldsymbol{G}_{1:l-1}\hat{\boldsymbol{\alpha}}_{l,1:l-1} = \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big)\big(\boldsymbol{G}_{l:L}\boldsymbol{\alpha}_{l:L} + \boldsymbol{\varepsilon}\big) \tag{21}$$

where $\boldsymbol{P}_{l-1}$ is the projection matrix for the maximum likelihood fitting, i.e., $\boldsymbol{P}_{l-1} = \boldsymbol{G}_{1:l-1}\boldsymbol{G}^\dagger_{1:l-1}$. Accordingly, $(\boldsymbol{I} - \boldsymbol{P}_{l-1})$ is itself a projection matrix onto the orthogonal space. Hence, conditioned on the channel coefficient we have that

$$\boldsymbol{r}_l|\boldsymbol{\alpha}_{l:L} \sim \mathcal{CN}\Big(\big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big)\boldsymbol{G}_{l:L}\boldsymbol{\alpha}_{l:L}, \frac{\sigma^2}{2}\big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big)\Big). \tag{22}$$

Assume now that $\boldsymbol{\eta}_l$ is known. The least squares estimate of the channel coefficient is then $\hat{\alpha}_l = \frac{\boldsymbol{g}_l^H \boldsymbol{r}_l}{\|\boldsymbol{g}_l\|^2}$, hence $\hat{\alpha}_l | \boldsymbol{\alpha}_{l:L}$ is circularly-symmetric complex Gaussian with $\mathbb{E}[\hat{\alpha}_l] = \frac{\boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{G}_{l:L} \boldsymbol{\alpha}_{l:L}}{\|\boldsymbol{g}_l\|^2}$ and $\mathrm{Var}[\hat{\alpha}_l] = \frac{\sigma^2 \boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{g}_l}{2\|\boldsymbol{g}_l\|^4}$. Hence, the statistic $\gamma_l := \frac{2|\hat{\alpha}_l|^2}{\mathrm{Var}[\hat{\alpha}_l]}$ follows a non-central χ^2 distribution with 2 degrees of freedom and non-centrality parameter

$$\mu_l := \frac{2|\mathbb{E}[\hat{\alpha}_l]|^2}{\mathrm{Var}[\hat{\alpha}_l]} = \frac{4|\boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{G}_{l:L} \boldsymbol{\alpha}_{l:L}|^2}{\sigma^2 \boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{g}_l} \tag{23}$$

and the result follows immediately by observing that $p_{\mathrm{d},l}(\boldsymbol{\alpha}_{l:L}) = \Pr(\gamma_l \geq \gamma_{\mathrm{th}} | \boldsymbol{\alpha}_{l:L}) = Q_1(\sqrt{\mu_l}, \sqrt{\gamma_{\mathrm{th}}})$.

B. Proof of Corollary 1

Let $\beta_l := |\boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{G}_{l:L} \boldsymbol{\alpha}_{l:L}|^2$ and notice that the distributional assumptions of α_l implies that $\sqrt{\beta_l}$ is Rayleigh distributed with squared scale parameter ζ_l. Define $A_l := \frac{4}{\sigma^2 \boldsymbol{g}_l^H \big(\boldsymbol{I} - \boldsymbol{P}_{l-1}\big) \boldsymbol{g}_l}$ such that $\mu_l = A_l \beta_l$. The marginal distribution for detection, also interpreted as the expected detection probability is $p_{\mathrm{d},l} = \Pr(\gamma_l \geq \gamma_{\mathrm{th}})$. By the law of total probability, $p_{\mathrm{d},l}$ can be computed as

$$\begin{aligned} p_{\mathrm{d},l} &= \int_0^\infty Q_1(\sqrt{A_l \beta_l}, \sqrt{\gamma_{\mathrm{th}}}) \frac{\beta_l}{\zeta_l} \exp\Big(\frac{-\beta_l}{2\zeta_l}\Big) \mathrm{d}\beta_l \\ &= \int_0^\infty x Q_1(\sqrt{A_l \zeta_l} x, \sqrt{\gamma_{\mathrm{th}}}) \exp\Big(\frac{-x^2}{2}\Big) \mathrm{d}x, \end{aligned} \tag{24}$$

using variable substitution $x := \sqrt{\frac{\beta_l}{\zeta_l}}$. Then

$$p_{\mathrm{d},l} = \exp\Big(\frac{-\gamma_{\mathrm{th}}}{2A_l\zeta_l + 2}\Big) = \exp\Big(\frac{\log(p_{\mathrm{fa}})}{A_l\zeta_l + 1}\Big), \tag{25}$$

is found by computing the integral [40].

APPENDIX C
FISHER INFORMATION DERIVATIONS

In this appendix, we provide the partial derivatives needed in the Fisher analysis Section III-B. Before providing the partial derivatives, for convenience of notation, we begin by defining

$$\boldsymbol{g}^{\mathrm{n}}_{\tau_l} = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \boldsymbol{1}_{\tilde{T}} \otimes \big(\boldsymbol{n} \odot \boldsymbol{d}(\tau_l)\big) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l), \tag{26a}$$

$$\begin{aligned} \boldsymbol{g}^{\mathrm{n}}_{\theta_l^{\mathrm{az}}} = \boldsymbol{1}_{\tilde{T}} \otimes \boldsymbol{d}(\tau_l) \otimes \Big(& \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \frac{\partial \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}} \\ &+ \langle \frac{\partial \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}}, f \rangle \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l) \Big). \end{aligned} \tag{26b}$$

$$\begin{aligned} \boldsymbol{g}^{\mathrm{n}}_{\theta_l^{\mathrm{el}}} = \boldsymbol{1}_{\tilde{T}} \otimes \boldsymbol{d}(\tau_l) \otimes \Big(& \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \frac{\partial \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}} \\ &+ \langle \frac{\partial \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}}, f \rangle \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l) \Big). \end{aligned} \tag{26c}$$

where we have defined $\boldsymbol{n} = [0, 1, \ldots, N-1]^\top$, and where the partial derivatives of the array response vector are given as

$$\frac{\partial \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}} = j \Big(\boldsymbol{P}_{\mathrm{u}}^\top \frac{\partial \boldsymbol{\kappa}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}} \Big) \odot \exp\big(j \boldsymbol{P}_{\mathrm{u}}^\top \boldsymbol{\kappa}(\boldsymbol{\theta}_l)\big), \tag{27a}$$

$$\frac{\partial \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}} = j \Big(\boldsymbol{P}_{\mathrm{u}}^\top \frac{\partial \boldsymbol{\kappa}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}} \Big) \odot \exp\big(j \boldsymbol{P}_{\mathrm{u}}^\top \boldsymbol{\kappa}(\boldsymbol{\theta}_l)\big), \tag{27b}$$

with partial derivatives of the wavenumber vector given as $\frac{\partial \boldsymbol{\kappa}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}} = \frac{2\pi}{\lambda}[-\sin(\theta_l^{\mathrm{az}}) \sin\big(\theta_l^{\mathrm{el}}\big), \cos(\theta_l^{\mathrm{az}}) \sin\big(\theta_l^{\mathrm{el}}\big), 0]^\top$, and $\frac{\partial \boldsymbol{\kappa}(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}} = \frac{2\pi}{\lambda}[\cos(\theta_l^{\mathrm{az}}) \cos\big(\theta_l^{\mathrm{el}}\big), \sin(\theta_l^{\mathrm{az}}) \cos\big(\theta_l^{\mathrm{el}}\big), -\sin\big(\theta_l^{\mathrm{el}}\big)]^\top$.

Now the partial derivatives for the non-RIS signal are simply given by $\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \Re\{\alpha_l\}} = \boldsymbol{g}_l^{\mathrm{n}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \Im\{\alpha_l\}} = j\boldsymbol{g}_l^{\mathrm{n}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \tau_l} = -j2\pi\Delta_f \alpha_l \boldsymbol{g}^{\mathrm{n}}_{\tau_l}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \theta_l^{\mathrm{az}}} = \alpha_l \boldsymbol{g}^{\mathrm{n}}_{\theta_l^{\mathrm{az}}}$, and $\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \theta_l^{\mathrm{el}}} = \alpha_l \boldsymbol{g}^{\mathrm{n}}_{\theta_l^{\mathrm{el}}}$. The elements of the FIM can then easily be derived, e.g., $\Re\{(\frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \Re\{\alpha_i\}})^H \frac{\partial \boldsymbol{\mu}^{\mathrm{n}}}{\partial \Re\{\alpha_j\}}\} = \Re\{\langle \boldsymbol{g}_i^{\mathrm{n}}, \boldsymbol{g}_j^{\mathrm{n}} \rangle\}$, $i, j = 1, \ldots, L$. We omit listing all the elements of the FIM for conciseness of presentation.

We define for convenience the notation

$$\boldsymbol{g}^{\mathrm{r}}_{\tau_l} = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \boldsymbol{\nu}(\boldsymbol{\phi}_l) \otimes \big(\boldsymbol{n} \odot \boldsymbol{d}(\bar{\tau}_l)\big) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0), \tag{28a}$$

$$\boldsymbol{g}^{\mathrm{r}}_{\theta_l^{\mathrm{az}}} = \langle \frac{\partial \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{az}}}, \boldsymbol{f} \rangle \boldsymbol{\nu}(\boldsymbol{\phi}_l) \otimes \boldsymbol{d}(\bar{\tau}_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0), \tag{28b}$$

$$\boldsymbol{g}^{\mathrm{r}}_{\theta_l^{\mathrm{el}}} = \langle \frac{\partial \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l)}{\partial \theta_l^{\mathrm{el}}}, \boldsymbol{f} \rangle \boldsymbol{\nu}(\boldsymbol{\phi}_l) \otimes \boldsymbol{d}(\bar{\tau}_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0), \tag{28c}$$

$$\boldsymbol{g}^{\mathrm{r}}_{\phi_l^{\mathrm{az}}} = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \frac{\partial \boldsymbol{\nu}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{az}}} \otimes \boldsymbol{d}(\bar{\tau}_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0), \tag{28d}$$

$$\boldsymbol{g}^{\mathrm{r}}_{\phi_l^{\mathrm{el}}} = \langle \boldsymbol{a}_{\mathrm{u}}^*(\boldsymbol{\theta}_l), f \rangle \frac{\partial \boldsymbol{\nu}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{el}}} \otimes \boldsymbol{d}(\bar{\tau}_l) \otimes \boldsymbol{a}_{\mathrm{u}}(\boldsymbol{\theta}_0), \tag{28e}$$

where the partial derivatives of the RIS response vector is

$$\Big[\frac{\partial \boldsymbol{\nu}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{az}}}\Big]_{\tilde{t}} = \boldsymbol{\omega}_{\tilde{t}} \Big(\frac{\partial \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{az}}} \odot \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_0) \Big), \tag{29a}$$

$$\Big[\frac{\partial \boldsymbol{\nu}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{el}}}\Big]_{\tilde{t}} = \boldsymbol{\omega}_{\tilde{t}} \Big(\frac{\partial \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_l)}{\partial \phi_l^{\mathrm{el}}} \odot \boldsymbol{a}_{\mathrm{r}}(\boldsymbol{\phi}_0) \Big), \tag{29b}$$

for $\tilde{t} = 1, \ldots, \tilde{T}$. Now, the partial derivatives for the RIS signal are represented as $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \Re\{\bar{\alpha}_l\}} = \boldsymbol{g}_l^{\mathrm{r}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \Im\{\bar{\alpha}_l\}} = j\boldsymbol{g}_l^{\mathrm{r}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \bar{\tau}_l} = -j2\pi\Delta_f \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}_{\tau_l}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \theta_l^{\mathrm{az}}} = \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}_{\theta_l^{\mathrm{az}}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \theta_l^{\mathrm{el}}} = \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}_{\theta_l^{\mathrm{el}}}$, $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \phi_l^{\mathrm{az}}} = \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}_{\phi_l^{\mathrm{az}}}$, and $\frac{\partial \boldsymbol{\mu}^{\mathrm{r}}}{\partial \phi_l^{\mathrm{el}}} = \bar{\alpha}_l \boldsymbol{g}^{\mathrm{r}}_{\phi_l^{\mathrm{el}}}$, and the elements of the FIM can easily be found using this.

We provide the partial derivatives used to compute the Jacobian matrices, $\boldsymbol{T}^{\mathrm{n}}$ and $\boldsymbol{T}^{\mathrm{r}}$:

$$\frac{\partial \tau(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = \frac{2(\boldsymbol{c}_l - \boldsymbol{p})}{c\|\boldsymbol{c}_l - \boldsymbol{p}\|}, \tag{30a}$$

$$\frac{\partial \bar{\tau}(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = \frac{\boldsymbol{p}_{\mathrm{r}} - \boldsymbol{c}_l}{c\|\boldsymbol{p}_{\mathrm{r}} - \boldsymbol{c}_l\|} + \frac{\boldsymbol{c}_l - \boldsymbol{p}}{c\|\boldsymbol{p} - \boldsymbol{c}_l\|}, \tag{30b}$$

$$\frac{\partial \theta^{\mathrm{az}}(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = \frac{x_l^{\mathrm{su}}[\boldsymbol{Q}_{\mathrm{u}}]_{:,2} - y_l^{\mathrm{su}}[\boldsymbol{Q}_{\mathrm{u}}]_{:,1}}{(x_l^{\mathrm{su}})^2 + (y_l^{\mathrm{su}})^2}, \tag{30c}$$

$$\frac{\partial \theta^{\mathrm{el}}(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = -\frac{\|\boldsymbol{x}_l^{\mathrm{su}}\|^2[\boldsymbol{Q}_{\mathrm{u}}]_{:,3} - z_l^{\mathrm{su}}(\boldsymbol{c}_l - \boldsymbol{p})}{\|\boldsymbol{x}_l^{\mathrm{su}}\|^3 \Big(1 - \Big(\frac{z_l^{\mathrm{su}}}{\|\boldsymbol{x}_l^{\mathrm{su}}\|}\Big)^2\Big)^{1/2}}, \tag{30d}$$

$$\frac{\partial \phi^{\mathrm{az}}(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = \frac{x_l^{\mathrm{sr}}[\boldsymbol{Q}_{\mathrm{r}}]_{:,2} - y_l^{\mathrm{sr}}[\boldsymbol{Q}_{\mathrm{r}}]_{:,1}}{(x_l^{\mathrm{sr}})^2 + (y_l^{\mathrm{sr}})^2}, \tag{30e}$$

$$\frac{\partial \phi^{\mathrm{el}}(\boldsymbol{c}_l)}{\partial \boldsymbol{c}_l} = -\frac{\|\boldsymbol{x}_l^{\mathrm{sr}}\|^2[\boldsymbol{Q}_{\mathrm{r}}]_{:,3} - z_l^{\mathrm{sr}}(\boldsymbol{c}_l - \boldsymbol{p}_{\mathrm{r}})}{\|\boldsymbol{x}_l^{\mathrm{sr}}\|^3 \Big(1 - \Big(\frac{z_l^{\mathrm{sr}}}{\|\boldsymbol{x}_l^{\mathrm{sr}}\|}\Big)^2\Big)^{1/2}}. \tag{30f}$$

REFERENCES

[1] F. Liu, Y. Cui, C. Masouros, *et al.*, “Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[2] A. Liu, Z. Huang, M. Li, *et al.*, “A survey on fundamental limits of integrated sensing and communication,” *IEEE Commun. Surv. Tutor.*, vol. 24, no. 2, pp. 994–1034, 2022.

[3] S. P. Chepuri, N. Shlezinger, F. Liu, G. C. Alexandropoulos, S. Buzzi, and Y. C. Eldar, “Integrated sensing and communications with reconfigurable intelligent surfaces: From signal modeling to processing,” *IEEE Signal Process. Mag.*, vol. 40, no. 6, pp. 41–62, 2023.

[4] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, “Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications,” *IEEE Signal Process. Mag.*, vol. 39, no. 2, pp. 135–158, 2022.

[5] Y. Cui, F. Liu, X. Jing, and J. Mu, “Integrating sensing and communications for ubiquitous iot: Applications, trends, and challenges,” *IEEE Network*, vol. 35, no. 5, pp. 158–167, 2021.

[6] C. Chen, H. Song, Q. Li, F. Meneghello, F. Restuccia, and C. Cordeiro, “Wi-Fi sensing based on IEEE 802.11bf,” *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 121–127, 2023.

[7] W. A. Ahmad, J. Wessel, H. J. Ng, and D. Kissinger, “IoT-ready millimeter-wave radar sensors,” in *IEEE Global Conf. Artif. Intell. Internet Things (GCAIoT)*, 2020, pp. 1–5.

[8] F. Munier, Z. Xiong, R. Shreevastav, *et al.*, “Positioning of RedCap devices in 5G networks,” *IEEE Commun. Mag.*, vol. 62, no. 8, pp. 110–116, 2024.

[9] D. Tagliaferri, M. Manzoni, M. Mizmizi, *et al.*, “Cooperative coherent multistatic imaging and phase synchronization in networked sensing,” *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2905–2921, 2024.

[10] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, “Radio localization and mapping with reconfigurable intelligent surfaces: Challenges, opportunities, and research directions,” *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, 2020.

[11] Y. He, Y. Chen, Y. Hu, and B. Zeng, “Wifi vision: Sensing, recognition, and detection with commodity mimo-ofdm wifi,” *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8296–8317, 2020.

[12] M. Mercuri, E. Arnieri, R. De Marco, P. Veltri, F. Crupi, and L. Boccia, “Reconfigurable intelligent surface-aided indoor radar monitoring: A feasibility study,” *IEEE J. Electromagn., RF Microw. Med. Bio.*, vol. 7, no. 4, pp. 354–364, 2023.

[13] B. Kehoe, S. Patil, P. Abbeel, and K. Goldberg, “A survey of research on cloud robotics and automation,” *IEEE Trans. Automat. Sci. Eng.*, vol. 12, no. 2, pp. 398–409, 2015.

[14] T. Jiang, M. Jin, Q. Guo, Y. Liu, Y. Li, and J. Yao, “Full-duplex ISAC-enabled D2D underlaid cellular networks: Joint transceiver beamforming and power allocation,” *IEEE Trans. Cogn. Commun. Netw.*, pp. 1–1, 2025.

[15] W. Tang, M. Z. Chen, X. Chen, *et al.*, “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 1, pp. 421–439, 2021.

[16] A. Aubry, A. De Maio, and M. Rosamilia, “Reconfigurable intelligent surfaces for N-LOS radar surveillance,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 735–10 749, 2021.

[17] K. Meng, Q. Wu, R. Schober, and W. Chen, “Intelligent reflecting surface enabled multi-target sensing,” *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8313–8330, 2022.

[18] K. Keykhosravi, G. Seco-Granados, G. C. Alexandropoulos, and H. Wymeersch, “RIS-enabled self-localization: Leveraging controllable reflections with zero access points,” in *2022 IEEE Int. Conf. Commun.*, 2022, pp. 2852–2857.

[19] H. Chen, H. Kim, M. Ammous, *et al.*, “RISs and sidelink communications in smart cities: The key to seamless localization and sensing,” *IEEE Commun. Mag.*, vol. 61, no. 8, pp. 140–146, 2023.

[20] H. Chen, P. Zheng, M. F. Keskin, T. Al-Naffouri, and H. Wymeersch, “Multi-RIS-enabled 3D sidelink positioning,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2024.

[21] Z. Yang, H. Zhang, H. Zhang, *et al.*, “MetaSLAM: Wireless simultaneous localization and mapping using reconfigurable intelligent surfaces,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2022.

[22] 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*, vol. 110, no. 9, pp. 1401–1422, 2022.

[23] Y. Huang, J. Yang, W. Tang, C.-K. Wen, S. Xia, and S. Jin, “Joint localization and environment sensing by harnessing NLOS components in RIS-aided mmWave communication systems,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2023.

[24] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, “Foundations of MIMO radar detection aided by reconfigurable intelligent surfaces,” *IEEE Trans. Signal Process.*, vol. 70, pp. 1749–1763, 2022.

[25] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, “MetaRadar: Multi-target detection for reconfigurable intelligent surface aided radar systems,” *IEEE Trans. Wirel. Commun.*, vol. 21, no. 9, pp. 6994–7010, 2022.

[26] H. Kim, A. Fascista, H. Chen, *et al.*, “RIS-aided monostatic sensing and object detection with single and double bounce multipath,” in *2023 IEEE Int. Conf. Commun. Workshops*, 2023, pp. 1883–1889.

[27] E. Grossi, H. Taremizadeh, and L. Venturino, “Radar target detection and localization aided by an active reconfigurable intelligent surface,” *IEEE Signal Process. Lett.*, vol. 30, pp. 903–907, 2023.

[28] M. Rihan, E. Grossi, L. Venturino, and S. Buzzi, “Spatial diversity in radar detection via active reconfigurable intelligent surfaces,” *IEEE Signal Process. Lett.*, vol. 29, pp. 1242–1246, 2022.

[29] G. Zhang, D. Zhang, Y. He, J. Chen, F. Zhou, and Y. Chen, “Multi-person passive WiFi indoor localization with intelligent reflecting surface,” *IEEE Trans. Wirel. Commun.*, vol. PP, pp. 1–1, Oct. 2023.

[30] J. Hu, H. Zhang, K. Bian, M. D. Renzo, Z. Han, and L. Song, “Metasensing: Intelligent metasurface assisted RF 3D sensing by deep reinforcement learning,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 2182–2197, 2021.

[31] Y. He, D. Zhang, and Y. Chen, “High-resolution WiFi imaging with reconfigurable intelligent surfaces,” *IEEE Internet Things J.*, vol. 10, no. 2, pp. 1775–1786, 2023.

[32] J. An, C. Xu, Q. Wu, *et al.*, “Codebook-based solutions for reconfigurable intelligent surfaces and their open challenges,” *IEEE Wirel. Commun.*, pp. 1–8, 2022.

[33] F. Saggese, V. Croisfelt, R. Kotaba, K. Stylianopoulos, G. C. Alexandropoulos, and P. Popovski, “On the impact of control signaling in RIS-empowered wireless communications,” *IEEE Open J. Commun. Soc.*, vol. 5, pp. 4383–4399, 2024.

[34] G. San Antonio, D. R. Fuhrmann, and F. C. Robey, “MIMO radar ambiguity functions,” *IEEE J. Sel. Top. Signal Process.*, vol. 1, no. 1, pp. 167–177, 2007.

[35] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, “Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems,” *IEEE Trans. Wirel. Commun.*, vol. 17, no. 8, pp. 4939–4954, 2018.

[36] H. Kim, H. Chen, M. F. Keskin, *et al.*, “RIS-enabled and access-point-free simultaneous radio localization and mapping,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2023.

[37] P. Zheng, H. Chen, T. Ballal, M. Valkama, H. Wymeersch, and T. Y. Al-Naffouri, “JrCUP: Joint RIS calibration and user positioning for 6G wireless systems,” *IEEE Trans. Wirel. Commun.*, vol. 23, no. 6, pp. 6683–6698, 2024.

[38] M. Emadi, E. Miandji, and J. Unger, “OMP-based DOA estimation performance analysis,” *Digital Signal Processing*, vol. 79, pp. 57–65, 2018, ISSN: 1051-2004.

[39] H. Wymeersch and G. Seco-Granados, “Adaptive detection probability for mmWave 5G SLAM,” in *2nd 6G Wirel. Summit*, 2020, pp. 1–5.

[40] P. Sofotasios, M. Valkama, T. Tsifitsis, Y. Brychkov, S. Freear, and G. Karagiannidis, “Analytic solutions to a Marcum Q-function-based integral and application in energy detection of unknown signals over multipath fading channels,” in *9th Int. Conf. Cogn. Radio Oriented Wirel. Netw.*, 2014.

[41] H. Madsen and P. Thyregod, *Introduction to general and generalized linear models*, 1st ed. CRC press, 2010.

[42] U. Spagnolini, *Statistical Signal Processing in Engineering*. John Wiley & Sons, Ltd, 2018, ISBN: 9781119294016.

[43] D. Malioutov, M. Cetin, and A. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, 2005.

[44] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inf. Theo.*, vol. 53, no. 12, pp. 4655–4666, 2007.

[45] Z. Guo, X. Wang, and W. Heng, “Millimeter-wave channel estimation based on 2-D beamspace MUSIC method,” *IEEE Trans. Wirel. Commun.*, vol. 16, no. 8, pp. 5384–5394, 2017.

[46] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, “Generalized optimal sub-pattern assignment metric,” in *2017 20th Int. Conf. Inf. Fusion*, IEEE, 2017, pp. 1–8.

[47] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[48] P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, and cramer-rao bound,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 5, pp. 720–741, 1989.

[49] T. W. Anderson, “Asymptotic theory for principal component analysis,” *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 122–148, 1963.

[50] T.-J. Shan, M. Wax, and T. Kailath, “On spatial smoothing for direction-of-arrival estimation of coherent signals,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 4, pp. 806–811, 1985.