Multi-sensor Spatial Association using Joint Range-Doppler Features

Anant Gupta, *Student Member, IEEE*, Ahmet Dundar Sezer, *Member, IEEE*, and Upamanyu Madhow, *Fellow, IEEE*

Abstract—**We investigate the problem of localizing multiple targets using a single set of measurements from a network of radar sensors. Such "single snapshot imaging" provides timely situational awareness, but can utilize neither platform motion, as in synthetic aperture radar, nor track targets across time, as in Kalman filtering and its variants. Associating measurements with targets becomes a fundamental bottleneck in this setting. In this paper, we present a computationally efficient method to extract 2D position and velocity of multiple targets using a linear array of FMCW radar sensors by identifying and exploiting inherent geometric features to drastically reduce the complexity of spatial association. The proposed framework is robust to detection anomalies, and achieves order of magnitude lower complexity compared to conventional methods. While our approach is compatible with conventional FFT-based range-Doppler processing, we show that more sophisticated techniques for range-Doppler estimation lead to reduced data association complexity as well as higher accuracy estimates of target positions and velocities.**

Index Terms—**Sensor Networks, Aggregation, Approximation Algorithms, Single Snapshot Localization**

I. INTRODUCTION

RECENT advances in low-cost design and fabrication enable the potential application of high-accuracy millimeter wave (mmWave) radar sensors to a variety of commercial sectors, including automotive, drones and robotics [1], [2]. The large available bandwidths enable high range resolution, while the small wavelength enhances Doppler and microDoppler resolution. In this paper, we explore the utility of a network of such sensors in providing timely situational awareness for highly dynamic environments, by considering estimation of the kinematic state of the scene (i.e., the positions and velocities of targets) via a single set of measurements obtained by a network of sensors. We do not rely on tracking targets across time, or on platform motion to synthesize larger apertures.

The specific problem we consider is that of localizing multiple targets in a 2D scene using a linear array of radar sensors. Figure 1 shows a scenario with two targets being observed with a linear array of four spatially separated sensors positioned along x-axis. Each sensor collects the relative range and Doppler observations for the targets in the scene.

Since these observations are not ordered *a priori*, each range-Doppler measurement must first be associated with a target, and then the measurements associated with a given target from multiple sensors can be used to estimate its position and velocity. Since the number of possible associations grows exponentially in the number of sensors, it is critical to develop efficient algorithms for spatial association. It is also important to build in robustness to missed detections, since millimeter waves can be easily occluded by objects in the scene.
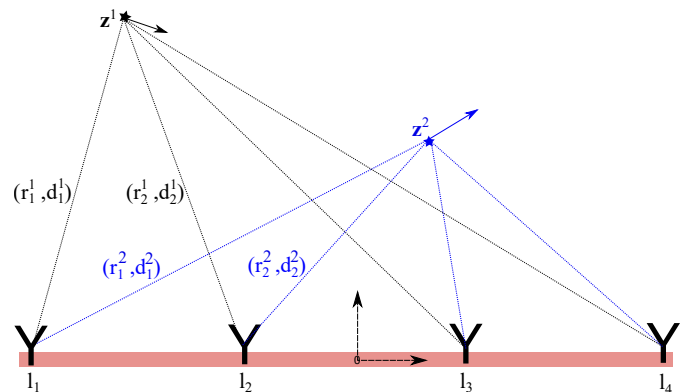

Fig. 1. 2D System model with a linear array of radar sensors placed on x-coordinates, $[l_1, l_2, l_3, l_4]$. The kinematic states z_1, z_2 of two targets are to be estimated using the unordered range and doppler observations from the sensors.

A. Contributions

Our goal is to develop robust and computationally efficient algorithms for single snapshot spatial data association. The main contributions of our study are as follows:
(1) We examine the geometric relations between instantaneous range, Doppler, and sensor locations, and show that features obtained via those geometric relations simplify the association problem. Specifically, we observe and exploit linear relationships between functions of the range-Doppler observations for a target across the linear array of sensors.
(2) We provide a low-complexity solution for the association problem by introducing a new graph-search based algorithm which prunes the set of feasible associations based on geometric relationships. In particular, our proposed algorithm considers a cost function based on the linear geometric relationships together with the triangle inequality constraint for the range observations at pairs of sensors and eliminates a significant number of possible associations. In addition, our approach accounts for detection anomalies such as missed detections

and false alarms while reducing the complexity.

(3) We compare our proposed algorithm against conventional algorithms in the literature and evaluate performance in terms of localization accuracy, cardinality errors, robustness, and complexity. Also, we show that using an enhanced accuracy estimation algorithm (i.e., NOMP [3]) instead of conventional FFT-based approach improves localization accuracy and reduces association complexity as the number of targets and sensors increases.

B. Related Work

The majority of prior work addresses *temporal* data association, which focuses on association of new measurement with existing target tracks. A number of techniques have been proposed in this regard, including Random Finite Set based sampling methods [4], fuzzy clustering [5], and convex optimization [6]. These methods rely on the temporal continuity of target state to assist in associating observations across multiple time frames. Most of these methods are designed for a single sensor case, and extensions to multiple sensor settings are not well-known. In this paper, our focus is on *spatial* association, where the data from multiple spatially separated sensors needs to be associated within the same time frame. While the problem of spatial association studied in this paper has received relatively less attention, we provide a brief overview of the most widely used algorithms in the literature that can be extended for the spatial problem.

The association problem between a pair of sensors can be optimally solved using the well-known Hungarian algorithm [7]. However, a naive extension to multiple sensors by factorizing into pairwise (2D) associations over consecutive sensors does not work well in the presence of detection anomalies such as miss, false alarm, clutter, and close-target interactions [8].

The multi-sensor association problem can be formulated as the Maximum A-Posteriori (MAP) estimation of most likely chain of observations across sensors. In order to solve this problem, a graphical model is defined, where a node represents sensor detection and an edge between nodes represents association hypothesis with a certain probability [9]. Association between sensors is obtained by solving the Minimum Cost Maximum Flow (MCF) problem over this graph. A variety of methods such as Linear Programming [10], Dynamic Programming [11], [12], and push-relabel maximum flow [9] have been proposed to efficiently solve the MCF problem. Although those methods solve the optimization in polynomial time, they require specialized mechanisms such as expansion of the observation set over successive iterations to resolve detection anomalies. Moreover, the complexity of the MCF problem grows quickly as $O(N^3 \log N)$, where N is the number of sensors [9]. In comparison with prior work, our approach reduces complexity by leveraging the high accuracy of sensor observations and their geometric properties.

Probabilistic approaches such as the gated Nearest Neighbor (NN) [13] method sequentially associate observations across the sensors. At each sensor, each observation is associated with its closest match to the state predicted by the chain of observations from the past sensors. However, using only the single most likely observation to form association is vulnerable to clutter and anomalies in noisy scenarios. In addition, a single association error can cause significant contamination in final state estimate. This problem is well known in the literature on Simultaneous Localization and Mapping (SLAM), and various improvements such as Multiple Hypothesis Tracking [14], K-best assignment [15], and JPDAF [13] have been proposed. In contrast, we propose an alternative search approach based on geometric fitting criteria which do not depend on such probabilistic models and avoid the contamination of state.

Bottom up approaches based on grid search over a set of candidate target states have been suggested in the literature [16]. In [17], an approach based on enumerating all possible candidates followed by pruning and merging shows promising results. Randomized adaptive search procedures such as random consensus sampling (RANSAC) [18], Interpretation Tree [19], Joint Compatibility Branch and Bound [20] have been shown to mitigate the impact of detection anomalies. These methods utilize a suitably defined metric to check the consistency of a set of associated observations, and employ branch and bound type search strategies to reduce the search complexity. Our graphical approach uses similar pruning techniques to perform the graph search, but with the additional use of geometric constraints and a geometric fitting error metric for guiding the search.

Outline: The rest of the paper is organized as follows. In Section II, we introduce the association problem in the single snapshot localization setting. In Section III, our graph association algorithm is presented. Then, the proposed algorithm is evaluated over different system parameters in Section IV and Section V concludes the paper.

Notation: $a, \boldsymbol{a}, A, \mathcal{A}$ represent scalar, vector, matrix and set respectively. We use [.] to construct vector, matrix and {.} to construct set. $\times, \cup, \cap$ denote the cartesian product, union and, intersection of two sets and $\varnothing$ denotes a *NULL* value. $n(\mathcal{A})$ represents the number of non-empty elements in set $\mathcal{A}$. $\circ$ denotes element-wise multiplication between vectors. A^T denotes transpose of matrix A and $\wedge$ denotes logical "and" operator.

II. PROBLEM DESCRIPTION

A. System Model

Consider a linear array of N_S radar sensors in a two-dimensional (2D) scene with N_T targets as in Figure 1. Without loss of generality, we assume that the sensor array is static and located along x-axis and centered at origin. The absolute kinematic state of the targets can be obtained by using the target location relative to this sensor array along with its own odometer information.

The kinematic state (i.e., instantaneous position and velocity information of all targets) of the scene is given by

$$\mathcal{Z} = \{\boldsymbol{z}^k\}_{k=1}^{N_T}$$

where $\boldsymbol{z}^k = (x^k, y^k, v_x^k, v_y^k)$ is the kinematic state of target k with an instantaneous velocity of (v_x^k, v_y^k) at position (x^k, y^k).

The range-Doppler of target k observed at sensor i, can be expressed in terms of the desired kinematic state as follows,

$$r_i^k = \sqrt{(x^k - l_i)^2 + (y^k)^2}, \quad d_i^k = \frac{(x^k - l_i)v_x^k + y^k v_y^k}{r_i^k}. \tag{1}$$

where l_i is the x-coordinate of sensor i. We denote this non-linear mapping as $(r_i^k, d_i^k) = \mathcal{T}_i(\boldsymbol{z}^k)$.

B. Single Snapshot Localization

In order to extract range and Doppler information of target k, each sensor i uses the signal (i.e., $m_i^{obs}(t)$) reflected back from the scene in monostatic mode. In this study, we focus on localization of the scene using a single snapshot. For that reason, the kinematic state of the scene is assumed to be constant for a certain time interval and the scene localization is performed based on the range and Doppler information gathered during that time interval. Based on those, the Maximum Likelihood Estimator (MLE) for the scene including all N_S sensors can be expressed as,

$$\hat{\mathcal{Z}}_{ML} = arg \max_{\mathcal{Z}} \prod_{i=1}^{Ns} \mathcal{L}\left(m_i^{obs} | \mathcal{T}_i(\mathcal{Z})\right) \tag{2}$$

where m_i^{obs} corresponds to the observed signal in a single snapshot and $\mathcal{L}\left(m_i^{obs} | \mathcal{T}_i(\mathcal{Z})\right)$ is the conditional log likelihood of the observed signal for scene $\mathcal{Z}$.

The optimization problem in (2) is difficult in general since the number of targets (i.e., N_T) is not known and a brute force search for $\mathcal{Z}$ incurs exponential complexity in the number of targets; that is, $n(\mathcal{D}(\boldsymbol{z}))^{N_T}$ for a grid $\mathcal{D}(\boldsymbol{z})$. In addition, the observations contain a variety of anomalies such as clutter, missed detections, and false alarms, which further complicates the solution.

In order to facilitate the solution of the problem in (2), the problem is divided into two stages as follows:

1) Estimation: The Range-Doppler pairs of $M_i \leq N_T$ non-occluded targets are estimated from received signal m_i^{obs} at sensor i using efficient algorithms proposed in the literature [21]. The estimate at sensor i for k^{th} target can be modeled as follows,

$$(r_i^k) = (r_i^k)^{true} + w_i^R + \tilde{b}_i^k, \tag{3a}$$
$$(d_i^k) = (d_i^k)^{true} + w_i^D + \bar{b}_i^k \tag{3b}$$

where $w_i^R \sim \mathcal{N}(0, \sigma_{r_i}^2)$ and $w_i^D \sim \mathcal{N}(0, \sigma_{d_i}^2)$ denote independent Gaussian distributed noises with zero mean and $\tilde{b}_i^k$ and $\bar{b}_i^k$ denote the bias errors introduced due to proximity with any other $M_i - 1$ targets in the scene. The noise variance depends on estimation accuracy at the given SNR which, in turn, depends on target radar cross section (RCS), path loss, and antenna directivity. For simplicity, we assume equal received signal power across all targets in the scene.

We denote the set of estimated range-Doppler pairs at sensor i by $\Theta_i = \{\bigcup_{k=1}^{M_i} \boldsymbol{\theta}_i^k\}$ where $\boldsymbol{\theta}_i^k = \left[(r_i^j), (d_i^j)\right]^T$. Index $k \in [1, M_i]$ in θ_i^k denotes the index of kth measurement with respect to the M_i measurements for sensor i, whereas, index j in $[(r_i^j), (d_i^j)]^T$ denotes the global index of the jth target.

The different superscripts are used to highlight the fact that the order of targets for which the range-Doppler measurements are obtained at the sensors is unknown. Indeed, our aim in this work is to find the correct ordering/association of the range-Doppler measurements.

2) Association problem: The estimation of kinematic state $\mathcal{Z}$ requires the association of those un-ordered range-Doppler pairs, Θ_i, collected across all sensors. An association chain is defined as the ordered set of range-Doppler observations, $\mathcal{A} : \{\{\boldsymbol{\theta}_i\}_{i=1}^{N_S} | \boldsymbol{\theta}_i \in \tilde{\Theta}_i\}$ which is constructed from the *NULL* augmented sets; that is, $\tilde{\Theta}_i = \Theta_i \cup \varnothing$. $\boldsymbol{\theta}_i = \varnothing$ corresponds to the *NULL* state, which represents absence of observation at sensor i (e.g., due to missed detection). Figure 2 shows a graphical representation of an association problem with three targets observed using $N_S = 4$ sensors. Sensors 1 and 2 observe all targets $M_1 = M_2 = 3$, sensor 3 misses target $\boldsymbol{z}^2$ and sensor 4 has a false alarm. The desired association chain for target $\boldsymbol{z}^1$ across four sensors is shown in the shaded region.

The spatial association problem can be formulated as the following maximum a posteriori (MAP) estimation problem,

$$\boldsymbol{\mathcal{A}}^* = \underset{\boldsymbol{\mathcal{A}} \subset \tilde{\Theta}_1 \times \cdots \times \tilde{\Theta}_{N_S}}{argmax} \log P(\boldsymbol{\mathcal{A}}) P(\Theta | \boldsymbol{\mathcal{A}}) \tag{4}$$

$$\text{such that} \quad \mathcal{A}^i \cap \mathcal{A}^j = \varnothing \quad \forall i \neq j, \quad n(\mathcal{A}^k) \geq 2$$

where $\Theta = \bigcup_{i=1}^{N_S} \Theta_i$ denotes the set of all range-Doppler observations, $\boldsymbol{\mathcal{A}} = \{\mathcal{A}^1, \mathcal{A}^2, \cdots\}$ denotes a subset of association chains chosen from the set of all possible potential chains, $\tilde{\Theta}_1 \times \tilde{\Theta}_2 \times \cdots \tilde{\Theta}_{N_S}$. The optimal solution $\boldsymbol{\mathcal{A}}^*$ consists of the set of chains which jointly maximizes overall log likelihood while the constraints ensure that no two chains share a common observation and each chain contains at least two observations.

When the targets are well-separated, the bias terms in (3a) and (3b) vanish and the likelihood for the individual targets becomes independent across multiple targets. In this case, the log likelihood in (4) simplifies to

$$\log P(\boldsymbol{\mathcal{A}}) P(\Theta | \boldsymbol{\mathcal{A}}) = \sum_{\mathcal{A} \in \boldsymbol{\mathcal{A}}} \log P(\mathcal{A}) + \log P(\Theta | \mathcal{A})$$

where $P(\Theta | \mathcal{A}) = \prod_{i=1}^{N_S} P(\Theta_i | \mathcal{A})$ is the probability of detecting the range-Doppler pairs which can be modeled by a Bernoulli distribution,

$$P(\Theta_i | \mathcal{A}) = \begin{cases} \alpha & \text{, if target missed at sensor } i, \mathcal{A}_i = \varnothing \\ 1 - \alpha & \text{, else} \end{cases}$$

where α denotes the probability of detection errors in (5) and is set to nominal value $\alpha = 0.05$ [9]. This model accounts for the occurrence of both miss and false alarms across the sensors in the likelihood, which is given by

$$P(\Theta | \mathcal{A}) = \alpha^{N_S - n(\mathcal{A})} (1 - \alpha)^{n(\mathcal{A})}.$$

Also, $P(\mathcal{A})$ is the likelihood of chain modeled using the perceived range-Doppler pairs, $(\hat{r}_i, \hat{d}_i) = \mathcal{T}_i(\hat{\boldsymbol{z}})$ for a target state $\hat{\boldsymbol{z}}_k$ predicted by the chain (see Section III-B2). By

ignoring the constant terms which preserve the MAP solution, we define the normalized negative log likelihood as follows,

$$\mathcal{L}(\mathcal{A}) = \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} \left(\frac{(\hat{r}_i - r_i)^2}{\sigma_r^2} + \frac{(\hat{d}_i - d_i)^2}{\sigma_d^2} \right) + n(\mathcal{A}) \log \frac{\alpha}{1 - \alpha} \tag{5}$$

where $\boldsymbol{\theta}_i = [r_i, d_i]^T$ is the observation from i^{th} sensor in association chain $\mathcal{A}$ and σ_r^2 and σ_d^2 are the nominal variance terms for range and Doppler, respectively (see Appendix A for details). The first term in (5) denotes the squared error between the estimated and observed range-Doppler pairs in the chain while the second term penalizes the selection of smaller chains which prevents formation of duplicate chains for the same target. Hence, the association problem is reduced to the following constrained minimization problem:

$$\mathcal{A}^* = \operatorname*{argmin}_{\mathcal{A} \subset \tilde{\Theta}_1 \times \cdots \times \tilde{\Theta}_{N_S}} \sum_{\mathcal{A} \in \mathcal{A}} \mathcal{L}(\mathcal{A}) \tag{6}$$
$$\text{such that} \quad \mathcal{A}^i \cap \mathcal{A}^j = \varnothing \; \forall i \neq j, \quad n(\mathcal{A}) \geq 2$$

The joint minimization problem over all potential association chains in (6) is difficult in general. For that reason, we use an iterative approach where the most likely chains of observations are identified and removed from observation set Θ sequentially,

$$\operatorname*{argmin}_{\mathcal{A} \in \tilde{\Theta}_1 \times \cdots \times \tilde{\Theta}_{N_S}} \mathcal{L}(\mathcal{A}) \quad \text{such that} \quad n(\mathcal{A}) \geq 2 \,. \tag{7}$$

Without any prior knowledge of association between the nodes, the number of potential chains $\Theta_1 \times \Theta_2 \times \cdots \Theta_{N_S}$ still grows exponentially. However, the formulation in (7) enables the utilization of various network optimization methods to identify the most likely chain. Once the associated chains of range-Doppler observations are found across sensors, the kinematic state of the scene can be easily obtained by solving the inverse kinematic problem $[\hat{x}, \hat{y}, \hat{v}_x, \hat{v}_y] = \mathcal{T}^{-1}(\mathcal{A})$ using Gauss-Newton algorithm [21].

III. GRAPHICAL ASSOCIATION

In order to solve the association problem in (7), we formulate the spatial association problem using graphical models and present our low-complexity graphical search method to obtain association chains efficiently via geometric relations.

A. Graph Generation

To begin with, we define a target-based graph to perform data association with following elements:

- Node $\boldsymbol{\theta}_i^k$ represents the kth range-Doppler pair at sensor i. Nodes for a given sensor are arranged along a single column of the graph as shown in Figure 2.
- Edge $e_{ij}^{kl} = [\boldsymbol{\theta}_i^k, \boldsymbol{\theta}_j^l]$ denotes the linkage between pairs of observation across sensor i and sensor j, which can correspond to a feasible target z_{ij}^{kl} referred to as "*candidate*" location.
- Chain $\mathcal{A}^j$ is represented by the sequence of two or more nodes spanning distinct sensors, which is associated to a single target, $\hat{z}^j$.

Geometric Constraint: A significant portion of the edges can be easily discarded in the graph generation phase by using the following geometric constraint on target's range (for noiseless case),

$$C_{\text{G}}(e_{ij}) : \quad (r_i - r_j < l_{ij}) \wedge (r_i + r_j > l_{ij}) \tag{8}$$

where $l_{ij} = |l_i - l_j|$ represents the separation between sensor i and sensor j.

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is initialized with vertices for all range-Doppler pairs $\mathcal{V} = \{\Theta_i\}_{i=1}^{N_S}$ and edges $\mathcal{E}$ between any two consecutive nodes that satisfy condition $C_{\text{G}}(e_{i-1,i}^{k,l}), \forall k \in [1, M_{i-1}], \forall l \in [1, M_i]$ for all $i \in \{2, \ldots, N_S\}$ given in (8).

B. Spatial Association using Geometric Features

In this subsection, we describe the solution of the association problem presented in (7) using graph $\mathcal{G}$ by exploiting geometric relations between range, Doppler, and sensor geometry. For clarity of exposition, we focus on the association procedure of a single target $z = [x, y, v_x, v_y]$ and, therefore, drop the superscript k for the sake of simplicity.

1) Geometric Relations: The range of target observed at i^{th} sensor is given by

$$r_i = \sqrt{(x - l_i)^2 + (y)^2} \,. \tag{9}$$

The Doppler component is the rate of change of range and it is given by,

$$d_i = \dot{r}_i = \frac{(x - l_i)v_x + yv_y}{r_i}$$
$$r_i d_i = (x - l_i)(v_x) + yv_y \,. \tag{10}$$

For a linear array of sensors, the range and Doppler measurements for a target satisfy the following relations based on (9) and (10):

$$r_i^2 = r_j^2 - 2x(l_i - l_j) + (l_i^2 - l_j^2) \tag{11a}$$
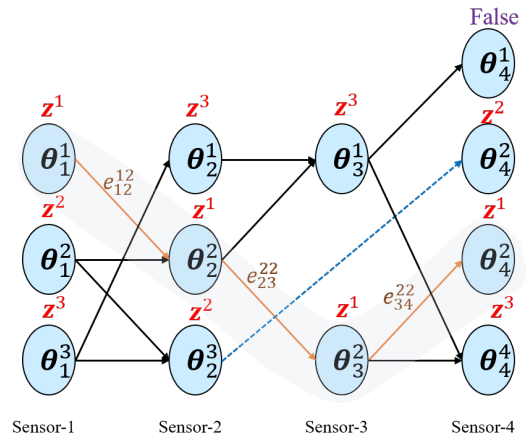$$r_i d_i = r_j d_j - (v_x)(l_i - l_j) \tag{11b}$$

Fig. 2. Target-based observation graph for a scene with 3 targets and 4 sensors. Sensors $1, 2$ observe all 3 targets in different orders. Sensor 3 misses the observation of target state z^2 while sensor 4 contains a false observation. Desired association chain, $\mathcal{A}$ is shown by the shaded set of nodes.

where r_i (r_j) and d_i (d_j) are the range and Doppler estimated at the ith (j^{th}) sensor, respectively. l_i (l_j) is the x-coordinate of ith (j^{th}) sensor. (9) and (10) indicate that for the noiseless setting, the range-Doppler products and range squared are linear with respect to target's velocity and position at x-coordinate, respectively. Therefore, the correct associations can be identified by fitting the observations to those geometric relations.

2) State Prediction and Fitting Error: The presence of noise in (r_i, d_i) causes high error in these geometric relations due to the quadratic dependence. An estimate of target state parameters $\hat{x}, \hat{v}_x$ can be obtained by minimizing that error between observed and predicted range and Doppler values. Let $\boldsymbol{q}_1 = [r_i d_i | (r_i, d_i) \in \mathcal{A}]$ and $\boldsymbol{l} = [l_i | \boldsymbol{\theta}_i \in \mathcal{A}]$ denote the vector of range-Doppler products using observations in chain $\mathcal{A}$ and the vector of corresponding sensor x-coordinates, respectively. Predicted fit $\hat{\boldsymbol{q}}_1$ can be expressed using the geometric relation in (11b) as follows:

$$\hat{\boldsymbol{q}}_1 = -v_x \boldsymbol{l} + \kappa_1 \mathbb{1} = H \boldsymbol{s}_1$$

where $H = [\boldsymbol{l}, \mathbb{1}]$, $\boldsymbol{s}_1 = [-v_x \; \kappa_1]^T$, and κ_1 is a constant. Then, the least squares estimate for $\hat{\boldsymbol{s}}_1$ is obtained as

$$\hat{\boldsymbol{s}}_1 = arg \min_{\boldsymbol{s}_1} \quad \|\boldsymbol{q}_1 - H\boldsymbol{s}_1\|^2 \qquad (12)$$
$$= (H^T H)^{-1} H^T \boldsymbol{q}_1.$$

Therefore, the least squares estimate is obtained as $\hat{v}_x = \boldsymbol{u}^T \boldsymbol{q}_1$ where $\boldsymbol{u} = -H(H^T H)^{-1} \boldsymbol{e}_1$ and $\boldsymbol{e}_1 = [1, 0]^T$.

Similarly, let $\boldsymbol{q}_2 = [r_i^2 | (r_i, d_i) \in \mathcal{A}]$ denote the vector of range squared observations in chain $\mathcal{A}$, predicted fit $\hat{\boldsymbol{q}}_2$ can be expressed using the geometric relation in (11a) as follows:

$$\hat{\boldsymbol{q}}_2 - \boldsymbol{l} \circ \boldsymbol{l} = -2x\boldsymbol{l} + \kappa_2 \mathbb{1} = H \boldsymbol{s}_2$$

where $\boldsymbol{s}_2 = [-2x \; \kappa_2]^T$ and κ_2 is a constant. The least squares estimate of $\hat{x}$ is obtained as

$$\hat{\boldsymbol{s}}_2 = arg \min_{\boldsymbol{s}_2} \quad \|\boldsymbol{q}_2 - \boldsymbol{l} \circ \boldsymbol{l} - H\boldsymbol{s}_2\|^2 \qquad (13)$$
$$= (H^T H)^{-1} H^T (\boldsymbol{q}_2 - \boldsymbol{l} \circ \boldsymbol{l}).$$

Hence, we obtain $\hat{x} = \boldsymbol{u}^T (\boldsymbol{q}_2 - \boldsymbol{l} \circ \boldsymbol{l})/2$.

The remaining state parameters (i.e., $\hat{y}$ and $\hat{v}_y$) are obtained using the geometric relations in (9) and (10) as

$$\hat{y} = \sqrt{\frac{1}{n(\mathcal{A})} \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} (r_i^2 - (\hat{x} - l_i)^2)},$$

$$\hat{v}_y = \frac{1}{n(\mathcal{A})} \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} \frac{r_i d_i - (\hat{x} - l_i)\hat{v}_x}{\hat{y}}.$$

The normalized geometric fitting error of a chain $\mathcal{A}$ can be computed using these estimates as follows:

$$F(\mathcal{A}) = \frac{\|\boldsymbol{q}_1 - \hat{\boldsymbol{q}}_1\|^2}{\eta_1} + \frac{\|\boldsymbol{q}_2 - \hat{\boldsymbol{q}}_2\|^2}{\eta_2} \qquad (14)$$

$$= \frac{\left\| \left(I - H(H^T H)^{-1} H^T\right) \boldsymbol{q}_1 \right\|^2}{\eta_1} \qquad (15)$$

$$+ \frac{\left\| \left(I - H^T(H^T H)^{-1} H\right) (\boldsymbol{q}_2 - \boldsymbol{l} \circ \boldsymbol{l}) \right\|^2}{\eta_2}$$

where η_1 and η_2 are normalization constants that are set based on CRB (see Appendix C for details) and (15) is obtained by substituting the predicted fits into (14). It is important to note that the error in (15) is additive over the observations in chain $\mathcal{A}$. Therefore, the extension of the chain cannot reduce the fitting error. In other words, $F(\mathcal{A})$ is monotonically non-decreasing over the length of chain $\mathcal{A}$. For that reason, the fitting error provides a simple measure of the geometric consistency of a chain, which can be used to traverse the graph and extract the chains efficiently.

3) Geometric Association: We now present a graph search procedure which obtains the associated chains by minimizing geometric fitting error $F(\mathcal{A})$ in (15) and negative log likelihood $\mathcal{L}(\mathcal{A})$ in (5). We apply the geometric relations by adding constraints on the desired chain, $\mathcal{A}$ to the optimization problem in (7) as follows,

$$\min_{\mathcal{A} \in \tilde{\Theta}_1 \times \cdots \times \tilde{\Theta}_{N_S}} \mathcal{L}(\mathcal{A})$$
$$\text{such that} \quad n(\mathcal{A}) \geq \gamma, \qquad (16a)$$
$$F(\mathcal{A}) < \tau_f^{n(\mathcal{A})} \qquad (16b)$$

The constraint in (16a) restricts the number of missed observations to be less than $N_S - \gamma$ and the constraint in (16b) only allows chains with good geometric fit to be selected. In order to provide a solution for the optimization problem in (16), we perform Depth First Search (DFS) over the graph generated in Section III-A to extract the chains, where those additional constraints help in reducing the search complexity. Our complete Spatial Association using Geometry Algorithm (SAGA) is outlined in Algorithm 1. Here is a brief description:

1) We start the graph search by setting $\gamma = N_S$ so that only chains that include observations from all sensors are extracted. For that reason, we consider a graph having edges between consecutive sensors only. This helps to reduce the chains encountered during initial DFS procedure (see Appendix E for details).

2) The DFS is guided by geometric fitting error $F(\mathcal{A})$. After each node is visited, the fitting error of candidate chain is calculated and the chain is ignored if it has a fitting error higher than predefined threshold $\tau_f^{N_S}$. Since the fitting error is non-decreasing over the length of the chain, most of the candidate chains are eliminated before reaching at the end of the graph, which reduces the complexity further. Details of DFS are shown in Appendix D. At the termination of the DFS, the corresponding chain of nodes is added to solution $\mathcal{A}^{\dagger}$ if it satisfies all the constraints in (16) and the negative log-likelihood of the association chain is below a predefined threshold (i.e., $\mathcal{L}(\mathcal{A}) < \tau_l^{n(\mathcal{A})}$). The nodes belonging to the selected chains are removed from the graph together with their corresponding edges to keep subsequent chains disjoint.

3) In order to deal with missed detection cases at sensors, the minimum chain length constraint (i.e.,γ) is relaxed in steps upto robustness level ρ. Due to that relaxation, the graph includes not only the edges between consecutive sensors but also the edges among the nodes that skip over h consecutive sensors. Those edges are called *Skip-*

h edges where $h = N_S - \gamma$. Then, the DFS procedure is repeated for different minimum chain length constraints. Consequently, in this procedure, *NULL* states are taken into account and the generated chain does not include any observation from a sensor that misses the corresponding target by skipping over the observations of that sensor via *Skip-h* edges. In addition, the DFS procedure implicitly accounts for *NULL* state in the beginning and end of a chain by starting searching from different nodes in consideration of minimum chain length constraint.

4) The thresholds (i.e., τ_f and τ_l) for the geometric fitting error and the likelihood depend on length of the chain $n(\mathcal{A})$ and their initial value is set based on CFAR criteria (see Appendix C for details). Using a tight initial threshold τ_f for $F(\mathcal{A})$ restricts the number of branches to be explored at each node to a smaller set. This reduces the initial complexity of DFS while allowing only a subset of association chains $\mathcal{A}^{\dagger} \subset \mathcal{A}^*$ to be found. The thresholds are later relaxed by a factor of $\beta > 1$ to allow the observations contaminated with noise to be selected. The relaxation is stopped when no further chains with length $n(\mathcal{A}) \geq N_S - \rho$ exist in the graph.

Algorithm 1 Spatial Association using Geometric Assistance

> **Input:** Graph $\mathcal{G}$, Robustness level ρ
> 1: INITIALIZE Chains $\mathcal{A}^{\dagger} = \emptyset$, $\boldsymbol{\tau} = [\tau_f, \tau_l]_{init}$
> 2: **repeat**
> 3:   REMOVE all Skip edges
> 4:   **for** $h = 0$ to ρ **do**
> 5:   Set minimum chain length: $\gamma = N_S - h$
> 6:   ADD SKIP-h EDGES TO GRAPH $\mathcal{G}$
> 7:   **for** $v \in \mathcal{V}$ **do**
> 8:   DFS from node v: $\mathcal{A} \leftarrow GA\text{-}DFS(v, \gamma, \boldsymbol{\tau})$
> 9:   **if** Valid Chain, $\mathcal{A}$ is found **then**
> 10: $\mathcal{A}^{\dagger} \leftarrow \mathcal{A}$
> 11:   Remove chain from graph $\mathcal{V} = \mathcal{V} - \{\mathcal{A}\}$
> 12:   **end if**
> 13:   **end for**
> 14:   **end for**
> 15:   Relax thresholds: $\boldsymbol{\tau} \leftarrow \beta\tau$
> 16: **until** Chains with length $n(\mathcal{A}) \geq N_S - \rho$ exists in $\mathcal{G}$
> **Output:** Selected chains $\mathcal{A}^{\dagger}$

Robustness: During chain length relaxation, a *Skip* edge is added between the observations across sensor i and sensor q if

1) Observations $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_q$ satisfy the geometric constraint $C_{\mathrm{G}}(e_{iq})$ in (8), and,
2) The target state predicted by $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_q$ differs by a predefined threshold τ_z from the ones predicted by using all observations on the paths that connect $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_q$.

$$C_{\mathrm{S}}(e_{iq}): \quad \left\| \hat{\boldsymbol{z}}_{\mathcal{A}_p} - \hat{\boldsymbol{z}}_{iq} \right\| > \tau_z, \forall \mathcal{A}_p : \{\boldsymbol{\theta}_i, \boldsymbol{\theta}_q\} \in \mathcal{A}_p \tag{17}$$

where $\mathcal{A}_p$ is in the form of $\mathcal{A}_p = \{\boldsymbol{\theta}_i, \boldsymbol{\theta}_j, \cdots, \boldsymbol{\theta}_q\}$ with $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_q$ at the edges of the path, τ_z is set based on CRB (see Appendix B), $\hat{\boldsymbol{z}}_{iq}$ indicates the predicted target

state based on $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_q$, and $\hat{\boldsymbol{z}}_{\mathcal{A}_p}$ shows the predicted target state using the observations in $\mathcal{A}_p$.

Enforcing the condition in (17) avoids the formation of multiple chains corresponding to the same target and avoids unnecessary increase in the number of edges. The number of skip connections introduced in the graph is controlled by the robustness level; that is, $0 \leq \rho \leq (N_S - 2)$, which sets the maximum number of missed detections that can be tolerated across the sensor array. In this way, addition of such edges provides a flexible mechanism to provide robustness against missed detection in the sensors while keeping search space in control.

Complexity: The non-decreasing property of $F(\mathcal{A})$ is used to discard unlikely chains in the early stages of DFS. This allows for rapid extraction of associations without requiring search over all possible chains in the graph. The minimum track length threshold, γ, is reset to its maximum value after each relaxation. Therefore, the skip edges in the graph can be removed at the end of the inner loop to reduce search complexity further. Therefore, our approach exploits the geometric structure of observations across multiple sensors to reduce search complexity.

C. Spatial Association using Edge-based State Likelihoods

Before evaluating the performance of our main algorithm, we describe an iterative search method, which relies on the fact that an approximate kinematic state estimate can be derived by using two connected observations in a graph. In other words, a state estimate can be obtained for each edge in a graph, which is a part of the association chain $\mathcal{A}$. Therefore, the search space for the association problem in (7) can be reduced to the set of edges.

The likelihood of a candidate $\boldsymbol{z}_e$ corresponding to an edge $e \in \mathcal{E}$ can be computed as,

$$\mathcal{L}(\boldsymbol{z}_e) = \sum_{i=1}^{N_S} \left[\min_{\boldsymbol{\theta} \in \Theta_i} \left(\frac{(r_i' - r_i)^2}{\sigma_r^2} + \frac{(d_i' - d_i)^2}{\sigma_d^2} \right) \right] \tag{18}$$

where $[r_i', d_i'] = \mathcal{T}_i(\boldsymbol{z}_e)$ is the perceived range and Doppler at sensor i for target state $\boldsymbol{z}_e$. Then, the most likely candidate can be selected by evaluating (18) over all edges and choosing the one that achieves the minimum negative log likelihood; that is, $\boldsymbol{z}^* = \boldsymbol{z}_{e^*}$ for $e^* = arg \min_{e \in \mathcal{E}} \mathcal{L}(\boldsymbol{z}_e)$. Then, the observations associated with $\boldsymbol{z}^*$ can be identified via the following neighborhood constraint:

$$\mathcal{N}(\boldsymbol{z}^*) = \bigcup_{i=1}^{N_S} \{(r_i, d_i) | (r_i - r_i^*) \leq \delta_r \wedge (d_i - d_i^*) \leq \delta_d\}$$

where $[r_i^*, d_i^*] = \mathcal{T}_i(\boldsymbol{z}^*)$ are the perceived range-Doppler at sensor i and δ_r and δ_d are the range and Doppler resolution parameters defined in Appendix A. The algorithm carrying out this Spatial Association using Edge-based State Likelihoods (*SAESL*) procedure is presented in Algorithm 2.

Since all edges in the graph are checked while selecting the candidates, this approach exhibits higher complexity than our proposed algorithm. Moreover, evaluation of state likelihood $\mathcal{L}(\boldsymbol{z}_e)$ in (18) is more expensive than evaluation of chain

Algorithm 2 SAESL Algorithm

1: INITIALIZE GRAPH WITH OBSERVATIONS Θ: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
2: AUGMENT GRAPH with skip edges
3: **for** $h = 0$ to ρ **do**
4:     ADD SKIP-h EDGES TO GRAPH $\mathcal{G}$
5: **end for**
6: INITIALIZE $\mathcal{Z} = \emptyset$
7: **while** $\mathcal{E} \neq \emptyset$ **do**
8:     FIND MOST LIKELY CANDIDATE, $\mathcal{Z} \leftarrow z^*$
    from edge $z^* = arg\min_{e \in \mathcal{E}} \mathcal{L}(z_e)$
9:     REMOVE ALL VERTICES EXPLAINED BY z^*,
    $\mathcal{V} \leftarrow \mathcal{V} - \mathcal{N}(z^*)$
10:     UPDATE EDGES $\mathcal{E}$
11: **end while**
12: RETURN Selected candidates $\mathcal{Z}$

---

likelihood $\mathcal{L}(\mathcal{A})$ in (5) as it involves a minimization over all other observations. In Section IV, we use this algorithm as a benchmark against our proposed algorithm.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed spatial association algorithm, *SAGA* against the *SAESL* algorithm through various performance metrics. We consider a linear array of N_S FMCW radar sensors each of which collects range and Doppler observations from the scene. The FMCW radar system parameters are set based on the ones that are used in typical low cost automotive systems at mm-Wave frequencies [22]: bandwidth $B = 0.5$ GHz, carrier frequency $f_c = 60$ GHz, chirp duration $78\mu s$, $N_{ch} = 64$ chirps, and sampling rate $f_s = 0.82$ MHz. This provides range and Doppler resolutions of $\delta r = 0.3$ and $\delta d = 0.5 m/s$, respectively, and maximum range and Doppler of $19.2m$ and $\pm 16 m/s$, respectively, suitable for short range situational awareness applications. In the simulations, a single snapshot of the scene is considered with multiple targets having equal received SNR at all sensors. The kinematic states of targets are randomly selected based on uniform distributions $x \sim \mathcal{U}(-8m, 8m), y \sim \mathcal{U}(2m, 12m), v_x \sim \mathcal{U}(-10m/s, 10m/s), v_y \sim \mathcal{U}(-10m/s, 10m/s)$.

It is important to note that when range and Doppler separation between two targets gets small, the estimation algorithm either provides a merged estimate or results in detection anomalies such as miss and false alarm. In order to differentiate the scenes with such estimation errors due to range-Doppler proximity, we consider two different scenarios with two different scenes. The *well-separated* scene is generated by enforcing a minimum separation between the range and Doppler of the targets at all sensors. The *adverse* scene does not have such constraints and contains additional missed detection anomalies by randomly removing measurements from the sensors with probability P_{miss}. Unless stated otherwise, the nominal values of system parameters are presented in Table I.

A. Localization Accuracy

In this subsection, we analyze the localization accuracy of kinematic state estimates obtained using associated sensor

TABLE I
SIMULATION PARAMETERS

Number of targets	$N_T = 20$
Number of radar sensors	$N_S = 6$
SNR	-10 dB
Sensor Array Width	$L_W = 4$ m
Simulated misses	$P_{miss} = 0.05$
Robustness Level	$\rho = 4$
Max error threshold	$d = 0.16$ m

observations. This depends on the accuracy of underlying range-Doppler estimates. The position and velocity estimation errors for state estimates $\hat{\mathcal{Z}}$ are computed as follows:

$$D_p(\hat{\boldsymbol{Z}}) = \frac{1}{n(\hat{\boldsymbol{Z}})} \sum_{\hat{z} \in \hat{\boldsymbol{Z}}} \min_{z \in \boldsymbol{Z}^{\text{true}}} d_p(z, \hat{z})^2$$

$$D_v(\hat{\boldsymbol{Z}}) = \frac{1}{n(\hat{\boldsymbol{Z}})} \sum_{\hat{z} \in \hat{\boldsymbol{Z}}} \min_{z \in \boldsymbol{Z}^{\text{true}}} d_v(z, \hat{z})^2$$

where $d_p(z, z') = \sqrt{(x - x')^2 + (y - y')^2}$ and $d_v(z, z') = \sqrt{(v_x - v_x')^2 + (v_y - v_y')^2}$ are the errors in position and velocity, respectively. The CRBs for Range-Doppler and Position-Velocity estimates are evaluated in Appendix A and Appendix B, respectively. Figure 3 shows the Root Mean Square Error (RMSE) in range-Doppler estimated at sensor level for different number of targets in a *well-separated* case. We observe that range-Doppler RMSE at individual sensors achieves CRB at a SNR= -15 dB threshold. The RMSE for position-velocity estimates obtained from sensor observations also achieve their CRB at the same SNR threshold. This shows that association using *SAGA* does not introduce any additional errors to the localization process when SNR is above this threshold. However, the RMSE increases sharply below the SNR threshold due to the difficulty in associating noisy range-Doppler pairs. Therefore, we use nominal $SNR = -10$ dB in our simulations to perform further analysis.

Cardinality Error and OSPA: For multiple targets, the number of *valid* targets identified by the system is also an important performance metric. An estimated target $\hat{z}$ is classified to be *valid* only if it lies within a region "close" to the true targets, $\min_{z \in \boldsymbol{Z}^{\text{true}}} \|\hat{z} - z\| < \bar{d}$ where $\bar{d}$ sets the maximum error threshold. The cardinality error is defined as the difference between actual number of targets and the number of estimated target; that is, $N_T - N_e = |\boldsymbol{Z}^{\text{true}}| - |\hat{\boldsymbol{Z}}|$. That error is caused due to the detection anomalies in the estimation algorithm at sensor level as well as during the association stage. In such cases, the localization accuracy by itself does not capture the true performance of the system. Therefore, we use the optimal subpattern assignment (OSPA) metric [23], which combines the localization and cardinality error into a single performance metric and is given by

$$\text{OSPA}(\hat{\boldsymbol{Z}}) = \sqrt{\frac{1}{n(\hat{\boldsymbol{Z}})} \left(\sum_{i=1}^{m} \min\left(d_c(\hat{z}_i), \bar{d}\right)^2 + |N_e - N_T|\bar{d}^2 \right)}$$

where m is the number of *valid* targets, $N_e - N_T$ is the cardinality error and, $d_c(\hat{z}_i)$ is the localization error computed
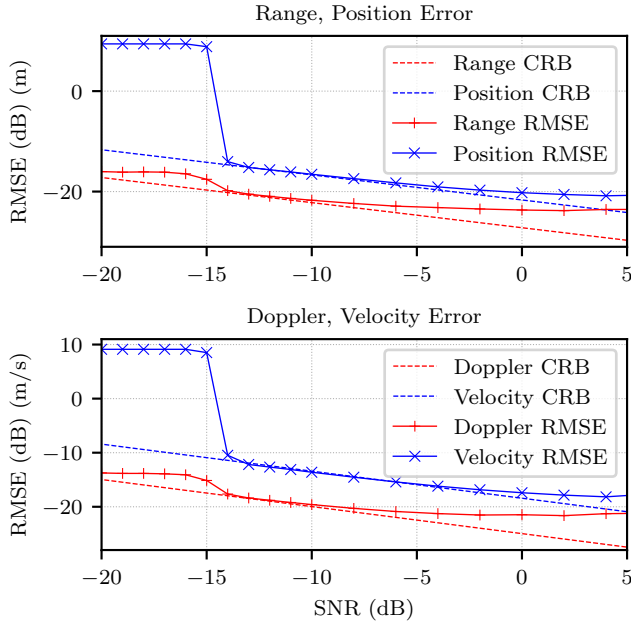
Fig. 3. Range-Dopper estimation accuracy and Position-Velocity estimation accuracy versus SNR. The position-velocity RMSE converges to the CRB bound as SNR increases and the SNR at which this convergence occurs is called as SNR threshold. The SNR threshold provides an indicator for the localization performance when multiple targets, $N_T > 1$ are present.

relative to *closest* true target given as

$$d_c(\hat{z}_i) = \min_{z \in \mathcal{Z}^{\text{true}}} d_p(z, \hat{z}_i)^2 + d_v(z, \hat{z}_i)^2 .$$

Figure 4 shows the OSPA error along with the localization and cardinality errors with increasing scene density in the *well-separated* case. Both localization error and cardinality error start to increase as the scene gets denser. The *SAGA* and *SAESL* schemes have comparable localization error for each target that is validated, but SAGA underestimates the number of targets (i.e., $N_e < N_T$) while SAESL overestimates it (i.e., $N_e > N_T$). Since the localization error is computed only over the reduced set of *valid* targets, we compute the OSPA metric which effectively combines both quantities. The OSPA metric is dominated by localization error when the scene is sparse and cardinality errors when the scene is dense. We observe that *SAGA* has slightly worse overall performance compared to *SAESL* as the number of targets increases at low $SNR = -15$dB. This is because the geometric fit that SAGA relies on is impaired at lower SNR. As we increase SNR to -10dB, the geometric fit and hence performance of SAGA improves, reducing the performance difference with SAESL. Moreover, *SAGA* obtains the association with significantly lower complexity than *SAESL*, as discussed in the next section.

B. Complexity Reduction

In this section, we analyze the computational savings achieved by the proposed *SAGA* algorithm and provide comparison against traditional approaches. Figure 5 shows that the geometric pruning criteria in (8) reduce the number of edges in the graph by an order of magnitude as N_T increases. Also, the worst case complexity of *SAGA* can be expressed in terms
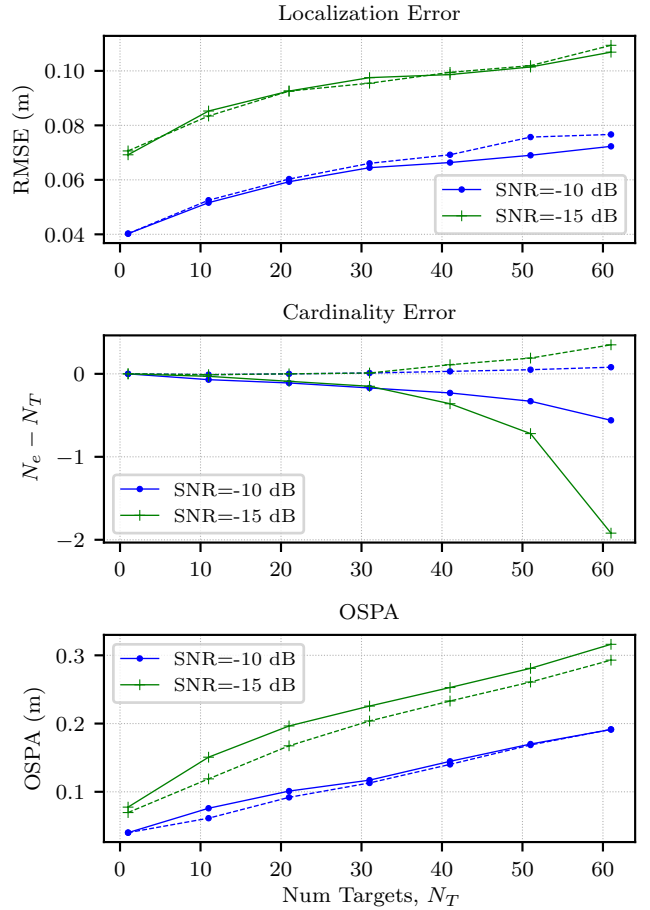

Fig. 4. Overall localization accuracy versus number of targets at $SNR = -15, -10$ dB. The solid and dotted lines represent the performances of *SAGA* and *SAESL* association algorithms, respectively.

of the number of association chains visited over the graph. We observe that the number of chains visited is lower than the number of total pruned edges in the graph and lies close to the lower bound $N_S N_T$. This shows the effectiveness of geometric features in solving the association problem with significantly lower complexity than the worst-case $N_T^{N_S}$ complexity.

In order to effectively compare the performance, we now consider adverse scenes in which the sensor observations contain detection anomalies. When the miss probability is low, *SAGA* rapidly extracts all chains. As the number of missed detections increases, the robust scheme automatically increases the number of iterations by allowing relaxation of constraints in DFS graph search. In contrast, *SAESL* always requires a large number of iterations.

SAGA provides robustness to missed detections by selectively adding skip edges to the graph. This mechanism reduces the OSPA error in adverse scenarios at the expense of increased computational complexity. The level of robustness can be tuned using a parameter ρ which is set based on the adversity of the scene. Figure 5 also shows the estimation performance for different robustness levels with increasing scene adversity (i.e., increasing miss detections). OSPA error reduces with higher robustness levels. However, low robustness level (e.g., $\rho = 1$) is sufficient to obtain good performance

Fig. 6. Association complexity versus the number of targets averaged over 100 trials using nominal parameters with robustness levels $\rho = 0$ and $\rho = 4$. Total number of FLOPS is denoted by blue line while the runtime is in red.
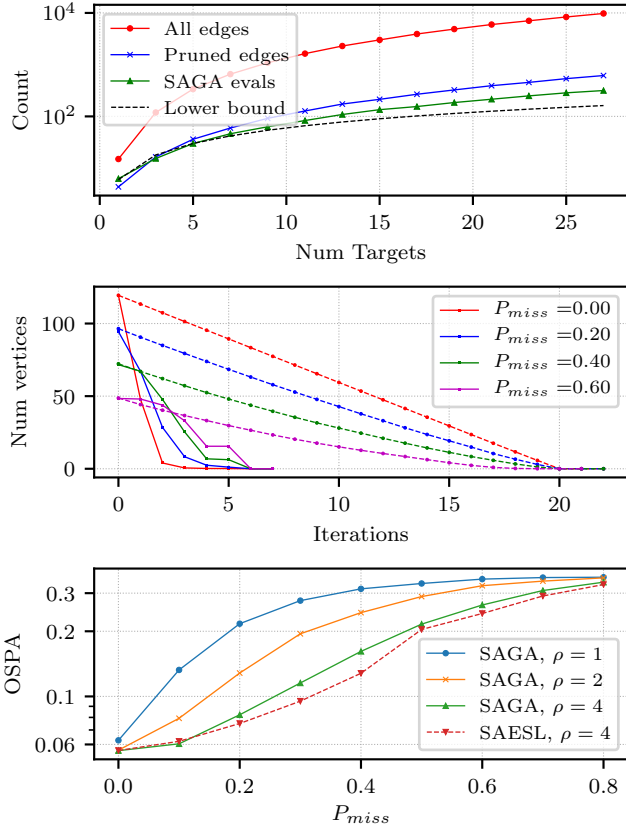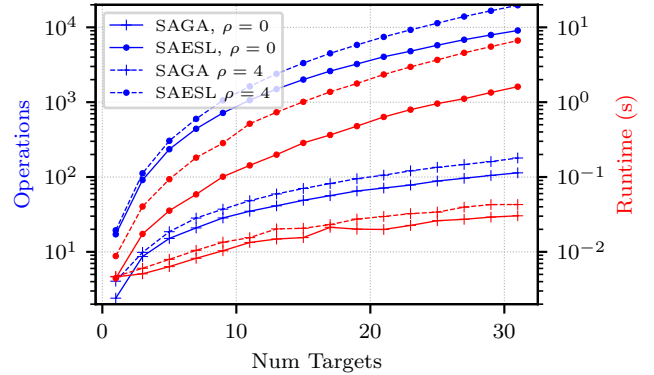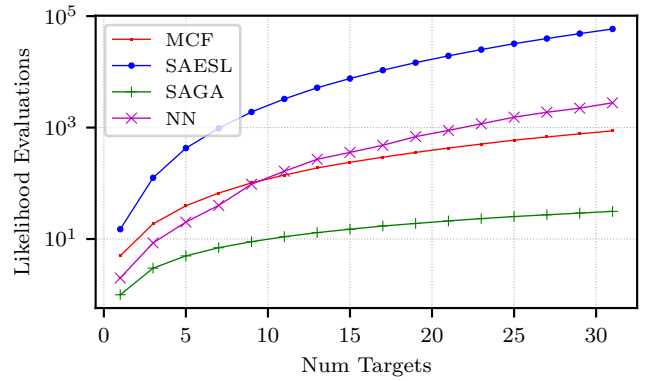

Fig. 5. (Top) Graph edge count and number of association chains visited by SAGA algorithm. (Middle) Graph size at end of each iteration of association algorithm for different P_{miss}. *SAGA* is denoted by solid line while *SAESL* is denoted by dotted line. (Bottom) OSPA versus P_{miss} with different robustness levels ρ

Fig. 7. Number of evaluations of Likelihood $\mathcal{L}(\mathcal{A})$ with increasing number of targets. *SAGA* has the lowest complexity across all scene densities while the complexities *NN* and *MCF* lie between those of *SAGA* and *SAESL*.

at typical miss detection probability $P_{miss} < 0.05$. Similarly, a higher robustness level helps to reduce the cardinality errors when the scene contains higher number of targets. The highest robustness level is $\rho = 4$, which corresponds to the minimum chain length constraint in (16a) with $n(\mathcal{A}) \geq 2$.

Runtime Comparison: We now compare the computational complexity of our approach against *SAESL*. Computing the number of operations that occur during association is difficult, since the number of chains visited depends on a variety of factors such as the fitting error thresholds and minimum chain length. However, given the same sensor estimates for the simulated scenes, we compare the relative complexities of *SAGA* against other methods in Figure 6 in terms of total number of Floating Point operations (FLOPS) and the total runtime. We observe that *SAGA* exhibits an order of magnitude lower complexity compared to the *SAESL* algorithm. Moreover, this improvement increases as the number of targets increases, which highlights the advantage of our approach. In addition, as we increase the robustness level (e.g., from $\rho = 0$ to $\rho = 4$), the relative increase in complexity for the proposed *SAGA* algorithm is far less than for the *SAESL* approach.

We also compare the complexity against traditional methods such as gated Nearest neighbor filter (*NN*) and Minimum cost flow (*MCF*). The NN association scheme [13] builds the association chain by starting with a local kinematic state

estimate from a pair of sensor observations and sequentially adding the nearest measurement from other sensors to update this state. The MCF association scheme [9] identifies the most likely set of chain by solving the minimum cost maximum flows over the graph. The cost of each edge is set based on its relative likelihood similar to our *SAESL* method. We use an optimized implementation [24] of *MCF* for comparison purposes.

In order to compare the complexity of those algorithms, we count the number of times that the primary objective function (i.e., the likelihood cost in (7)) is computed during the graph search procedure. Figure 7 compares the complexity across algorithms as a function of the number of targets. The proposed *SAGA* algorithm requires the lowest number of likelihood evaluations–significantly lower than for the naive *SAESL* iterative search method. This shows our algorithm can effectively predict the correct chain using the geometric fitting criteria. The *MCF* and *NN* algorithms have similar complexities, lying between those of *SAGA* and *SAESL*.

Figure 8 shows the overall runtime of the algorithms as the scene density increases. We observe that *SAGA* is faster than the other methods by an order of magnitude. Since, the FLOPS count is not available for these other methods, we only compare the overall runtime, which follows a similar trend and provides a reasonable estimate of algorithmic complexity.
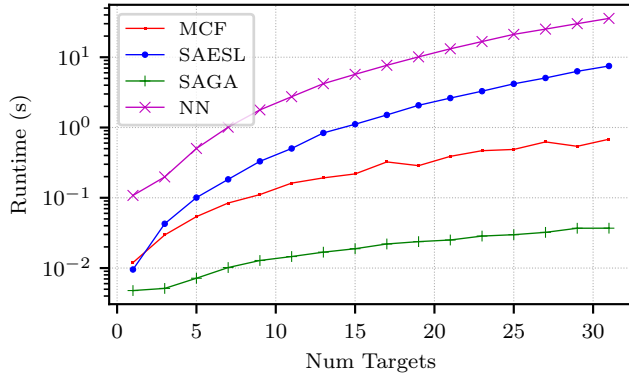
Fig. 8. Runtime comparison with traditional algorithms.

Fig. 9. Estimation accuracy (thin) and cardinality error (thick) versus number of targets at SNR = -15 dB.

C. Benefit of Super-Resolution

Our algorithm extracts the geometric relationships between range-Doppler measurements based on the sensor array geometry and builds association chains by adding likely observations at new sensors to existing chains. In this section, we investigate the role of enhanced accuracy of range and Doppler estimates obtained using the NOMP [3] super-resolution algorithm in spatial association compared to coarse estimates obtained using DFT. Figure 9 compares the localization and cardinality errors. We see that the localization accuracy using NOMP estimates achieves the CRB when the number of targets is moderate, whereas DFT has higher RMSE as expected. However, the RMSE of NOMP deviates away from CRB as the number of targets increases and approaches the accuracy of DFT-based estimates for dense scenes.

It is important to note that our association algorithm works even with the coarse DFT-based estimates. However, NOMP provides an accuracy boost at the *input* of the association algorithm, which enables identification of more targets and results in lower cardinality errors compared to DFT.

The increased accuracy of NOMP estimates also results in smaller association time relative to DFT, due to the reduction in geometric fitting errors. This reduction in association time comes, of course, at the expense of additional computation during range-Doppler estimation. Figure 10 compares the runtime of the estimation and association stages with different number of sensors for $N_T = 20$ targets. We observe that the association time with NOMP estimates is 10 times lower than the one with DFT estimates, while the estimation overhead is about $2-3$ times higher. Figure 10 shows that the computational complexity of association starts dominating that of estimation as the number of targets and sensors increases. Thus, the overall complexity reduction due to NOMP-based estimation, relative to DFT-based estimation, becomes more pronounced with a denser scene and a larger number of sensors.

D. Array Geometry

In this subsection, we analyze the localization performance of linear sensor arrays from the perspective of data association. We consider the adverse scene with $P_{miss} = 0.2$ to emphasize
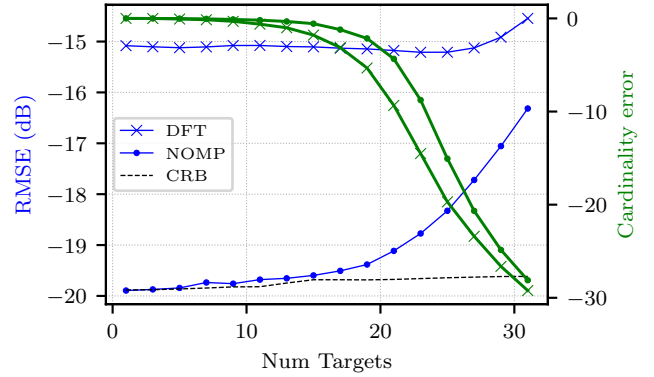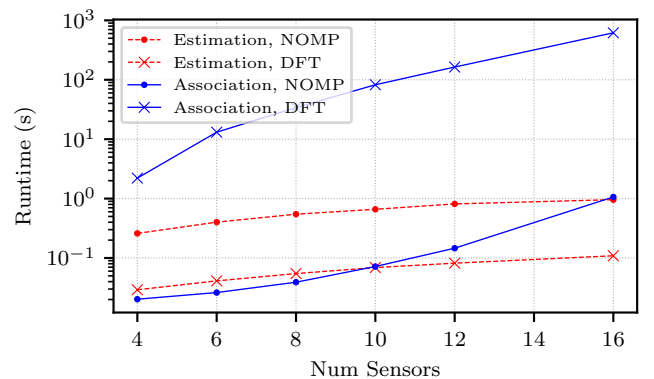

Fig. 10. Runtime comparison of association (solid) & estimation (dotted) stages versus number of sensors.

our findings. The array width and the number of sensors affect both localization accuracy and association complexity.

Increasing the array width generates more spatial diversity in range-Doppler measurements across sensors. This helps to reduce the OSPA error for a given number of sensors. On the other hand, larger distance among the sensors weakens the pruning criteria for the graph edges used in (8), resulting in a denser graph with a higher number of potential associations between sensors. Therefore, the overall localization performance improves with wider arrays at the expense of slightly more association complexity. The available sensor width is an important design constraint in practical applications (e.g., the length of the side profile of a vehicle). We therefore analyze the effect of the number of sensors, keeping the array width fixed to $L_W = 4$ m.

We find that increasing the number of sensors improves association performance as well as association complexity. Figure 11 shows OSPA versus number of sensors for *SAESL* and *SAGA*. While the OSPA for *SAESL* association decreases monotonically with the number of sensors, we observe that the OSPA for *SAGA* with robustness level ρ achieves minimum OSPA with $N_S = \rho + 3$ sensors, and increases for $N_S > \rho + 3$. This is due to missed observations preventing the formation of chains with minimum length constraint $N_S - \rho$. For an array with N_S sensors and a robustness level of ρ, the expected

number of missed targets can be expressed as

$$\mathbb{E}[\text{miss}] = \sum_{k=1}^{\min(N_S-2,\rho+1)} \binom{N_S}{k} P_{miss}^k (1 - P_{miss})^{N_S-k}.$$

Figure 12 shows that the number of missed targets observed using our approach closely matches this expected value for various values of ρ and N_S.

Thus, as we increase the number of sensors, while we improve localization accuracy, we must increase the robustness level used in the *SAGA* algorithm (setting it to $\rho = N_s - 3$) in order to avoid increase in cardinality errors. While this does result in increased computational complexity, it is still significantly lower than that of the *SAESL* algorithm. We leave as an open issue the design of more sophisticated methods for selection of a subset of sensors during the association stage to reduce the complexity further.

V. CONCLUSION

We have shown that simple constraints relating range-Doppler observations to sensor geometry can be exploited to significantly reduce the complexity of spatial association. Our system-level simulations demonstrate that the proposed framework for spatial association based on these geometric constraints is robust to noisy observations and detection anomalies, and that it scales well with the number of sensors and targets. Our approach is compatible with standard FFT-based range-Doppler processing, but enhanced accuracy estimation at each sensor (i.e., super-resolution of range and Doppler) significantly improves both localization accuracy and association complexity. The geometric constraints used to simplify the association problem rely on a linear placement of the sensor array, which is reasonable, for example, when the sensors are placed along the side, front or back of a vehicle. An interesting open question is whether such geometric concepts can be extended to simplify association for more general sensor array configurations. Additional important topics for future investigation include extending these ideas to more complex target models (e.g., for extended targets, and targets causing both specular and diffuse reflection), and combining them with complementary strategies utilizing platform and/or target motion across multiple snapshots.
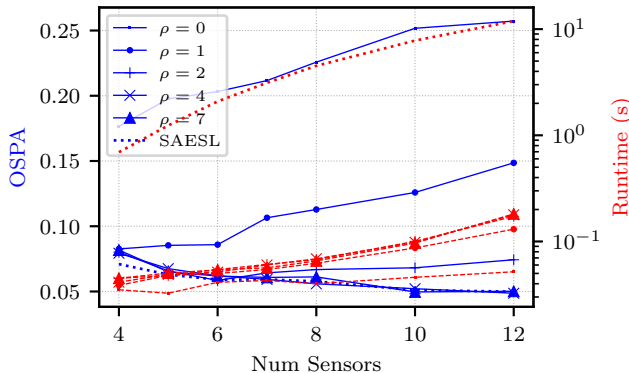

Fig. 11. Association versus number of sensors for SAESL (thick dotted) and SAGA (thin solid) with different robustness levels.
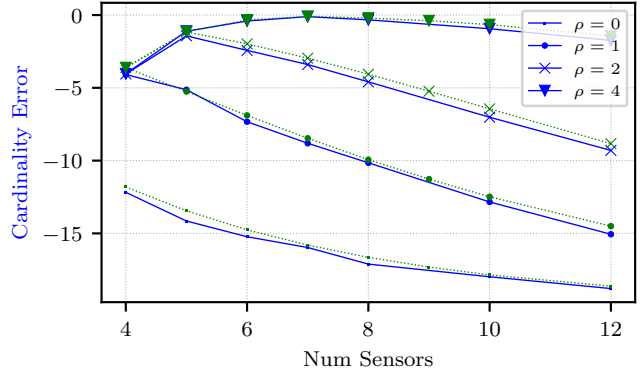

Fig. 12. Comparison of simulated (solid) and theoretical (dotted) cardinality error.

APPENDIX A
CRB FOR RANGE AND DOPPLER

The Cramér Rao Bound provides an estimation theoretic lower bound on the sample covariance of range-Doppler estimates; that is, $Cov(\boldsymbol{\theta}_i) \geq I(\boldsymbol{\theta}_i)^{-1}$ where $I(\boldsymbol{\theta}_i)$ is Fisher Information Matrix (FIM) given by,

$$I(\boldsymbol{\theta}_i) = \mathbb{E}\left[\left(\nabla_{\boldsymbol{\theta}_i}\mathcal{L}(m_i^{obs}|\boldsymbol{\theta}_i)\right)\left(\nabla_{\boldsymbol{\theta}_i}\mathcal{L}(m_i^{obs}|\boldsymbol{\theta}_i)\right)^H\right]$$

where $\mathcal{L}(m_i^{obs}|\boldsymbol{\theta}_i)$ is the log likelihood of the observed signal for a given target range-Doppler $\boldsymbol{\theta}_i$. For an FMCW radar, this expression simplifies to [21],

$$I(\boldsymbol{\theta}_i) = \kappa \left(\frac{A^2}{\sigma^2}\right) \begin{bmatrix} 1/\delta_r^2 & 0 \\ 0 & 1/\delta_d^2 \end{bmatrix} \tag{19}$$

where κ is a constant, and δ_r and δ_d are the Rayleigh range and Doppler resolutions, respectively. $\frac{A^2}{\sigma^2}$ in (19) is the SNR of the received signal, $m_i^{obs}(t)$ with $m_i^{obs}(t) = Ae^{j\phi_{FMCW}(t)} + w(t)$ at sensor i, where $w(t) \sim \mathcal{N}(0, \sigma^2)$. We set the nominal variance of range-Doppler estimates based on the value of CRB at $SNR = -20$ dB; that is, $\sigma_{r_i} = \frac{\delta_r\sigma}{\kappa A}$ and $\sigma_{d_i} = \frac{\delta_d\sigma}{\kappa A}$.

APPENDIX B
CRB FOR POSITION AND VELOCITY

Using the range-Doppler model in Section 3, we evaluate the single target CRB for kinematic parameters $\bar{\boldsymbol{z}}$ using the log likelihood of range-Doppler observations $\mathcal{A} = \{\boldsymbol{\theta}_i\}_{i=1}^{N_S}$ given kinematic state $\bar{\boldsymbol{z}}$, which is

$$\mathcal{L}\left(\{\boldsymbol{\theta}_i|\bar{\boldsymbol{z}}\}_{i=1}^{N_S}\right) = \sum_{i=1}^{N_S} \left(\frac{(\bar{r}_i - r_i)^2}{\sigma_{r_i}^2} + \frac{(\bar{d}_i - d_i)^2}{\sigma_{d_i}^2}\right)$$

where $\boldsymbol{\theta}_i = (r_i, d_i)$ is the observed range-Doppler pair for sensor i, $(\bar{r}_i, \bar{d}_i) = \mathcal{T}_i(\bar{\boldsymbol{z}})$ is true range-Doppler pair for given target state $\bar{\boldsymbol{z}}$ and $\sigma_{r_i}^2$ and $\sigma_{d_i}^2$ are, respectively, the range and Doppler CRBs obtained in (19). The FIM for $\bar{\boldsymbol{z}}$ can be evaluated as

$$I(\bar{\boldsymbol{z}}) = \mathbb{E}\left[\nabla_{\boldsymbol{z}}\mathcal{L}\left(\{\boldsymbol{\theta}_i|\bar{\boldsymbol{z}}\}_{i=1}^{N_S}\right)\right].$$

The CRB obtained from inverse FIM is used to find position and velocity CRB as follows,

$$CRB_p = I(\bar{\boldsymbol{z}})_{(1,1)}^{-1} + I(\bar{\boldsymbol{z}})_{(2,2)}^{-1},$$
$$CRB_v = I(\bar{\boldsymbol{z}})_{(3,3)}^{-1} + I(\bar{\boldsymbol{z}})_{(4,4)}^{-1}.$$

The CRB of velocity is a function of both range and Doppler variances whereas the CRB of position only depends on the variance of range. We use the nominal range and Doppler CRB values to set the minimum separation distance threshold, $\tau_z = 10\sqrt{CRB_p + CRB_v}$ between targets. This threshold is also used to check similarity between chains in the association algorithm.

APPENDIX C
ASSOCIATION CONSTRAINT RELAXATION

The choice of initial stopping thresholds τ_f^n and τ_l^n and scaling factor β for subsequent relaxations in *SAGA* algorithm governs the total complexity of association algorithm. In order to initialize the association algorithm, we set tight thresholds for $\mathcal{L}(\mathcal{A})$ and $\mathcal{F}(\mathcal{A})$. Assuming the range-Doppler observations have small error (i.e., $w_i^R \ll r_i$, $w_i^D \ll d_i$ in (3)), the expected negative log likelihood in (5) can be approximated as

$$\mathcal{L}(\mathcal{A}) \approx \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} \left(\frac{(w_i^R)^2}{\sigma_r^2} + \frac{(w_i^D)^2}{\sigma_d^2} \right).$$

Since $w_i^R \sim \mathcal{N}(0, \sigma_r^2)$ and $w_i^D \sim \mathcal{N}(0, \sigma_d^2)$ are standard Normal distributed random variables, $\mathcal{L}(\mathcal{A}^k)$ has chi-squared distribution, $\chi_{2n(\mathcal{A})}^2$ with $2n(\mathcal{A})$ degrees of freedom. Then, the expected fitting error in (14) can be approximated as

$$F(\mathcal{A}) = \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} \frac{\left((\hat{r}_i \hat{d}_i) - (r_i d_i) \right)^2}{\eta_1} + \frac{\left((\hat{r}_i)^2 - (r_i)^2 \right)^2}{\eta_2}$$

$$\approx \sum_{\boldsymbol{\theta}_i \in \mathcal{A}} \frac{\left(r_i w_i^D + d_i w_i^R \right)^2}{\eta_1} + \frac{\left(2 r_i w_i^R \right)^2}{\eta_2} \quad (20)$$

where ($\hat{r}_i$ and $\hat{d}_i$) denote the perceived range-Doppler pair at sensor i for predicted state $\boldsymbol{z}$ and $\boldsymbol{\theta}_i = (r_i, d_i)$ denotes the observed range-Doppler pair at sensor i. The normalization factors η_1, η_2 are set to the variance of numerator terms which is,

$$(\eta_1)_i = \text{Var}[r_i w_i^D + d_i w_i^R] \approx \sigma_{r_i}^2 d_i^2 + r_i^2 \sigma_{d_i}^2 + \sigma_{r_i}^2 \sigma_{d_i}^2$$

$$(\eta_2)_i = \text{Var}[2 r_i w_i^R] \approx 4 r_i^2 \sigma_{r_i}^2.$$

Using those values to normalize (20) results in $\mathcal{F}(\mathcal{A}^k) \sim \chi_{2n(\mathcal{A})}^2$ being chi-squared distributed with $2n(\mathcal{A})$ degrees of freedom. Hence, the thresholds for the association algorithm are determined as follows,

$$\tau_f^{n(\mathcal{A})} : \Pr(F(\mathcal{A}) > \tau_f^{n(\mathcal{A})}) = P_{\text{FA}}$$

$$\tau_l^{n(\mathcal{A})} : \Pr(\mathcal{L}(\mathcal{A}) > \tau_l^{n(\mathcal{A})}) = P_{\text{FA}}$$

where P_{FA} is the nominal false alarm rate set to $P_{\text{FA}} = 0.01$.

Note that while the normalization factors η_1, η_2 depend on r_i, d_i, we set this based on the maximum range, Doppler values to get a conservative initial value. This does not cause a problem since the sucessive relaxation procedure loosens that threshold so that chains with high fitting error can be extracted.

The relaxation factor, β should be set appropriately. Choosing a high value causes faster convergence but might lead to false chains being identified. On the other hand, a low value delays the extraction of *loose* chains. In the simulations, we find that $\beta = 2$ performs well.

APPENDIX D
DEPTH FIRST SEARCH

A depth first search algorithm is outlined in Algorithm 3. At each node, the DFS procedure traverses through all branches which have geometric fitting error below the maximum error threshold $\tau_f^{N_S}$. On reaching the end of the graph, we select the chain if it satisfies the likelihood, fitting error, and minimum chain length constraints. In addition, we check for possible chain termination at each node after going through all its branches. This step implicitly accounts for the NULL state at the end of a chain.

Algorithm 3 Geometry Assisted Depth First Search

1: **procedure** GA-DFS($v, \mathcal{A}, \gamma, \boldsymbol{\tau}$)
2:      Get list of children of v that *geometrically fit*,
$$B(v) = \left\{ v_j : \mathcal{F}([\mathcal{A}, v_j]) < \tau_f^{N_S} \right\}$$
3:      **if** $B(v) \neq \emptyset$ **then**
4:          Sort $B(v)$ using geometric fitting error, $F([\mathcal{A}, v_j])$
5:          **for** $v_j \in B(v)$ **do**
6:              BRANCH out a new chain $\mathcal{A}^j : \mathcal{A} \leftarrow v_j$
7:              $\mathcal{A}^o \leftarrow$ GA-DFS($v_j, \mathcal{A}^j, \gamma, \boldsymbol{\tau}$)
8:              Exit loop if valid chain $\mathcal{A}^o$ is found.
9:          **end for**
10:      **end if**
11:      CHECK IF CHAIN CAN BE TERMINATED AT v
12:      **if** $n(\mathcal{A}) \geq \gamma, \mathcal{L}(\mathcal{A}) < \tau_l^{n(\mathcal{A})}, F(\mathcal{A}) < \tau_f^{n(\mathcal{A})}$ **then**
13:          SELECT $\mathcal{A}^o \leftarrow \mathcal{A}$,
14:      **end if**
15:      **Output:** $\mathcal{A}^o$
16: **end procedure**

APPENDIX E
MINIMUM AMBIGUITY ASSOCIATION

Lemma 1. *In the ideal detection scenario (i.e., no miss or false alarms), the number of candidate locations generated between a pair of sensors is minimum for consecutive sensors.*

Proof. Recall that candidate locations are generated when range perceived at a pair of sensors satisfy conditions in (8). For a candidate, $\boldsymbol{z}_{ij}^{pq}$ generated by incorrectly associated observations, $\boldsymbol{\theta}_i^p, \boldsymbol{\theta}_j^q$, across consecutive sensors i, j, the following relations hold,

$$r_i^p - r_j^q < l_{ij}, \quad r_i^p + r_j^q > l_{ij}. \quad (21)$$

Now consider $\bar{q}^{th}$ observation at sensor k adjacent to sensor j which corresponds to same target as $\boldsymbol{\theta}_j^q$, the following hold,

$$r_j^q - r_k^{\bar{q}} < l_{jk} \quad \text{(using (8))}$$

$$r_j^q + l_{jk} > r_k^{\bar{q}} \quad (l_{jk} \geq 0, \text{ Triangle inequality})$$

Using these along with (21) we obtain,

$$r_i^p - r_k^{\bar{q}} < l_{ik}, \quad r_i^p + r_k^{\bar{q}} > l_{ik}$$

Hence any candidate produced between consecutive sensors i, j also generates a candidate between sensors i, k by skipping over intermediate sensor j. Hence,

$$\sum_{p=1}^{n(\Theta_i)} \sum_{q=1}^{n(\Theta_{i+1})} n(\boldsymbol{z}_{i,i+1}^{pq}) \leq \sum_{p=1}^{n(\Theta_i)} \sum_{q=1}^{n(\Theta_k)} n(\boldsymbol{z}_{i,k}^{pq})$$

Therefore, the number of candidates generated between a pair of sensors is minimum for consecutive sensors. □

Association complexity is due to the presence of unwanted candidate targets which need to be discarded based on their likelihood. When a target is observed at all sensors, it is sufficient to associate observations along consecutive sensors. Lemma 1 states that the association of observations along consecutive sensors generates the lowest number of phantoms during graph search. Hence, the number of potential ambiguities is minimized when the graph search procedure is conducted across consecutive sensors first.

REFERENCES

[1] D. Barrett, D. Wang, A. Ahmad, and V. Mahimkar, "Using mmWave sensors to enhance drone safety and productivity," Texas Instruments, Tech. Rep., 2017.
[2] D. Barrett and A. Alvarez, "mmWave radar sensors in robotics applications," Texas Instruments, Tech. Rep., 2017.
[3] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum." *IEEE Trans. Signal Processing*, vol. 64, no. 19, pp. 5066–5081, 2016.
[4] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Transactions on intelligent vehicles*, vol. 3, no. 1, pp. 30–45, 2017.
[5] Q. Wang, Z. Chen, Z. Wang, Z. Zhao, and Z. Wu, "Clustering data association using data relevance in spatial domain for doppler-only rsn localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 6, pp. 3018–3031, 2018.
[6] J. L. Williams and R. A. Lau, "Multiple scan data association by convex variational inference," *IEEE Transactions on Signal Processing*, vol. 66, no. 8, pp. 2112–2127, 2018.
[7] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.
[8] B. Wu and R. Nevatia, "Tracking of multiple, partially occluded humans based on static body part detection," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 1, 2006, pp. 951–958.
[9] L. Zhang, Y. Li, and R. Nevatia, "Global data association for multi-object tracking using network flows," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.
[10] H. Jiang, S. Fels, and J. J. Little, "A linear programming approach for multiple object tracking," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, June 2007, pp. 1–8.
[11] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes, "Globally-optimal greedy algorithms for tracking a variable number of objects," in *CVPR 2011*, June 2011, pp. 1201–1208.
[12] J. Berclaz, F. Fleuret, E. Turetken, and P. Fua, "Multiple object tracking using k-shortest paths optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1806–1819, Sep. 2011.
[13] Y. Bar-Shalom and X.-R. Li, *Multitarget-multisensor tracking: principles and techniques*. YBS, Storrs, CT, 1995, vol. 19.
[14] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.
[15] K. G. Murty, "Letter to the editor—An algorithm for ranking all the assignments in order of increasing cost," *Operations Research*, vol. 16, no. 3, pp. 682–687, 1968.
[16] F. Folster and H. Rohling, "Data association and tracking for automotive radar networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 4, pp. 370–377, 2005.
[17] S. Venkateswaran and U. Madhow, "Localizing multiple events using times of arrival: A parallelized, hierarchical approach to the association problem," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5464–5477, 2012.
[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[19] W. E. L. Grimson and T. Lozano-Perez, "Localizing overlapping parts by searching the interpretation tree," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 4, pp. 469–482, 1987.
[20] J. Neira and J. D. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 890–897, 2001.
[21] A. Gupta, U. Madhow, and A. Arbabian, "Super-resolution in position and velocity estimation for short-range MM-Wave radar." in *50th Asilomar Conference on Signals, Systems and Computers*, Nov. 2016, pp. 1144–1148.
[22] T. Wagner, R. Feger, and A. Stelzer, "Wide-band range-doppler processing for FMCW systems," in *2013 European Radar Conference*, Oct. 2013, pp. 160–163.
[23] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo, "A metric for performance evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3452–3457, 2011.
[24] L. Perron and V. Furnon, "Or-tools," Google. [Online]. Available: https://developers.google.com/optimization/

Anant Gupta received the B.Tech. degree in electronics and electrical communication engineering and the M.Tech. degree in telecommunication systems engineering from IIT Kharagpur in 2013. He received the M.S. and Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara (UCSB) in 2016 and 2020 respectively. He is currently working as Senior Engineer in the wireless R&D team at Qualcomm, San Diego. His research interests include wireless sensing, signal processing and machine learning.
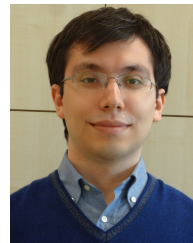
Ahmet Dundar Sezer received the B.S., M.S., and Ph.D. degrees in Electrical and Electronics Engineering from Bilkent University, Ankara, Turkey, in 2011, 2013, and 2018, respectively. He is currently a Post-Doctoral Researcher at the University of California, Santa Barbara, CA, USA. His current research interests include signal processing, wireless communications, and optimization.

Upamanyu Madhow is Professor of Electrical and Computer Engineering at the University of California, Santa Barbara. His current research interests focus on next generation communication, sensing and inference infrastructures centered around millimeter wave systems, and on robust machine learning. He received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1985, and his Ph. D. degree in electrical engineering from the University of Illinois, Urbana-Champaign in 1990. He has worked as a research scientist at Bell Communications Research, Morristown, NJ, and as a faculty at the University of Illinois, Urbana-Champaign. Dr. Madhow is a recipient of the 1996 NSF CAREER award, and co-recipient of the 2012 IEEE Marconi prize paper award in wireless communications. He has served as Associate Editor for the IEEE Transactions on Communications, the IEEE Transactions on Information Theory, and the IEEE Transactions on Information Forensics and Security. He is the author of two textbooks published by Cambridge University Press, *Fundamentals of Digital Communication* (2008) and *Introduction to Communication Systems* (2014).