# mmSnap: Bayesian One-Shot Fusion in a Self-Calibrated mmWave Radar Network

Anirban Banik[†], Lalitha Giridhar[†], Aaditya Prakash Kattekola[†], Anurag Pallaprolu[†],
Yasamin Mostofi[†], Ashutosh Sabharwal[§] and Upamanyu Madhow[†]
[†]University of California, Santa Barbara, CA, U.S.A.  [§]Rice University, TX, U.S.A
{abanik, lalitha, aadityaprakash, apallaprolu, ymostofi, madhow}@ucsb.edu  ashu@rice.edu

*Abstract*—We present *mmSnap*, a collaborative RF sensing framework using multiple radar nodes, and demonstrate its feasibility and efficacy using commercially available mmWave MIMO radars. Collaborative fusion requires network calibration, or estimates of the relative poses (positions and orientations) of the sensors. We experimentally validate a self-calibration algorithm developed in our prior work, which estimates relative poses in closed form by least squares matching of target tracks within the common field of view (FoV). We then develop and demonstrate a Bayesian framework for one-shot fusion of measurements from multiple calibrated nodes, which yields instantaneous estimates of position and velocity vectors that match smoothed estimates from multi-frame tracking. Our experiments, conducted outdoors with two radar nodes tracking a moving human target, validate the core assumptions required to develop a broader set of capabilities for networked sensing with opportunistically deployed nodes.

*Index Terms*—mmWave radar network, self-calibration, radar fusion

## I. INTRODUCTION

The integration of low-cost millimeter-wave (mmWave) MIMO radar in the 60 and 77 GHz bands with a communication infrastructure (e.g., cellular or WiFi) enables the creation of collaborative RF sensing networks, as illustrated in Figure 1. This approach can overcome the coverage limitations of individual radar nodes, which face line of sight (LoS) obstructions and limited field of view (FoV). When radar nodes with overlapping FoVs collaborate, they can significantly enhance the quality of inference by fusing their measurements. For instance, a single mmWave radar node can provide a precise estimate of Doppler (radial velocity), which is the projection of a moving target's velocity along the LoS. However, it cannot instantaneously estimate the target's full vector velocity. By combining Doppler measurements from two radar nodes, the vector velocity can be estimated instantaneously, except in degenerate configurations. Similarly, individual nodes can make accurate position estimates by combining high-resolution range data (due to the large bandwidth available) with angle measurements, though the latter is limited by the aperture size of the radar. Collaborative sensing allows multiple nodes to combine their range estimates, enhancing the resolution of position estimates. Such fine-grained fusion relies
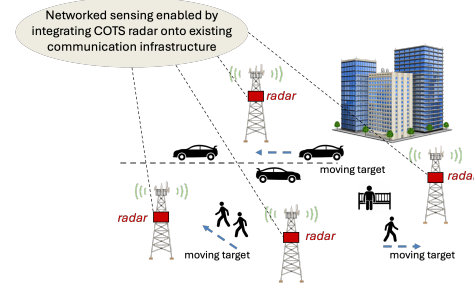
Fig. 1: A collaborative networked RF sensing infrastructure.

on precise network calibration, requiring accurate estimates of the nodes' relative positions and orientations (poses). In opportunistically deployed networks, where such information is typically unavailable, calibration becomes a major challenge but remains essential for effective collaborative sensing.

**Contributions:** We introduce *mmSnap*, a framework for self-calibration and Bayesian one-shot fusion in a radar network, designed to deliver instantaneous ("snap") position and velocity through fusion in a millimeter-wave (mmWave) radar network. Self-calibration is based on an optimal algorithm for joint tracking and pose estimation introduced in our prior work [1], in which closed form estimates for the relative poses for any two nodes tracking a target in their shared FoV are obtained by least squares alignment of the estimated tracks. Then, assuming a calibrated network, we develop a Bayesian framework for estimating instantaneous target position and velocity based on measurements from collaborating nodes. We present preliminary experimental results using COTS 77 GHz mmWave radar nodes in an outdoor setting, validating that the self-calibration algorithm from [1] yields accurate pose estimates from human target tracks, and that post-calibration, our Bayesian one-shot fusion method produces instantaneous estimates close to smoothed multi-frame tracking outputs. While these initial experiments are limited to two radar nodes and a single moving target, they establish the feasibility of core building blocks for collaborative RF sensing at scale.

**Related work:** The use of distributed radar for estimating vector velocity has been explored in several studies [2]–[5]: [2] assesses single-shot ego-velocity for vehicles, while [3] uses a linear least-squares method for velocity estimation amidst Doppler ambiguities. In [4], Doppler and angular data from two radars are fused, and [5] proposes a maximum

Fig. 2: The end-to-end mmSnap pipeline.

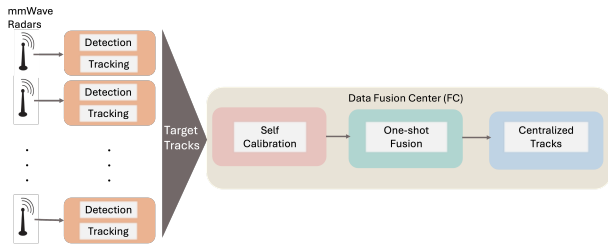| Parameter | Value |
|---|---|
| Carrier Frequency ($f_c$) | 76 GHz |
| ADC Sampling Rate | 10 MHz |
| Number of Transmitters | 3 |
| Number of Receivers | 4 |
| Bandwidth | 4.24 GHz |
| Chirp Slope | 83 MHz/$\mu$s |
| Frame Duration | 150 ms |
| Number of Chirps per Frame | 60 |
| Chirp Ramp Time | 60 $\mu$s |
| Range Resolution | 3.54 cm |
| Doppler Resolution | 0.183 m/s |
| Max Unambiguous Range | 18.07 m |

TABLE I: Radar Parameters for mmSnap Experiments

likelihood method, for velocity estimation which is prone to struggle in noisy conditions. Unlike our framework, which includes self-calibration and Bayesian fusion, these studies assume known radar poses relative to a global frame, which may be unrealistic. Further, the maximum likelihood techniques that they employ struggle in the presence of geometric degeneracy. A more general problem of scene estimation and point cloud based environmental mapping with multiple radars is considered in [6], with Doppler information utilized for registration across radars. Our approach is computationally much simpler because it focuses on self-calibration and one-shot fusion for point target models, demonstrating the efficacy of approximating extended human targets as detection-cluster centroids for both tasks. Finally, recent efforts in self-calibration using a target in the common FoV include brute force optimization for 3D pose estimation using a corner reflector target in [7], and singular value decomposition (SVD) for 2D pose estimation using human targets in [8]. Our self-calibration approach is similar to the latter, but we utilize the complex number framework in our prior work [1], which yields relative poses in closed form.

To the best of our knowledge, this is the first work to experimentally demonstrate a framework for self-calibration and Bayesian one-shot fusion in radar networks geared toward urban outdoor settings. While our approach is applicable to general moving objects, we validate the framework through controlled experiments with human motion. Results across varied radar configurations show that our approach is well-suited to collaborative sensing with opportunistic node deployment, where significant variation in the relative poses of nodes with overlapping FoVs is expected.

The remainder of this paper is structured as follows: Section II presents the full mmSnap system model and self-calibration algorithm; Section III details the Bayesian framework for one-shot fusion. Experimental results for self-calibration and fusion in a two-radar network are provided in each of these sections. Section IV offers concluding remarks and outlines future directions.

## II. MMSNAP SYSTEM MODEL

A possible realization of the end-to-end mmSnap pipeline is illustrated in Figure 2. Each radar node carries out conventional point cloud detection individually. Depending on the available computation, it may also perform target tracking on its own, as shown in Figure 2, or this functionality could be migrated to the fusion center. Each node shares its detection

and tracking (if performed at the node) results with the fusion center. Track matching for common targets is used for self-calibration. One-shot fusion of detected points is used to provide low-latency estimates of target position/velocity, while also serving as inputs to a centralized tracker. The nodes are coarsely time-synchronized at the level of a radar frame.

Our experiments employ 77GHz AWR2243 MIMO radar boards from Texas Instruments, with each board outputting raw ADC data in radar cube format. The chirp parameters used in the experiments, and the corresponding range and radial velocity resolutions, are detailed in Table I. The radar boards are equipped with three transmit and four receive antennas, resulting in 12 virtual antennas with an inter-antenna spacing of $d = \lambda/2$. However, the virtual array size for azimuth angle estimation (which is what we focus on) is 8, corresponding to a spatial frequency resolution of $2\pi/8 = \pi/4$. While the mmSnap architecture in Fig. 2 assumes real-time communication with a fusion center, our experiments use offline processing of data collected from multiple sensors.

### A. Signal Processing Pipeline

As illustrated in Fig. 3, the local radar signal processing at each node is standard. The raw ADC data is transformed from its original radar cube format into Range-Doppler-Angle point clouds for each frame by applying Fast Fourier Transforms (FFTs) along the corresponding dimensions. To focus on moving targets, static clutter removal is performed, followed by a two-dimensional Ordered-Statistics Constant False Alarm Rate (2D OS-CFAR) detection applied across frames for further refinement. In the tracking stage, DBSCAN clustering extracts point cloud centroids, which are then processed by an Extended Kalman Filter (EKF) for continuous tracking. For human targets within 10 m, the point cloud shows significant Doppler variation due to limb movements. We design a DBSCAN variant to extract a cluster center representing torso position and motion, allowing the self-calibration and one-shot fusion algorithms, based on point target models, to remain simple and effective.
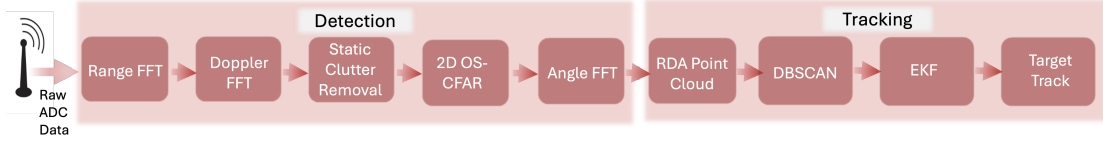
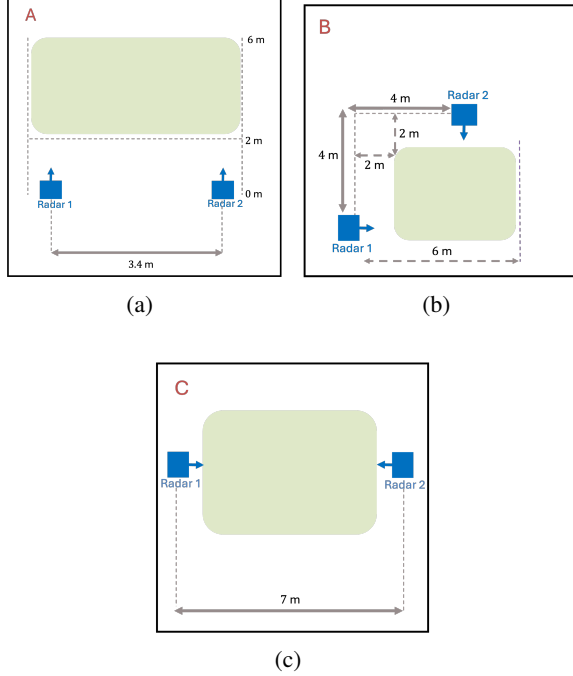Fig. 3: mmSnap Signal Processing Chain from Raw ADC data to Target Tracks.



Fig. 4: Illustration of the experimental setups used to evaluate the mmSnap pipeline, labeled A through C. The green regions demarcate the area within which the target moves, following two different trajectories: either a straight line or random motion.
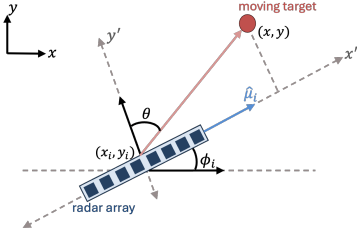


Fig. 5: Measurement model for a single moving target seen by a radar with respect to a global frame of reference

We consider a two-radar outdoor network tracking a single moving target, following either a straight line or a random trajectory. Three two-node configurations (A-C in Fig.4), differing in the relative orientation of the arrays in the nodes, are investigated. The array broadside for each node is indicated by an arrow, and the shaded region represents the area within which the target moves.

### B. Measurement Model

We begin by introducing the basic measurement model for a radar node $S_i$, where $i \in 1, \ldots, N$ in an $N$-node network,

tracking a single moving target, as illustrated in Fig. 5. As shown in Fig. 5, the positions of radar node $S_i$ and the target with respect to a global reference frame are denoted by 2D vectors $\mathbf{p_i} = (x_i, y_i)^T$ and $\mathbf{p} = (x, y)^T$, respectively, and the target velocity is denoted by $\dot{\mathbf{p}} = (v_x, v_y)^T$. The target range is given by

$$r_i = \|\mathbf{p} - \mathbf{p_i}\| = \sqrt{(x - x_i)^2 + (y - y_i)^2} \qquad (1)$$

The radial velocity measured via Doppler is obtained by projecting the target velocity vector onto the unit vector pointing from the radar node to the target:

$$v_i = \left\langle \dot{\mathbf{p}}, \frac{\mathbf{p} - \mathbf{p_i}}{\|\mathbf{p} - \mathbf{p_i}\|} \right\rangle = \frac{v_x (x - x_i) + v_y (y - y_i)}{r_i} \qquad (2)$$

where $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b}$ denotes the inner product of vectors $\mathbf{a}$ and $\mathbf{b}$. For a $\lambda/2$-spaced virtual array, the spatial frequency of the target is given by

$$\omega_i = \pi \sin \theta = \frac{\pi \langle \mathbf{p} - \mathbf{p_i}, \hat{\mu}_\mathbf{i} \rangle}{r_i} = \frac{\pi((x - x_i) \cos \phi_i + (y - y_i) \sin \phi_i)}{r_i} \tag{3}$$

where $\hat{\mu}_\mathbf{i} = (\cos \phi_i, \sin \phi_i)^T$ denote the unit vector along the radar array. In order to apply the preceding measurement model across radar nodes, we need a global frame of reference, which is established via self-calibration, as described next.

### C. Self Calibration

Self-calibration is accomplished by matching the "local" tracks seen by different nodes for a target in their common FoV. We evaluate pairwise self-calibration in our 2-node experiments, but as noted in [1], these can be stitched together for network-wide calibration. Expressing 2D coordinates as complex numbers, denote by $z_i[k] = u_i[k] + jv_i[k]$, $i = 1, 2$, the target's position in the $k$th radar frame as estimated in node $S_i$'s coordinate system. Letting $S_1$ serve as the global reference without loss of generality, the optimal estimate for the relative (complex-valued) position $p_{21}$ and relative orientation $\phi_{21} \in [0, 2\pi)$ of $S_2$ with respect to $S_1$ is obtained by least squares minimization of the track matching error:

$$J(p, \phi) = \sum_{k=1}^{K} |z_1[k] - p - e^{j\phi} z_2[k]|^2 \qquad (4)$$

The solution is obtained in closed form as [1]:

$$\hat{\phi}_{21} = -\angle \widehat{\mathbf{z}}_1^H \widehat{\mathbf{z}}_2$$

$$\widehat{p}_{21} = \overline{z}_1 - e^{j\hat{\phi}_{21}} \overline{z}_2 \tag{5}$$

where $\overline{z}_i$ is the track centroid, $\widehat{\mathbf{z}}_i$ is the centered track from $S_i$'s perspective, defined as $\widehat{\mathbf{z}}_i = (z_i[1] - \overline{z}_i, \ldots, z_i[K] - \overline{z}_i)^T$,

| | A | B | C |
|---|---|---|---|
| **RMSE (m)** | | | |
| Straight Trajectory | 0.5664 | 0.5657 | 0.1783 |
| Random Trajectory | 0.7076 | 0.6331 | 0.4426 |

TABLE II: Trajectory RMSE values for different self-calibration scenarios
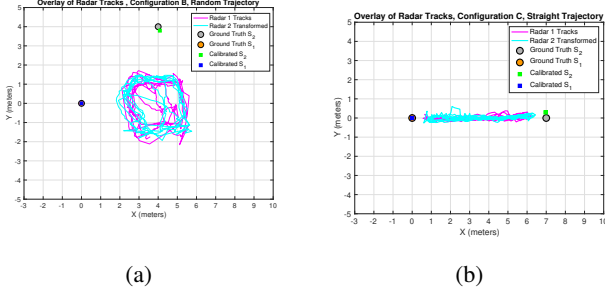


(a)　　　　　　　　(b)

Fig. 6: Overlaid tracks after self-calibration for radar node configurations B and C, with random and straight target trajectories, respectively.

and $\mathbf{x}^H = (\mathbf{x}^*)^T$ denotes the conjugate transpose of a vector $\mathbf{x}$. The minimum cost achieved by this solution is given by

$$J_{min}(2,1) = ||\widehat{\mathbf{z}}_1||^2 + ||\widehat{\mathbf{z}}_2||^2 - 2|\widehat{\mathbf{z}}_1^H \widehat{\mathbf{z}}_2| \qquad (6)$$

The trajectory root-mean-square error (RMSE) after self-calibration is given by

$$RMSE = \sqrt{J_{min}/K} \qquad (7)$$

The rationale for self-calibration is the difficulty of estimating the ground truth for relative poses in large-scale opportunistic deployments. Consequently, the track fitting RMSE serves as a useful, albeit indirect, quantitative indicator of calibration accuracy.

For experimental validation of self-calibration, we collect 90s sequences of a single moving human target within the overlapping FoV of the two radars. For radar frame duration of 150 ms, each sequence yields a trajectory of length $K = 600$ frames for self-calibration. We show in Table II the RMSE as computed using (7) for different node configurations (Fig. 4) and trajectories (straight line and random). While low RMSE indicates good track alignment and hence acceptable accuracy in relative poses, it is not a direct measure of pose accuracy, since it also depends on node configuration and the nature of the trajectory. For instance, straight-line trajectories yield lower RMSE across configurations, yet pose estimates remain comparable to those from random trajectories. In setup B, the estimates are $(3.93\,\text{m}, 3.90\,\text{m}, 274.9°)$ for the straight and $(4.08\,\text{m}, 3.80\,\text{m}, 271.5°)$ for the random trajectory, which are comparable to the ground truth of $(4\,\text{m}, 4\,\text{m}, 270°)$. In Fig. 6, we visually demonstrate the self-calibration algorithm by overlaying the track seen by $S_1$ and that seen by $S_2$, after transforming the latter to $S_1$'s frame based on its estimated relative pose. Since the tracking by each radar is imperfect, the overlay is also noisy. However, averaging across multiple frames yields accurate estimates of the relative pose.

## III. One-shot Fusion

We now extend the measurement model from Section II-B to a multi-radar setting, derive the Bayesian one-shot fusion framework, and present experimental comparisons with smoothed multi-frame estimates.

### A. Bayesian Framework

Consider $N$ radars $S_i, i = 1, ..., N$ tracking a common target moving in their overlapping FoV. We assume that the network is calibrated (i.e., relative poses are known), and without loss of generality, set node $S_1$ as our reference, so that $\mathbf{p_1} = (0,0)^T, \phi_1 = 0°$. The range ($r_i$), radial velocity ($v_i$) and spatial frequency ($\omega_i$) estimates for $S_i, i = 1, ..., N$, are then given by (1)-(3). The corresponding measurements are modeled as conditional Gaussian as follows

$$\begin{aligned} R_i &\sim \mathcal{N}(r_i, \sigma_R^2) \\ \Omega_i &\sim \mathcal{N}(\omega_i, \sigma_\Omega^2) \quad, \forall i \in \{1, 2, ..., N\} \\ V_i &\sim \mathcal{N}(v_i, \sigma_V^2) \end{aligned} \qquad (8)$$

where $\sigma_R^2$, $\sigma_\Omega^2$, and $\sigma_V^2$ are the measurement noise variances in range, spatial frequency, and radial velocity, respectively. The target state $\theta$ is taken to be the instantaneous position and velocity with respect to the global frame, $\theta = (x, y, v_x, v_y)^T$. For one-shot fusion, we estimate the state based on the observation vector $\mathcal{Y}$ corresponding to the actual range, radial Doppler and spatial frequency measurements of the target by the $N$ radars

$$\mathcal{Y} = (R_1, R_2, ..., R_N, \Omega_1, \Omega_2, ..., \Omega_N, V_1, V_2, ..., V_N)^T$$

Assuming that the measurement noises in (8) are independent across sensors and across modalities, we model $\mathcal{Y}$ is conditionally Gaussian given the state $\theta$.

The maximum likelihood (ML) estimate of $\theta$,

$$\widehat{\theta}_{ML} = \arg\max_\theta p(\mathcal{Y} \mid \theta) \qquad (9)$$

is given by minimizing the negative log likelihood $L_{ML}(\theta) = -\log p(\mathcal{Y} \mid \theta)$, and is the solution to the following nonlinear least squares problem

$$\widehat{\theta}_{ML} = \arg\min_\theta L_{ML}(\theta)$$

$$= \arg\min_\theta \left[ \frac{1}{\sigma_R^2} \sum_{i=1}^N (R_i - r_i)^2 + \frac{1}{\sigma_\Omega^2} \sum_{i=1}^N (\Omega_i - \omega_i)^2 \right. $$
$$\left. + \frac{1}{\sigma_V^2} \sum_{i=1}^N (V_i - v_i)^2 \right] \qquad (10)$$

*Limitation:* The ML one-shot estimate is usually effective, but can yield unstable estimates in degenerate settings, as discussed later. We therefore regularize the nonlinear least squares problem in (10) using a Bayesian framework using a Gaussian prior for the state:

$$\begin{aligned} x &\sim \mathcal{N}(0, \sigma_x^2), y \sim \mathcal{N}(0, \sigma_y^2) \\ v_x &\sim \mathcal{N}(0, \sigma_{v_x}^2), v_y \sim \mathcal{N}(0, \sigma_{v_y}^2) \end{aligned} \qquad (11)$$
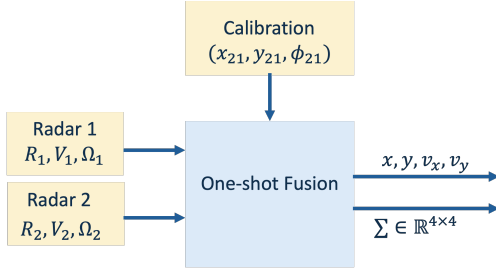
Fig. 7: The one-shot fusion pipeline for a two-radar network.

The Bayesian estimate of $\theta$ corresponds to maximizing the posterior density

$$\widehat{\theta}_{Bayes} = \arg \max_\theta \ p(\theta \mid \mathcal{Y}) = \arg \max_\theta \ p(\mathcal{Y} \mid \theta)p(\theta) \tag{12}$$

or minimizing the corresponding negative log likelihood ratio $L_{Bayes}(\theta) = -log(P(\mathcal{Y} \mid \theta)P(\theta))$. In our Gaussian setting, this yields the following *regularized nonlinear least squares* problem:

$$\widehat{\theta}_{\text{Bayes}} = \arg \min_\theta \left[ \frac{1}{\sigma_R^2} \sum_{i=1}^N (R_i - r_i)^2 + \frac{1}{\sigma_\Omega^2} \sum_{i=1}^N (\Omega_i - \omega_i)^2 + \right.$$
$$\left. \frac{1}{\sigma_V^2} \sum_{i=1}^N (V_i - v_i)^2 + \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{v_x^2}{\sigma_{v_x}^2} + \frac{v_y^2}{\sigma_{v_y}^2} \right] \tag{13}$$

This can be solved iteratively using the Levenberg-Marquardt method, but the solution is sensitive to initial estimates.

**Initial position estimate:** Let $(x_{t_i}, y_{t_i})$ be the target's position estimates w.r.t. $S_i$, which is located at $(x_i, y_i)$ w.r.t. $S_1$, then the initial estimate for the target position is

$$x^{initial} = \frac{1}{N} \sum_{i=1}^N (x_{t_i} \cos \phi_i - y_{t_i} \sin \phi_i + x_i)$$
$$y^{initial} = \frac{1}{N} \sum_{i=1}^N (y_{t_i} \cos \phi_i + y_{t_i} \sin \phi_i + y_i) \tag{14}$$

**Initial velocity estimate:** The initial estimate for the target velocity is taken as

$$v_x^{initial} = \frac{1}{N} \sum_{i=1}^N (V_i \cos (\phi_i + \theta_i))$$
$$v_y^{initial} = \frac{1}{N} \sum_{i=1}^N (V_i \sin (\phi_i + \theta_i)) \tag{15}$$

**Covariance Estimation:** Downstream tasks such as centralized tracking require estimates of posterior covariances as well as posterior means. Although the $4 \times 4$ posterior covariance matrix can be estimated analytically by Laplace's method [9], we consider a direct numerical approach based on computing the posterior density over a grid for the state $(x, y, v_x, v_y)$, centered around the Bayesian estimate ($\widehat{\theta}_{\text{Bayes}}$).

Figure 7 illustrates the one-shot fusion pipeline for a distributed setup with two radar nodes.

### B. Experimental Results

We evaluate the one-shot fusion algorithm for the two-node network for the configuration and target trajectories described in Fig. 4. Due to the absence of exact ground truth for target position and velocity at each frame, we use track-level fusion estimates [10] as a benchmark for our one-shot estimates. Viewing self-calibration as "training" of the radar network, and
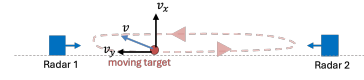


Fig. 8: Geometric degeneracy in configuration C

one-shot fusion as "inference" based on the trained network, we use different tracks for these two stages. Thus, for each configuration, we use the straight line trajectory for self-calibration when evaluating one-shot fusion for the random trajectory, and vice versa.

We set less informative priors ($\sigma_x = \sigma_y = 3\,m$) for position estimates as they are more accurate when using range and angle measurements. In contrast, Doppler measurements only provide estimates of radial velocity. Therefore, we apply stronger priors ($\sigma_{v_x} = \sigma_{v_y} = 3.5\,m/s$) by incorporating our knowledge of average human walking and running speeds. For measurement noise, we use $\sigma_R = 0.035\,m$, $\sigma_\Omega = \pi/4$, and $\sigma_V = 0.1807\,m/s$, which are the range, spatial frequency, and Doppler resolutions, respectively.

Geometric degeneracy for one-shot vector velocity estimation occurs whenever the target lies on or near the straight line connecting the two radars, since both radars are then measuring the same radial velocity (with sign flipped). This corresponds to poorly conditioned maximum likelihood fusion. While this can happen for any of the three configurations shown in Fig. 4, configuration C, in which the radar arrays face each other, is the most vulnerable, since this happens for a significant portion of the common FoV (Fig. 8). Bayesian fusion incorporating realistic priors for target motion addresses such degeneracies and significantly reduces RMSE values, as shown in Table III.

More generally, the quality of one-shot position and velocity estimates depends on the FoV diversity provided by collaborating nodes. Ideally, the proportion of the common FoV containing targets approximately along the line between radars (leading to velocity estimate degeneracy) or at highly correlated angles of arrival (causing poor position estimates) should be small. Among the three configurations, configuration B offers the best FoV diversity, while configuration C offers the least. This is consistent with the position and velocity RMSE values in Table III: configuration B achieves the best estimates, while configuration C yields the poorest velocity estimates. As illustrated in Figure 10, however, geometric degeneracy can still occur in configuration B, reflected in increased uncertainty in the posterior distribution.

To further validate the accuracy of the proposed one-shot fusion approach, we provide a detailed comparison of the position $(X, Y)$ and velocity estimates $(V_X, V_Y)$ for a random trajectory using configuration B, where self-calibration is performed with the straight trajectory. The estimates are obtained from: 1) the Extended Kalman Filter (EKF) at $S_1$, 2) the EKF at $S_2$ transformed into the reference frame of $S_1$, and 3) one-shot fusion, as shown in Fig. 9. We observe that the EKF estimates from both $S_1$ and $S_2$ align closely, confirming the effectiveness of self-calibration. Although the one-shot velocity estimates exhibit more variations than the
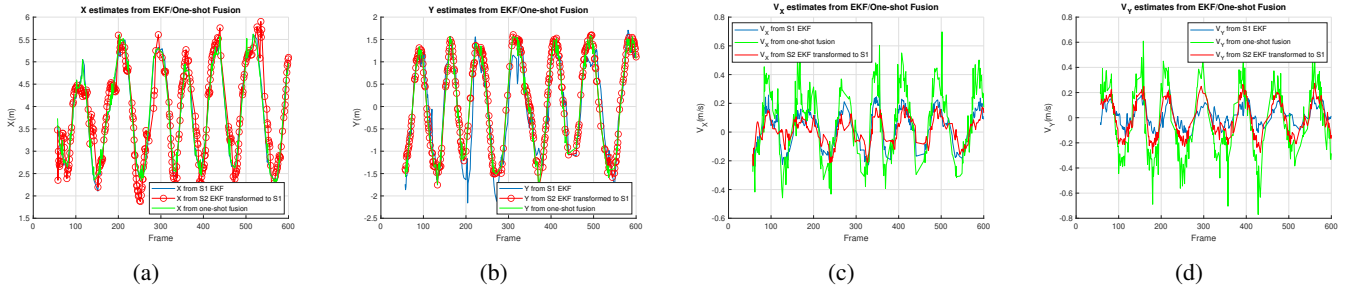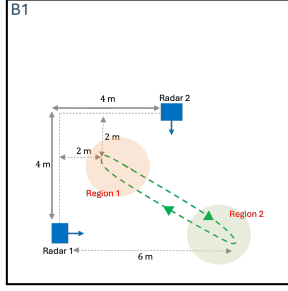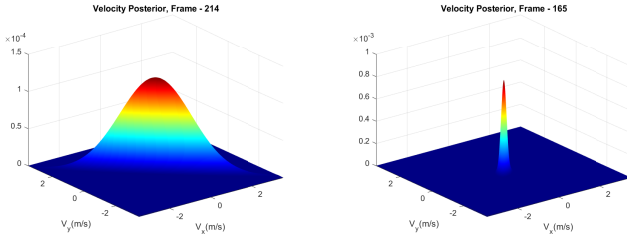
Fig. 9: $X, Y, V_X, V_Y$ estimates for Setup B random trajectory, after self-calibration using a straight trajectory



(a) Region 1 exhibits geometric degeneracy



(b) Region 1 Posterior



(c) Region 2 Posterior

Fig. 10: The velocity posterior distribution (b) for a target in region 1 is broader than the posterior (c) for a target in region 2, as expected.

| | A | | B | | C | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 1 | 2 | 1 | 2 |
| **Position RMSE (m)** | | | | | | |
| Bayesian | 0.3396 | 0.3252 | 0.1699 | 0.1438 | 0.3846 | 0.3690 |
| ML | 0.3404 | 0.3248 | 0.1701 | 0.1438 | 0.3839 | 0.3691 |
| **Velocity RMSE (m/s)** | | | | | | |
| Bayesian | 0.3609 | 0.3205 | 0.3300 | 0.2656 | 0.4139 | 0.4433 |
| ML | 0.3639 | 0.3239 | 0.3439 | 0.2667 | **9.7155** | **6.0543** |

TABLE III: RMSE for one-shot fusion (Bayesian vs ML) with track-level fusion estimates as benchmark. Values indicated in bold indicate cases when ML estimates fail

smoothed estimates provided by the EKF, they clearly follow the same trends. The state covariance matrix, derived from the posterior distribution, allows us to assess the uncertainties tied to the instantaneous state estimates. For example, in Fig. 10, we display the uncertainties in velocity across two regions using the velocity posterior distribution.

## IV. CONCLUSION

In this work, we present and experimentally demonstrate a framework for collaborative sensing in a distributed radar network. We show that simplified point target representations are sufficient for key tasks like self-calibration and one-shot fusion for position and velocity estimation. While our results focus on humans, the approach readily generalizes to other moving targets like vehicles via centroid extraction. The self-calibration accuracy is robust to radar configurations and target trajectory types, while Bayesian one-shot fusion effectively handles geometric degeneracies. This framework lays the foundation for collaborative inference in larger, opportunistically deployed sensor networks. As part of future work, we plan to extend mmSnap to address multi-target scenarios and also eliminate ghost targets caused by multi-path reflections.

## REFERENCES

[1] A. Banik, Y. Mostofi, A. Sabharwal, and U. Madhow, "Optimal self-calibration for collaborative sensing in mmwave radar networks," in *2024 58th Asilomar Conference on Signals, Systems, and Computers*, 2024, pp. 218–222.

[2] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous full-motion estimation of arbitrary objects using dual doppler radar," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 324–329.

[3] S.-G. Lee, J. Jung, and S.-C. Kim, "Enhanced velocity vector estimation using distributed radar system," in *2022 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, 2022, pp. 75–79.

[4] M. Ling, Y. Mao, and L. Zhao, "Velocity vector synthesis method based on distributed millimeter wave radar system," in *2024 16th International Conference on Communication Software and Networks (ICCSN)*, 2024, pp. 317–322.

[5] E. Insanic and P. R. Siqueira, "A maximum likelihood approach to estimation of vector velocity in doppler radar networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 553–567, 2012.

[6] H. Chen, Y. Liu, and Y. Cheng, "A robust robot perception framework for complex environments using multiple mmwave radars," *IEEE Journal of Selected Topics in Signal Processing*, vol. 18, no. 3, pp. 380–395, 2024.

[7] T. Grebner, M. Linder, N. Kern, P. Schoeder, and C. Waldschmidt, "6d self-calibration of the position and orientation of radar sensors in a radar network," in *2022 19th European Radar Conference (EuRAD)*, 2022, pp. 157–160.

[8] A. Shastri, M. Canil, J. Pegoraro, P. Casari, and M. Rossi, "mmscale: Self-calibration of mmwave radar networks from human movement trajectories," in *2022 IEEE Radar Conference (RadarConf22)*, 2022, pp. 1–6.

[9] D. J. C. MacKay, *Information Theory, Inference & Learning Algorithms*. USA: Cambridge University Press, 2002.

[10] S. Matzka and R. Altendorfer, "A comparison of track-to-track fusion algorithms for automotive sensor fusion," in *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2008, pp. 189–194.