Sensor fusion and extended multi-target tracking in joint sensing and communication networks

Sensor Fusion and Extended Multi-Target Tracking in Joint Sensing and Communication Networks

Elia Favarelli,† Elisabetta Matricardi,† Lorenzo Pucci,† Enrico Paolini,† Wen Xu,‡ and Andrea Giorgetti†

†Wireless Communications Laboratory, CNIT, DEI, University of Bologna, Italy, Email: elia.favarelli@unibo.it
‡Munich Research Center, Huawei Technologies Duesseldorf GmbH, Munich, Germany, Email: wen.dr.xu@huawei.com

***Abstract*—In this paper, we consider a joint sensing and communication (JSC) network in which multiple base stations (BSs) cooperate through a fusion center (FC) to detect and track the objects present in a supervised area. Every BS acts as a monostatic sensor capable of scanning the environment and sensing the targets while simultaneously communicating with user equipments (UEs). In particular, each BS generates range-angle maps, which are shared with a FC for data fusion and tracking via particle filter (PF) and multi-hypothesis tracker (MHT) algorithms. The performance of the proposed solutions is evaluated by varying the fraction of power and time devoted to sensing to manage the network overhead and offer a sensing/communication trade-off. Numerical results show that the proposed algorithms can successfully track multiple targets with different sizes and behavior in a vehicular scenario, ensuring, e.g., a root mean square error (RMSE) of the estimated position of a pedestrian less than 50 cm when considering three BSs.**

I. Introduction

The next-generation mobile radio systems are expected to support many emerging functionalities, some of which are considered radically new services. Among them, the ability to perform sensing via radio frequency (RF) signals represents an attractive feature. The evolution toward higher frequency bands and larger antenna arrays will enable this new functionality. The interest in sensing is triggered by its ability to support a variety of applications, such as traffic monitoring, autonomous driving, assisted living, and environment mapping, to mention a few [1]. Among them, safety in urban environments is an emerging area where sensing may represent a technology pillar. In this scenario, the exceptionally high density of base stations (BSs) envisioned in 6G mobile radio networks paves the way to accurate and reliable localization, e.g., to protect vulnerable road users. Furthermore, with the development of massive multiple-input multiple-output (MIMO) technology at mmWave, it has become possible to achieve not only high data rates but also perform very accurate position estimation of objects. This last feature is envisioned to help emerging communication challenges such as precise beam management [2], beam tracking to maintain high-quality links by accurately estimating the direction of the desired target [3], and predictive beam tracking using joint sensing and communication (JSC) in vehicular scenarios [4].

This work has been carried out in the framework of the CNIT National Laboratory WiLab and the WiLab-Huawei Joint Innovation Center.

However, to the authors' knowledge, the current literature on JSC lacks studies about the possibility of using multiple monostatic JSC systems in cooperation through a fusion center (FC) to perform the detection and tracking of multiple targets. For this reason, this work investigates the possibility of using sensor fusion techniques combined with tracking algorithms that exploit range-angle radar maps estimated by each BS. The ultimate goal of the proposed solution is to improve the whole network sensing performance, with the capability to control resources between communication and sensing functionalities. In particular, the main contributions are the following:

- We consider a JSC network consisting of monostatic BSs adopting orthogonal frequency division multiplexing (OFDM) signals and multiple antennas. Range-angle maps are generated by scanning the environment and simultaneously communicating with user equipments (UEs). The available power is split through a multi-beam radiation pattern, and interference among BSs is avoided through frequency division (FD).
- We propose a data fusion strategy that leverages the processing of soft maps at the FC to exploit measurement information gathered from the considered set of BSs.
- We compare the performance of two tracking algorithms, particle filter (PF) and multi-hypothesis tracker (MHT), and propose solutions to handle extended targets with non-linear motion.
- To investigate the trade-off between communication and sensing and manage the network overhead introduced by BS cooperation, we evaluate the impact of the fraction of power and time dedicated to sensing on system-level performance.

Throughout the paper, capital and lowercase boldface letters denote matrices and vectors, respectively; $\mathbf{I}_n$ is the $n \times n$ identity matrix; $\left\|\cdot\right\|_p$ stands for the p-norm; $\lceil\cdot\rceil$ represents the ceiling function; $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes a zero-mean circularly symmetric complex Gaussian random vector with covariance $\mathbf{\Sigma}$, and $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ denotes a real-valued Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance $\mathbf{\Sigma}$.

The rest of the paper is organized as follows. In Section II, the proposed sensing network is presented. Section III introduces the data fusion strategies and the tracking algorithms adopted. The considered case study is described in Section IV.

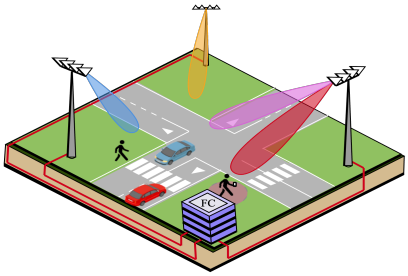

Fig. 1. An urban scenario with three monostatic JSC BSs aiming at monitoring road traffic and a vulnerable road user.

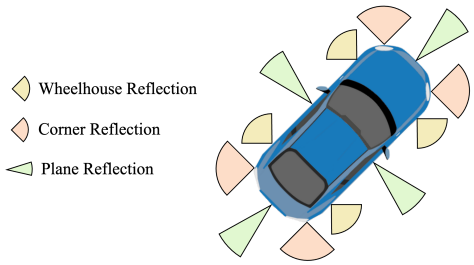

Fig. 2. Extended target model implemented to represent the vehicle reflections. Different colors represent various types of reflections.

The performance of the whole processing chain is evaluated in Section V, and conclusions are drawn in Section VI.

II. System Model

We consider a JSC network consisting of a set of monostatic MIMO BSs, based on OFDM and operating at mmWave. The BSs manage multi-cell interference adopting FD, and cooperate through a FC to detect and track the targets that are present in the surrounding environment by performing a scan while simultaneously communicating with UEs in their respective cells, as depicted in Fig. 1. Similarly to [5]–[7], every monostatic JSC system consists of two separated uniform linear arrays (ULAs) for transmission and reception, with N_{T} and N_{R} antennas, respectively, and with half-wavelength separation between the elements. In particular, we consider a 5G new radio (NR)-based system in which the transmitting antenna array is used to communicate and sense the environment simultaneously. This is done through a multi-beam radiation pattern that exploits the downlink communication signal for sensing by splitting the total available power in two directions. Differently, the receiving array is used to acquire back-scattered signals from the targets.

More in detail, every JSC system transmits 5G NR frames, each of them consisting of M OFDM symbols and K active subcarriers, to a UEs. However, as it will be explained in Section II-B, each BS scans the environment by changing the sensing direction and, in each direction, only a subset M_{s} of the total number of OFDM symbols is processed. Therefore, the transmitted signal for each direction can be represented by a matrix $\mathbf{X}_{\mathrm{s}} \in \mathbb{C}^{K \times M_{\mathrm{s}}}$, whose elements $x_k^{(m)}$ are complex modulation symbols.

With multiple antennas at the transmitter, a precoding operation is performed on the elements of $\mathbf{X}$ to map modulation symbols to the antennas. In particular, considering the beamforming vector $\mathbf{w}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$, the vector of transmitted symbols at each antenna, $\tilde{\mathbf{x}}_k^{(m)} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$, is given by $\tilde{\mathbf{x}}_k^{(m)} = \mathbf{w}_{\mathrm{T}} x_k^{(m)}$. As previously stated, the precoding operation aims at creating a multibeam radiation pattern to split the total available power between communication and sensing directions by using the same signal, as in [6], [8]. Hence, the considered beamforming vector is

$$\mathbf{w}_{\mathrm{T}} = \sqrt{\rho_{\mathrm{p}}}\, \mathbf{w}_{\mathrm{T,s}} + \sqrt{1-\rho_{\mathrm{p}}}\, \mathbf{w}_{\mathrm{T,c}} \tag{1}$$

where $\rho_{\mathrm{p}} \in [0,1]$ is a factor that represents the fraction of power reserved for sensing, thus determining a trade-off between communication and sensing functionalities, while $\mathbf{w}_{\mathrm{T,c}}$ and $\mathbf{w}_{\mathrm{T,s}}$ are respectively the communication and the sensing beamforming vectors, which are normalized with respect to the effective isotropic radiated power (EIRP), $P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}}$, and defined as [9]

$$\mathbf{w}_{\mathrm{T,c}} = \frac{\sqrt{P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}}}}{N_{\mathrm{T}}} \mathbf{a}_{\mathrm{T}}^{c}(\theta_{\mathrm{T,c}}), \quad \mathbf{w}_{\mathrm{T,s}} = \frac{\sqrt{P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}}}}{N_{\mathrm{T}}} \mathbf{a}_{\mathrm{T}}^{c}(\theta_{\mathrm{T,s}}). \tag{2}$$

In (2), P_{T} is the transmit power and $G_{\mathrm{T}}^{\mathrm{a}}$ is the transmit array gain along the beam steering direction, while $\mathbf{a}_{\mathrm{T}}(\theta_{\mathrm{T,c}}) \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ and $\mathbf{a}_{\mathrm{T}}(\theta_{\mathrm{T,s}}) \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ are the steering vectors for communication and sensing directions, respectively [6], [10]. Considering, without loss of generality, negligible inter-carrier interference (ICI) and inter-symbol interference (ISI), the vector $\tilde{\mathbf{y}}_k^{(m)} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$ of the received complex modulation symbols at antenna array after the OFDM demodulation, can be obtained as follows

$$\tilde{\mathbf{y}}_k^{(m)} = \mathbf{H}_k^{(m)} \tilde{\mathbf{x}}_k^{(m)} + \tilde{\mathbf{n}}_k \tag{3}$$

where $\mathbf{H}_k^{(m)} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ is the channel matrix for the mth OFDM symbol and the kth subcarrier, and $\tilde{\mathbf{n}}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{N}}^2 \mathbf{I}_{N_{\mathrm{R}}})$.

Then, starting from (3) and performing spatial combining in the considered sensing direction $\theta_{\mathrm{R,s}}$, through the receiving beamforming vector $\mathbf{w}_{\mathrm{R}} = \mathbf{a}_{\mathrm{R}}^{c}(\theta_{\mathrm{R,s}})$, we obtain the grid of the received modulation symbols $\mathbf{Y}_{\mathrm{s}} \in \mathbb{C}^{K \times M_{\mathrm{s}}}$, with elements $y_k^{(m)} = \mathbf{w}_{\mathrm{R}}^T \tilde{\mathbf{y}}_k^{(m)}$. The symbol grids $\mathbf{Y}_{\mathrm{s}}$ collected in each sensing direction are used to compute range-angle maps, as it will be detailed in Section II-B.

A. Target and Channel Model

The channel matrix in (3) considers the reflections generated by the targets. Every target can generate one or more back-scattered signals depending on whether it is a point-like target or an extended one. In this work, we consider both types. In particular, as will be presented in Section IV, we consider a scenario with one pedestrian modeled as a point-like target and two cars described by the extended target model shown in Fig. 2. As it can be observed, the car is modeled with 12 reflection points, 4 plane reflections due to front, back, and sides (with a narrow visibility function and large radar cross-section (RCS)), 4 due to the wheelhouses, and 4 due to corners,

as in [11], [12]. Each of these points has its visibility function[1], which depends on the impinging angle, and the RCS. Each scattering point is associated with a different Doppler shift and propagation delay. More details about the chosen values of RCS for the pedestrian and the cars will be provided in Section IV. The channel matrix in (3) in the case of L point reflectors, can be written as:

$$\mathbf{H}_k^{(m)} = \sum_{l=1}^{L} \alpha_l e^{j2\pi m T_\mathrm{s} f_{\mathrm{D},l}} e^{-j2\pi k \Delta f \tau_l} \mathbf{a}_\mathrm{R}(\theta_l) \mathbf{a}_\mathrm{T}^T(\theta_l) \quad (4)$$

where T_s is the OFDM symbol duration, Δf is the subcarrier spacing, $f_{\mathrm{D},l}$, τ_l, θ_l, and $\mathbf{a}_\mathrm{R}(\theta_l)$ are the Doppler shift, the round-trip delay, the direction of arrival (DoA), and the array response vector at the receiver of the lth back-scattered signal, respectively. The complex term $\alpha_l = |\alpha_l| e^{j\phi_l}$ includes phase shift and attenuation along the lth propagation path and its amplitude is

$$|\alpha_l| = \sqrt{\frac{c^2 \sigma_{\mathrm{rcs},l}}{(4\pi)^3 f_\mathrm{c}^2 d_l^4}} \quad (5)$$

with d_l the distance between the lth reflection points and the BS, and $\sigma_{\mathrm{rcs},l}$ its RCS; f_c and c are the carrier frequency and the speed of light, respectively. Starting from (5), the signal-to-noise ratio (SNR) measured at each receiving antenna, related to the lth reflection point is

$$\mathrm{SNR}_l = \rho_\mathrm{p} \cdot \gamma_l \cdot \frac{P_\mathrm{T} G_\mathrm{T}^\mathrm{a} G_\mathrm{R} |\alpha_l|^2}{\sigma_\mathrm{N}^2} \quad (6)$$

where G_R is the single element antenna gain at the Rx, $\gamma_l = |\mathrm{AF}(\theta_{\mathrm{T,s}} - \theta_l)|^2 \in [0,1]$ is the normalized array gain at the Tx to account for the imperfect alignment between the sensing direction and the target DoA, and $\sigma_\mathrm{N}^2 = N_0 K \Delta f$, with N_0 the one-sided noise power spectral density (PSD). Briefly returning to (4), it is noteworthy to highlight that the number of back-scattered signals L does not remain constant during the time. Indeed, the number of reflections generated by a car strongly depends on its relative angular position with respect to the BS since it is related to the visibility function previously introduced.

B. Beam-Scanning and Range-Angle Maps

Each JSC station senses the objects in the environment by performing a scan through the multi-beam radiation pattern defined by $\mathbf{w}_\mathrm{T}$ in (2). In particular, the communication beam points towards a UE, while the sensing beam changes during time, pointing towards several directions with a predefined angular step. In each sensing direction, M_s OFDM symbols are acquired and used to compute a range-angle map. Since the number of symbols per direction is fixed, the amount of time needed to complete a scan, T_scan, depends on the chosen number of sensing directions and the duration of the symbol T_s. After having acquired all the symbols to form the matrix $\mathbf{Y}_\mathrm{s}$ in Section II, the first operation we perform is an element-wise division between $\mathbf{Y}_\mathrm{s}$ and $\mathbf{X}_\mathrm{s}$ to remove the dependency from the transmitted symbols and get a new matrix $\mathbf{G}_\mathrm{s}$. The latter contains two complex sinusoids (for each target) whose frequencies are related to $f_{\mathrm{D},l}$ and τ_l, according to (4). For this reason, first, a double-periodogram (on the rows and the columns of $\mathbf{G}_\mathrm{s}$) is computed to get a range-Doppler map [6], [13], [14]. Then, starting from this map, a range-angle map is obtained by taking the column of the periodogram over which we have the maximum value and uniquely associating this to the considered sensing direction.[2]

Note that a scan of constant duration, T_scan, is considered, while the time between two consecutive measurements, T_meas, (range-angle maps update) can be varied. Hence, we indicate with $\rho_\mathrm{t} = T_\mathrm{scan}/T_\mathrm{meas}$ the fraction of time committed for sensing. By considering the power-division approach in (1), the overall fraction of resources dedicated to sensing is

$$\rho_\mathrm{sens} = \rho_\mathrm{p} \cdot \rho_\mathrm{t} \quad (7)$$

and the fraction of resources for communication is therefore $\rho_\mathrm{comm} = 1 - \rho_\mathrm{sens}$.

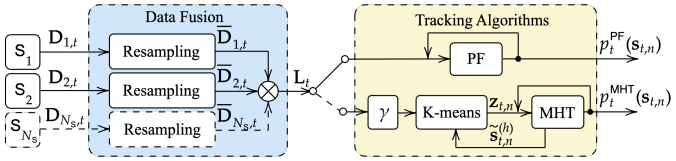

Fig. 3. Block diagram of the processing chain. The data fusion steps and the tracking algorithms proposed are depicted in blue and yellow, respectively.

III. Data Fusion and Tracking Algorithms

In Fig. 3 the processing chain is depicted. For each monostatic JSC BS, hereafter referred to as *sensor* and indexed with s, the range-angle map computed at time t and stored in the matrix $\mathbf{D}_{s,t}$, is collected by the FC, which at first performs a resampling operation through a linear uniform grid with resolution Δ_x and Δ_y, and the number of points N_x and N_y. The resampled range-angle maps $\overline{\mathbf{D}}_{s,t}$ are then fused using the element-wise product. Finally, the resulting map is stored in the matrix $\mathbf{L}_t$ and is used as target position information in the tracking algorithm measurement models described in the following.

A. Particle Filter

The PF is a popular and widely investigated tracking strategy that relies on the generation of a set of *particles* to approximate the target state (position and velocity) posterior distribution and tracks its behavior [15], [16]. Let us consider a set of N_obj objects with state vectors $\mathbf{s}_{t-1,n} = \left(s_{t-1,n,\mathrm{x}}, s_{t-1,n,\mathrm{y}}, s_{t-1,n,v_\mathrm{x}}, s_{t-1,n,v_\mathrm{y}}\right)$, where t and n are the

[1]This function has a value of one in the angular region where it can be seen best and decreases until zero in regions where the radar sensor receives no reflected energy.

[2]This choice is related to the assumption that if we consider a relatively small beamwidth, only one scatterer is likely to be present in the considered sensing direction. Since the inference of a target parameter turns out to be a frequency estimation, the periodogram represents (asymptotically) the log-likelihood.

time and target indexes. The first two vector elements represent the target position coordinates, whereas the last two represent the target velocity components. Let us also assume that the initial target states and the number of targets are known.

A set of particles $\mathbf{s}_{t-1,n}^{(q)}$, where q is the particle index, is then generated for each target to approximate its prior state distribution. The initial target state distribution can be written as

$$p_{t-1}^{\mathrm{PF}}(\mathbf{s}_{t-1,n}) = \sum_{q=1}^{N_\mathrm{p}} w_{t-1,n}^{(q)} \delta\left(\mathbf{s}_{t-1,n} - \mathbf{s}_{t-1,n}^{(q)}\right) \tag{8}$$

where N_p is the number of particles, $w_{t-1}^{(q)}$ is the weight of the qth particle at time $t-1$, and $\delta(\mathbf{s}_{t-1,n} - \mathbf{s}_{t-1,n}^{(q)})$ is a delta function centered at $\mathbf{s}_{t-1,n}^{(q)}$. All particle weights $w_{t-1,n}^{(q)}$ are initialized to $1/N_\mathrm{p}$.

In the prediction step, particles are propagated using a linear transition matrix and additive Gaussian noise. Specifically, we have $\mathbf{s}_{t,n}^{(q)} = \mathbf{F}\,\mathbf{s}_{t-1,n}^{(q)} + \boldsymbol{\nu}_{t,n}$, where $\mathbf{F}$ is the transition matrix and $\boldsymbol{\nu}_{t,n} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{\mathrm{PF}})$ with $\mathbf{Q}^{\mathrm{PF}}$ the prediction covariance matrix.

Next, we define a remapping function

$$\phi(\mathbf{s}_{t,n}^{(q)}) : \mathbb{R}^2 \mapsto \{1, \ldots, N_\mathrm{x}\} \times \{1, \ldots, N_\mathrm{y}\} \tag{9}$$

that allows us selecting the entry of $\mathbf{L}_t$ that is closest to the particle $\mathbf{s}_{t,n}^{(q)}$. This function is defined as

$$\phi(\mathbf{s}_{t,n}^{(q)}) = \left(\left\lceil \frac{s_{t,n,x}^{(q)} - b_\mathrm{x}}{\Delta_\mathrm{x}} \right\rceil, \left\lceil \frac{s_{t,n,y}^{(q)} - b_\mathrm{y}}{\Delta_\mathrm{y}} \right\rceil \right) \tag{10}$$

where b_x and b_y account for the position of the origin of the coordinate system. During the update step, particle weights are then determined using the equation

$$\widetilde{w}_{t,n}^{(q)} = w_{t-1,n}^{(q)} L_{t,\phi(\mathbf{s}_{t,n}^{(q)})} \tag{11}$$

where $L_{t,\phi(\mathbf{s}_{t,n}^{(q)})}$ is the element of $\mathbf{L}_t$ with indexes $\phi(\mathbf{s}_{t,n}^{(q)})$. At the end of the update step, particle weights are normalized according to $w_{t,n}^{(q)} = \widetilde{w}_{t,n}^{(q)} \big(\sum_{q=1}^{N_\mathrm{p}} \widetilde{w}_{t,n}^{(q)} \big)^{-1}$.

This procedure is then applied iteratively over time to perform multiple target tracking. To prevent weights drift, a resampling procedure is performed at each time step [17]. Finally, target state estimation $\hat{\mathbf{s}}_{t-1,n}^{\mathrm{PF}}$ is extracted using (8).

B. Multi-Hypothesis Tracker

An alternative to the soft decision taken by the PF sampling directly the matrix $\mathbf{L}_t$, can be the adoption of the MHT [18], [19]. This algorithm can be used in the framework of Bayesian filtering approaches [20]–[22], where the prior target distribution is assumed to be a Gaussian mixture, namely,

$$p_{t-1}^{\mathrm{MHT}}(\mathbf{s}_{t-1,n}) = \sum_{h=1}^{N_\mathrm{h}} w_{t-1,n}^{(h)} \mathcal{N}(\mathbf{s}_{t-1,n}^{(h)}, \mathbf{P}_{t-1,n}^{(h)}) \tag{12}$$

where N_h represents the number of components or hypotheses, $w_{t-1,n}^{(h)}$ its weight, and $\mathbf{s}_{t-1,n}^{(h)}$ and $\mathbf{P}_{t-1,n}^{(h)}$ the mean value and covariance of the hth hypothesis, respectively. Starting from this configuration, a Kalman linear prediction is performed to infer the target position in the consecutive time step:

$$\widetilde{\mathbf{s}}_{t,n}^{(h)} = \mathbf{F}\,\widetilde{\mathbf{s}}_{t-1,n}^{(h)}, \qquad \widetilde{\mathbf{P}}_{t,n}^{(h)} = \mathbf{F}\,\mathbf{P}_{t-1,n}^{(h)}\,\mathbf{F}^{\mathrm{T}} + \mathbf{Q}^{\mathrm{MHT}}.$$

MHT prediction is then exploited during the clustering procedure. To perform clustering on the matrix $\mathbf{L}_t$, an excision filter is implemented to elide all the values below a predefined threshold γ. Then, a k-means algorithm [23] is implemented to cluster the remaining points in a set of N_obj clusters, whose centroids $\mathbf{z}_{t,n}$ are used for MHT update. The k-means centroids are initialized with the predicted target state estimates to increase the clustering performance and reduce the computational cost [24]. MHT prediction is then corrected through the Kalman update procedure, exploiting the hard measurements $\mathbf{z}_{t,n}$, their covariance matrix $\mathbf{R}$, and considering the most likely data association hypotheses. In MHT, clutter measurements are modeled as a Poisson Point Process (PPP) with intensity λ, and during the update, the detection probability P_D is considered constant over the whole area of interest. Finally, maximum a posteriori (MAP) criterion is used to estimate the target state $\hat{\mathbf{s}}_{t-1,n}^{\mathrm{MHT}}$.

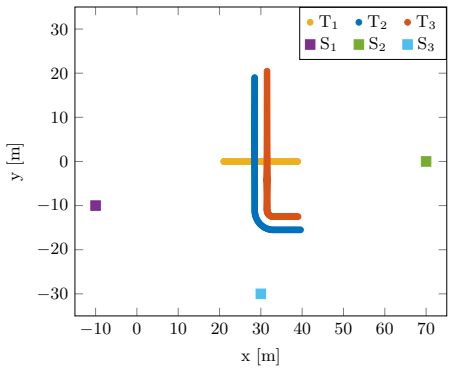

Fig. 4. The proposed case study with three BSs S_i. Blue, red and yellow traces are ground truth trajectories for two cars and a pedestrian, respectively.

IV. Case Study

The considered urban scenario, with two vehicles, one pedestrian, and three BSs, is shown in Fig 4. The targets' behavior is summarized in Tab. I. The pedestrian is modeled according to a Swerling I model with a mean RCS equal to $\bar{\sigma}_\mathrm{rcs} = 0.5\,\mathrm{m}^2$, whereas vehicles are represented as extended targets, according to the model in Section II-A. For the RCS values of each point reflector, we considered the measurements in the W band reported in [25].[3] From these measurements, it can be seen that approximately a sector of $10\,^\circ$ and $\bar{\sigma}_\mathrm{rcs} = 20\,\mathrm{dBsm}$ is present when looking at the vehicle's side. This is valid for all four plane reflectors, including front, back, and side views. The back-scattered power from other impinging angles arrives from either the wheelhouses, with $\bar{\sigma}_\mathrm{rcs} = 0\,\mathrm{dBsm}$, or from the corners, with $\bar{\sigma}_\mathrm{rcs} = 5\,\mathrm{dBsm}$.

[3]Unfortunately it was not possible to find RCS measurements at 28 GHz in the literature.

TABLE I
TARGETS BEHAVIOR IN THE VEHICULAR SCENARIO

		Pedestrian – T_1		
Motion	from [m]	to [m]	v_{ini} [m/s]	v_{fin} [m/s]
Uniform	(21,0)	(39,0)	1.8	1.8
		Vehicle #1 – T_2		
Motion	from [m]	to [m]	v_{ini} [m/s]	v_{fin} [m/s]
Decelerated	(28.5,19)	(28.5,5)	9	0
Static	(28.5,5)	(28.5,5)	0	0
Accelerated	(28.5,5.5)	(28.5,-11)	0	7
Circular	(28.5,-11)	(33,-15.5)	7	7
Uniform	(33,-15.5)	(40,-15.5)	7	7
		Vehicle #2 – T_3		
Motion	from [m]	to [m]	v_{ini} [m/s]	v_{fin} [m/s]
Decelerated	(39,-12.5)	(33,-12.5)	5	2.6
Circular	(33,-12.5)	(31.5,-11)	2.6	2.6
Accelerated	(31.5,-11)	(31.5,-5.5)	2.6	4.25
Decelerated	(31.5,-5.5)	(31.5,-4.15)	4.25	0
Static	(31.5,-4.15)	(31.5,-4.15)	0	0
Accelerated	(31.5,-4.15)	(31.5,20.5)	0	11

For what concerns system parameters, a 5G NR signal compliant with 3GPP Technical Specification in [26], is assumed. In particular, we consider a quadrature phase shift keying (QPSK) complex modulation alphabet, a carrier frequency $f_c = 28\,\text{GHz}$, a subcarrier spacing $\Delta f = 120\,\text{kHz}$, a total number of subcarrier $K = 3168$ (i.e., about $400\,\text{MHz}$ bandwidth), a number of OFDM symbols $M = 1120$ per frame, and a number of OFDM symbols for each sensing direction $M_s = 112$. The EIRP is set to $30\,\text{dBm}$, and the noise PSD is $N_0 = 4 \cdot 10^{-20}\,\text{W/Hz}$. All the three BSs are equipped with $N_T = N_R = 50$ antennas.

V. Numerical Results

In this section, the performance of the algorithms is evaluated in the case study considered. The sensing scan at each BS has duration $T_{scan} = 50\,\text{ms}$, and the overall scene is monitored for $10\,\text{s}$, resulting in a set of $N_m = 200$ measurements (maps collected). Resampling is performed through a linear uniform grid with resolution $\Delta_x = 0.1\,\text{m}$, and $\Delta_y = 0.1\,\text{m}$. The area of interest considered is in the range $x \in [20, 50]\,\text{m}$, and $y \in [-20, 40]\,\text{m}$, therefore the offsets are $b_x = 20\,\text{m}$, and $b_y = -20\,\text{m}$.

First, the system performance is investigated by varying the power amount dedicated to sensing, ρ_p from 0.01 to 0.5. Second, the algorithms are tested by varying the fraction of time for sensing ρ_t from 0.167 to 0.5. A set of $\mathcal{M} = 100$ Monte Carlo trials have been run for PF in each parameters configuration. The tracking performance of the nth target have been evaluated through the root mean square error (RMSE) defined as $\text{RMSE}_n = (\frac{1}{N_m}\sum_{t=1}^{N_m}\|\mathbf{s}_{t,n} - \hat{\mathbf{s}}_{t,n}\|_2^2)^{1/2}$, where $\mathbf{s}_{t,n}$ and $\hat{\mathbf{s}}_{t,n}$ are the true and estimated target states at the time instant t, respectively.

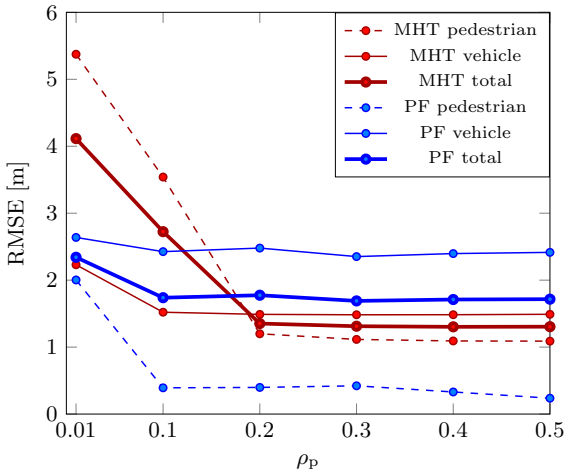

Fig. 5. Position estimation error by varying the fraction of power ρ_p.

In the PF, the number of particles is $N_p = 100$, and resampling is applied at each iteration. The prediction covariance matrix is set to $\mathbf{Q}^{PF} = \text{diag}(0.25, 0.25)$. In the MHT, the threshold is set to $\gamma = 5 \cdot 10^{-5} \cdot \max(\mathbf{L}_t)$, where $\max(\mathbf{L}_t)$ is the maximum value of the matrix $\mathbf{L}_t$. The prior covariance matrix is initialized as $\mathbf{P}_{t,n}^{(h)} = \text{diag}(0.5, 0.5, 0.5, 0.5)$, the measurements covariance matrix is set to $\mathbf{R} = \text{diag}(0.25, 0.25)$, the clutter intensity is $\lambda = 10^{-4}$, the detection probability is $P_D = 0.8$, and the prediction covariance matrix is $\mathbf{Q}^{MHT} = \text{diag}(0.25\,T_{scan}, 0.25\,T_{scan})$. For both the algorithms the transition matrix is defined as $\mathbf{F} = (1, 0, T_{scan}, 0; 0, 0, 0, T_{scan}; 0, 0, 1, 0; 0, 0, 0, 1)$.

A. Impact of ρ_p

In Fig. 5, the target localization error by varying the fraction of power reserved for sensing is reported. Dashed curves represent the performance with a pedestrian target, whereas solid lines represent the performance with a vehicle (extended target). In the considered scenario, target T_2 and T_3 are both vehicles; for this reason, RMSE is evaluated jointly for these two targets. Instead, the RMSE evaluated over the whole set of targets is shown with solid bold curves. The performance of PF is depicted in blue, whereas MHT performance is reported in red. We can observe that position estimation error decreases by increasing the fraction of power ρ_p devoted to sensing. Due to its reduced RCS, the pedestrian is more affected by the reduction of ρ_p compared to the vehicles. Moreover, it is possible to notice that PF performs better than MHT with reduced ρ_p due to its capability to process soft maps. However, MHT performs better in scenarios with high SNR, particularly with extended targets, due to its capability to cluster measurements that originate from the same target.

B. Impact of ρ_t

To relieve the network overhead due to soft-map sharing and the fraction of time dedicated to sensing, the effect of reducing the map update rate is evaluated. In Fig. 6, the localization error is reported by varying ρ_t; the higher the ρ_t, the more

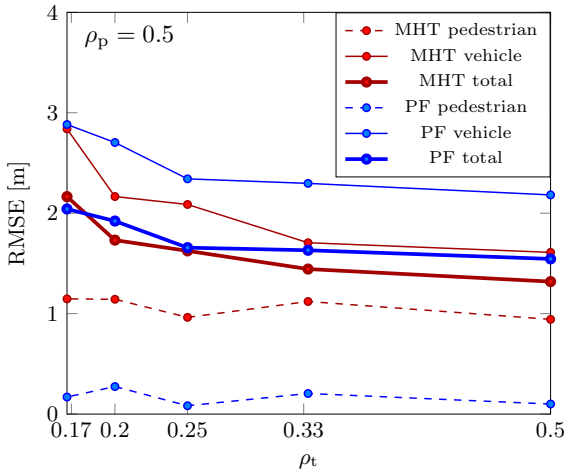

Fig. 6. Position estimation error by varying the fraction of time ρ_t.

frequently range-angle maps are estimated and exchanged with the FC. As we can see, tracking algorithms can cope with the lack of data due to infrequent updates, though the error slightly increases when $\rho_t < 0.25$, where tracking algorithms cannot follow the non-linear motion of the targets. It can also be noticed that vehicles are more sensitive to the reduction of ρ_t compared to pedestrians; that is because vehicles' motion is less predictable when the measurements interval increases ($T_{meas} = T_{scan}/\rho_t$). Moreover, we can notice that MHT and PF present similar performance and tolerance to variation of ρ_t. Fig. 6 also highlights that with $\rho_p = 0.5$, and $\rho_t = 0.25$, tracking performance is still remarkable despite using only a fraction $\rho_{sens} = 0.125$ of the resources.

VI. Conclusion

This work investigated the adoption of data fusion and tracking algorithms in a JSC system with cooperative monostatic sensors. Two different algorithms, PF and MHT, have been proposed to track targets starting from the range-angle maps. The effect of the fractions of time and power committed to sensing have been evaluated to offer a sensing/communication trade-off. Numerical results, specific for a 5G NR waveform, have shown that both solutions can track the target behavior with an RMSE lower than 0.5 m for the pedestrian, and lower than 1.5 m for the vehicles. Finally, the PF demonstrated superior robustness in low SNR regimes and better performance with targets that exhibit a low RCS, whereas MHT performed better for extended targets.

References

[1] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Towards dual-functional wireless networks for 6g and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, pp. 1728–1767, 2022.

[2] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, pp. 158–167, 2021.

[3] F. Liu, P. Zhao, and Z. Wang, "EKF-based beam tracking for mmWave MIMO systems," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2390–2393, 2019.

[4] Z. Du, F. Liu, and Z. Zhang, "Sensing-assisted beam tracking in V2I networks: Extended target case," in *IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, 2022, pp. 8727–8731.

[5] L. Pucci, E. Matricardi, E. Paolini, W. Xu, and A. Giorgetti, "Performance analysis of joint sensing and communication based on 5G New Radio," in *IEEE Work. on Adv. in Netw. Loc. and Nav. (ANLN), Globecom*, Madrid, Spain, Dec. 2021.

[6] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," in *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, July 2022, pp. 2043–2055.

[7] E. Favarelli, E. Matricardi, L. Pucci, E. Paolini, W. Xu, and A. Giorgetti, "Tracking and data fusion in joint sensing and communication networks," in *2022 IEEE Globecom Workshops*, Rio de Janeiro, Brasil, Dec. 2022, pp. 341–346.

[8] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 671–685, Jan. 2019.

[9] H. Asplund, D. Astely, P. von Butovitsch, T. Chapman, M. Frenne, F. Ghasemzadeh, M. Hagström, B. Hogan, G. Jöngren, J. Karlsson *et al.*, *Advanced Antenna Systems for 5G Network Deployments: Bridging the Gap Between Theory and Practice*. Academic Press, 2020.

[10] M. A. Richards, *Fundamentals of radar signal processing*. McGraw-Hill, 2005.

[11] M. Bühren and B. Yang, "Automotive radar target list simulation based on reflection center representation of objects," in *Proc. Int. Work. on Intelligent Transportation (WIT)*. Citeseer, 2006, pp. 161–166.

[12] ——, "Simulation of automotive radar target lists using a novel approach of object representation," in *IEEE Intelligent Veh. Symp.*, 2006, pp. 314–319.

[13] M. Braun, "OFDM radar algorithms in mobile communication networks," Ph.D. dissertation, Karlsruhe Institute of Technology, 2014.

[14] C. B. Barneto, T. Riihonen, M. Turunen, L. Anttila, M. Fleischer, K. Stadius, J. Ryynänen, and M. Valkama, "Full-duplex OFDM radar with LTE and 5G NR waveforms: challenges, solutions, and measurements," *IEEE Trans. Microw. Theory Tech.*, vol. 67, no. 10, pp. 4042–4054, Oct. 2019.

[15] P. Djuric, J. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, 2003.

[16] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms for particle filters: A computational complexity perspective," *EURASIP Journal on Advances in Signal Process.*, vol. 2004, no. 15, pp. 1–11, 2004.

[17] J. D. Hol, T. B. Schon, and F. Gustafsson, "On resampling algorithms for particle filters," in *IEEE Nonlinear Statistical Signal Process. Workshop*, 2006, pp. 79–82.

[18] B. Cheung, M. Rutten, S. Davey, and G. Cohen, "Probabilistic multi hypothesis tracker for an event based sensor," in *Int. Conf. on Information Fusion (FUSION)*, 2018, pp. 1–8.

[19] K. O. Arras, S. Grzonka, M. Luber, and W. Burgard, "Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities," in *IEEE Int. Conf. on Robotics and Autom.*, 2008, pp. 1710–1715.

[20] V. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *IEEE Pervasive Computing*, vol. 2, no. 3, pp. 24–33, 2003.

[21] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge university press, 2013, no. 3.

[22] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, 2003.

[23] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003, Biometrics.

[24] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for k-means clustering," *Pattern Recognition Letters*, vol. 25, no. 11, pp. 1293–1302, 2004.

[25] E. Bel Kamel, A. Peden, and P. Pajusco, "RCS modeling and measurements for automotive radar applications in the W band," in *Europ. Conf. on Ant. and Propag. (EuCAP)*, 2017, pp. 2445–2449.

[26] *5G NR: Physical channels and modulation*, 3GPP TS 38.211, 7 2020, version 16.2.0 Release 16.