

Less is More: Physical-Enhanced Radar-Inertial Odometry

Qiucan Huang, Yuchen Liang, Zhijian Qiao, Shaojie Shen, and Huan Yin

Abstract—Radar offers the advantage of providing additional physical properties related to observed objects. In this study, we design a physical-enhanced radar-inertial odometry system that capitalizes on the Doppler velocities and radar cross-section information. The filter for static radar points, correspondence estimation, and residual functions are all strengthened by integrating the physical properties. We conduct experiments on both public datasets and our self-collected data, with different mobile platforms and sensor types. Our quantitative results demonstrate that the proposed radar-inertial odometry system outperforms alternative methods using the physical-enhanced components. Our findings also reveal that using the physical properties results in fewer radar points for odometry estimation, but the performance is still guaranteed and even improved, thus aligning with the “less is more” principle.

I. INTRODUCTION

Radar-based state estimation has a rich history that traces its origins back to the previous century [1]. In recent years, it has become a new trend within the robotics community, primarily driven by significant advancements in addressing challenging conditions [2]. In contrast to the widely adopted visual and laser sensors, radar sensors offer a consistent source of sensing information that remains reliable under all weather conditions, even in adverse scenarios such as dark or foggy indoor environments. This robustness makes radar a dependable sensor, especially for mobile robots operating in demanding and challenging environments.

There is no such thing as a free lunch. The primary challenge arises from radar data being much noisier compared to other sensing modalities, such as data-rich visual images and high-quality laser points. This inherent noise poses significant challenges to radar-based state estimation, especially for invalid or dynamic point elimination and correspondence building with noises. Researchers have proposed several approaches to build a robust radar-based state estimation system. These approaches broadly align with two research directions. The first direction focuses on using novel radar sensors [3] or sensing modalities [4], designed to deliver dense radar data. It’s worth noting that these solutions often come with constraints, such as weight, power supply, and other domain-specific requirements, which restrict their deployment on general robotic platforms, particularly small-sized mobile robots intended for commercial use.

This work was supported in part by the HKUST Postgraduate Studentship, in part by the HKUST-DJI Joint Innovation Laboratory, and in part by the Hong Kong Center for Construction Robotics (InnoHK center supported by Hong Kong ITC).

The authors are with the Cheng Kar-Shun Robotics Institute, Hong Kong University of Science and Technology, Hong Kong SAR. E-mail: qhuan-gag@connect.ust.hk, yliangbk@connect.ust.hk, zqiaoac@connect.ust.hk, eeshaojie@ust.hk, eehyin@ust.hk

Corresponding author: Huan Yin

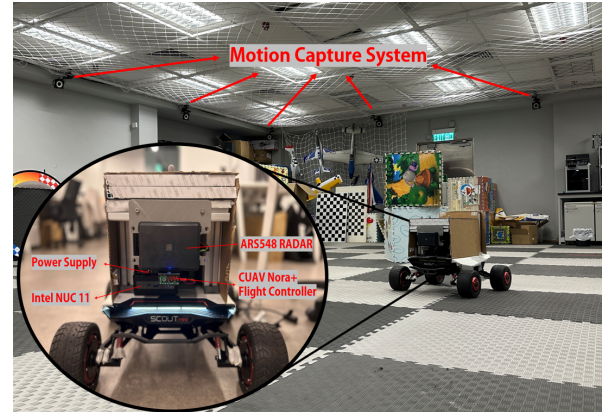


Fig. 1. The mobile platform is equipped with a radar sensor and within the experimental room, where a motion capture system is employed to ensure accurate ground truth data for evaluation.

The alternative direction involves the development of estimation systems tailored for conventional radar-on-a-chip sensors [5]–[10]. This approach aligns with the advanced methodologies employed in visual/LiDAR-inertial odometry systems (VIO/LIO) [11], [12]. However, radar points are generally regarded as sparse velocity-aided points in a typical visual or LiDAR matching manner, and the dynamic points and noisy correspondences are not well handled. We consider that a more comprehensive consideration of radar *physical properties* is desired to improve the radar-based state estimation with noises. In this study, we introduce a tightly coupled and physical-enhanced Radar-Inertial Odometry (RIO) method, in which radar’s physical properties are leveraged from the front-end radar point filter to the back-end optimization. Specifically, we utilize *Doppler velocities* and inertial measurement unit (IMU) to obtain static radar points, by formulating an IMU-aided velocity check scheme to filter radar points. This approach enhances the system’s robustness to noise generated from radar sensors. Additionally, we incorporate *radar cross section* (RCS), also known as radar signatures, to enhance point-to-point correspondence estimation within the proposed RIO system.

We conduct real-world experiments using both publicly available data and data collected in our motion capture room. Our proposed RIO system stands out by utilizing fewer radar points, a design choice informed by the intrinsic physical properties of radar. Remarkably, as validated in the experiments, using the physical-enhanced filters leads to more accurate metrics compared to those without filters, aligning with the “less is more” principle when guided by radar properties. In summary, our specific contributions and

findings can be outlined as follows:

- We design an IMU-aided velocity check scheme for Doppler velocities and static radar points can be obtained under this scheme.
- We leverage the RCS information to refine point-to-point correspondences, which improves the radar point matching-based pose estimation.
- The proposed RIO system is thoroughly evaluated through real-world ablation studies and comparisons, verifying the effectiveness of using radar physical properties in this study.

II. RELATED WORK

The survey paper [2] by Harlow *et al.* reviews the recent radar applications in robotics comprehensively. In this section, we will delve into the research works related to radar-(inertial) odometry.

Several datasets have been made available for Frequency Modulated Continuous Wave (FMCW) Navtech sensors, contributing significantly to the field of 360° radar sensing [3], [13], [14]. These datasets have played a pivotal role in advancing radar studies in recent years. Early Navtech radar-based robot odometry primarily relied on traditional visual-inspired keypoints and matching techniques [15], [16]. Adolfsson *et al.* introduced a high-precision radar odometry approach capable of achieving LiDAR-level precision. Moreover, they developed loop closure detection techniques for global mapping [17], [18]. Advanced machine learning techniques have been applied to enhance radar odometry [19], [20], rendering the entire framework data-driven and more resilient to noise. In addition to odometry, the Navtech radar-based community has also delved into place recognition [21], metric localization [22], and Simultaneous Localization and Mapping (SLAM) [23]. However, it's worth noting that Navtech radar sensors have relatively high weight (6 kg) and power requirements (30 W) compared to conventional embedded radar sensors, thereby constraining their application in small-sized robotic platforms.

In recent years, there have been notable contributions from Kramer *et al.* [24] and Cheng *et al.* [4] in the realm of generating dense sensing using lightweight and low-power radar-on-a-chip sensors, particularly widely-used single or cascaded chips. To achieve dense sensing on such chips, specialized domain knowledge is essential for customizing sensor readings. In the context of mapping, Li *et al.* [25] proposed a pose graph-based radar mapping using the ZF 4D imaging radar. The 4DRadarSLAM system [26] used the Oculii Eagle radar sensor, which offers a relatively dense angle resolution, facilitating more precise mapping.

Another research direction involves utilizing radar data produced by cheaper radar chips, with an integration using IMU sensors. For ego-velocity estimation, Kramer *et al.* introduced a method that combines Doppler velocity measurements and inertial measurements through sliding-window optimization [5]. Doer and Trommer devised an Extended Kalman Filter (EKF)-based radar inertial odometry approach for sensor fusion [6]. In 4D iRIOM [7], Doppler

velocity was employed to filter noisy data, and a distribution-to-multi-distribution matching approach inspired by LiDAR was developed for pose estimation [7]. Michalczyk *et al.* proposed multi-state and tightly-coupled state estimation systems, where radar points served as landmarks for data matching [8], [9]. In [10], the radar intensity was considered in the estimation of radar-inertial odometry. To address radar data noisiness, Lu *et al.* introduced a deep learning-based scheme for sensor fusion, thereby enhancing accuracy through a data-driven approach [27].

Our radar-inertial odometry system draws inspiration from previous studies within the research community [5]–[10]. The difference is the comprehensive utilization of both Doppler velocities and radar signatures (i.e., RCS) to enhance system performance. Our utilization includes front-end data filters, correspondence estimation, and residual functions. The efficacy of the proposed physical-enhanced RIO system has been validated in real-world scenarios, supported by ablation studies to provide further insights into its functionality.

III. METHODOLOGY

A. System Overview

An overview of the system is presented in Figure 2, which takes two data as input: sparse radar scans and 6Dof IMU data between two scans. The system is based on a sliding window approach that takes multiple frames to estimate the trajectory of RIO.

The complete state of our system is defined as follows:

$$\begin{aligned} \mathbf{X} &= [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_k^T, \mathbf{l}_0^T, \mathbf{l}_1^T, \dots, \mathbf{l}_n^T]^T \\ \mathbf{x}_i &= [\mathbf{p}_i^T, \mathbf{q}_i^T, \mathbf{v}_i^T, \mathbf{b}_{ai}^T, \mathbf{b}_{gi}^T]^T, \quad i \in [0, k] \\ \mathbf{l}_j &\in \mathbb{R}^3, \quad j \in [0, n] \end{aligned} \quad (1)$$

in which \mathbf{p}_i , \mathbf{q}_i and \mathbf{v}_i represent the position, rotation, and velocity of the i -th frame, respectively. In the rest of this paper, we use \mathbf{R}_i to represent the equivalent rotation matrix of the quaternion \mathbf{q}_i . The terms \mathbf{b}_{ai} , \mathbf{b}_{gi} denote the accelerometer bias and gyroscope bias in the i -th frame, respectively. Both \mathbf{b}_{ai} and \mathbf{b}_{gi} are modeled as random walk processes characterized by zero-mean Gaussian noise. The term \mathbf{l}_j refers to the position of the j -th landmark in the world frame.

The optimization problem for our system is formulated as follows,

$$\min_{\mathbf{X}} \left\{ \sum_{k \in I} \|\mathbf{r}_I(\mathbf{x}_k, \mathbf{x}_{k+1})\|_{\mathbf{P}_k^{k+1}} + \sum_{k \in D} w_D \|\mathbf{r}_D(\mathbf{x}_i, p_k)\|_2 + \sum_{k \in P} w_P \|\mathbf{r}_P(\mathbf{x}_i, \mathbf{l}_k, p_k)\|_2 \right\} \quad (2)$$

where p represents a radar point, and w_D , w_P are weights of residuals. The terms $\mathbf{r}_I(\cdot)$, $\mathbf{r}_D(\cdot)$ and $\mathbf{r}_P(\cdot)$ are three different residuals, with further illustrations provided in Section III-B, Section III-D, and Section III-E, respectively.

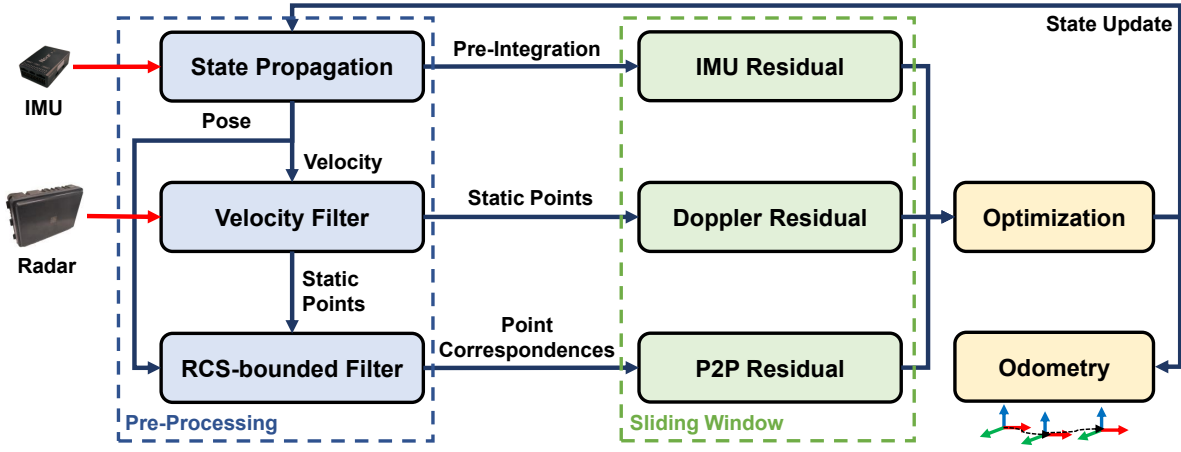


Fig. 2. The overview of our proposed physical-enhanced RIO system. The system follows a sliding window-based optimization framework, and with pre-processing modules to handle noisy points and correspondences. Specifically, Doppler velocities are used to obtain static radar points within an IMU-aided velocity check scheme; RCS information is leveraged to bound the neighboring search for point-to-point correspondence estimation.

B. IMU Pre-integration

Our system utilizes IMU data propagation to estimate continuous radar states. When there are changes in the previous states, re-propagation becomes necessary. However, conducting re-propagation across the entire sliding window can be computationally time-consuming. To circumvent the need for re-propagation, we employ IMU pre-integration, a technique that allows us to streamline the process, as introduced in [11]. By employing pre-integration, the residual of IMU measurement is defined as:

$$\mathbf{r}_I(\mathbf{x}_k, \mathbf{x}_{k+1}) = \begin{bmatrix} \mathbf{R}_k^T (\mathbf{p}_{k+1} - \mathbf{p}_k + \frac{1}{2}g\Delta t^2 - \mathbf{v}_k\Delta t) - \Delta \mathbf{p}_{k+1}^k \\ 2[\mathbf{q}_k^{-1} \otimes \mathbf{q}_{k+1} \otimes \Delta \mathbf{q}_{k+1}^k]_{xyz} \\ \mathbf{R}_k^T (\mathbf{v}_{k+1} + g\Delta t - \mathbf{v}_k) - \Delta \mathbf{v}_{k+1}^k \\ \mathbf{b}_{ak+1} - \mathbf{b}_{ak} \\ \mathbf{b}_{gk+1} - \mathbf{b}_{gk} \end{bmatrix} \quad (3)$$

where $[\cdot]$ extract the vector part of quaternion and the symbol \otimes denotes the Hamilton product; $\Delta \mathbf{p}_{k+1}^k$, $\Delta \mathbf{q}_{k+1}^k$ and $\Delta \mathbf{v}_{k+1}^k$ are results of IMU pre-integration; Δt is the time duration between two scans. Gravity, denoted as g , is assumed to align with the z -axis.

C. Radar Pre-processing

Radar scans frequently include a multitude of invalid points owing to multi-path reflection. Integrating the original data without pre-processing could lead to a failure in state estimation. Consequently, it becomes imperative to eliminate these invalid data points or noises.

The proposed initial step involves discarding data points that lie beyond the current radar settings, such as the field of view (FOV), which serves to reduce computational demands. Subsequently, we employ a radius filter as a criterion [7] for identifying invalid points, written as follows:

$$\mathbf{P} = \{p_i | \text{Num}(p_j) > N \text{ and } \|p_i - p_j\| \leq D\} \quad (4)$$

where N and D are predefined thresholds; \mathbf{P} represents the set of retained points; p_j denotes a neighbor of radar point

p_i . We eliminate points that are excessively far from other points by applying this filter, as scattered points are more likely to be invalid in obtained noisy data.

Radar scans occasionally include false points with error velocities, and the presence of such detection can potentially degrade the estimation performance. A natural approach is to use Doppler velocities for object removal. However, considering the sparse property of radar, velocity estimated solely from radar data may prove inaccurate in specific scenarios, such as when a sizable moving object obstructs the radar's view. To tackle this challenge, we design an IMU-aided velocity check scheme to filter false points, i.e., to obtain real static radar points in environments.

Specifically, we calculate the estimated Doppler velocity v_{est} of radar point p_j as follows:

$$v_{est} = \frac{p_j^T}{\|p_j\|} \cdot \mathbf{R}_E^T (\mathbf{R}_i^T \mathbf{v}_i + (\hat{\omega}_i - \mathbf{b}_{gt})^\wedge \mathbf{t}_E) \quad (5)$$

where \mathbf{R}_i and \mathbf{v}_i correspond to the rotation and velocity, of the i -th radar scan, which are obtained through state propagation. The extrinsic transformation from the radar to the IMU is denoted as \mathbf{R}_E and \mathbf{t}_E , respectively. The term \mathbf{b}_{gt} corresponds to the IMU gyroscope bias, and $\hat{\omega}_i$ denotes the angular velocity measured by the IMU. The operator $(\cdot)^\wedge$ represents the skew-symmetric matrix of the vector \mathbf{v} , which is defined as:

$$\mathbf{v} = [v_1, v_2, v_3]^T, \mathbf{v}^\wedge = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (6)$$

A radar point is recognized as an inlier if the 2-norm error between the measured Doppler velocity and the estimated Doppler velocity point is within a small threshold, signifying consistency between the two. Then outliers are subsequently excluded from consideration, leaving only the static points with correct velocity measurements for further estimation.

D. Residual on Doppler Velocity

A radar scan typically includes measurements of azimuth, elevation, range, and Doppler velocity for the detection of targets. In an ideal scenario, the Doppler velocity measured by the radar should align with the projection of the relative velocity between an object and the radar onto the direction from the object to the radar origin.

In the pre-processing step, dynamic points are filtered out with the IMU-aided scheme. With static radar points, the estimation process for the Doppler velocity of reserved static points is described by Equation 5. Then we define the residual of Doppler velocity measurements as follows:

$$\mathbf{r}_D(\mathbf{x}_i, p_k) = \frac{p_k^T}{\|p_k\|} \cdot \mathbf{R}_E^T (\mathbf{R}_i^T \mathbf{v}_i + (\hat{\omega}_i - \mathbf{b}_{gt})^\wedge \mathbf{t}_E) - v_{dk} \quad (7)$$

where v_{dk} is the Doppler velocity of p_k ; and p_k is the radar measurement in the i -th radar scan under a Cartesian coordinate system with the radar origin as its origin point.

E. RCS-bounded Filter for Point-to-Point Residual

Correspondence plays a pivotal role in point-based matching methods. Previous studies have demonstrated that tracking radar points between subsequent scans is possible. However, there will still be noises or invalid points in the reserved points after filters, the correspondence estimated by searching for the nearest neighbors in Euclidean space might not always be precise. To improve the robustness of matching, we integrate RCS in our approach.

We start with the definition of this physical property. According to [28], the relationship between received and transmitted power can be described as follows:

$$\frac{P_R}{P_E} = \frac{c G_E G_R \sigma}{(4\pi)^2 r^4} \quad (8)$$

in which σ is the value of RCS. P_R and P_E are received power and transmitted power; G_R and G_E are the received gain and transmitted gain. The term c is a constant coefficient, and r is the distance between the radar origin to the detected object.

Though the RCS of a simple object can be analytically determined, real-world environments are typically less than ideal, and the robot motions need to be considered. In the field of computer vision, optical flow [29] relies on the assumption that pixel intensities of an object remain consistent between consecutive frames. This assumption renders optical flow particularly well-suited for visual-inertial navigation systems, as demonstrated in previous work [11]. Taking inspiration from the principles of optical flow, we assume that slight changes in the pose between radar scans have minimal influence on the RCS, i.e., RCS of an object remains consistent between two consecutive radar scans.

We then conduct correspondence estimation between two consecutive scans guided by our assumption. To achieve this, we first transform all points in the current scan to the coordinate system of the previous scan. When considering

a target point p_j in these two scans, we design an RCS-bounded nearest neighbor point search, described as follows:

$$\begin{aligned} \mathbf{P}_{rcs} &= \{p_i | \|p_i - p_j\| \leq D \text{ and } \|\sigma_i - \sigma_j\| \leq d\} \\ \mathbf{P}_m &= \{p_k | \forall p_i \in \mathbf{P}_{rcs}, p_k \in \mathbf{P}_{rcs}, \|p_i - p_j\| \geq \|p_k - p_j\|\} \end{aligned} \quad (9)$$

in which D and d are the distance threshold and the RCS threshold; σ_i and σ_j represent the RCS values of the detected objects. \mathbf{P}_m presents the matched point. Considering that noisy correspondences will still remain after RCS and distance filter, we also employ a landmark selection strategy to make the estimation more robust. To achieve that, we select the closest one as the matched point and only consider the radar points that detected more than a certain time, i.e., only repeatable points in multiple frames are regarded as reliable landmarks for optimization.

Finally, with the estimated correspondences, the final 3D point-to-point (P2P) residual is defined as:

$$\mathbf{r}_P(\mathbf{x}_i, \mathbf{l}_k, p_k) = \mathbf{l}_k - (\mathbf{R}_i (\mathbf{R}_E p_k + \mathbf{t}_E) + \mathbf{p}_i) \quad (10)$$

where \mathbf{l}_k is the position of the landmark. p_k is the measurement of \mathbf{l}_k in the i -th scan. We solve the entire optimization problem (Equation 2) using the Ceres-Solver [30] with derived Jacobians.

IV. EXPERIMENTS

In this section, we first introduce the experimental setup of two datasets. Then quantitative results are presented with ablation studies and comparisons.

A. Setup

We demonstrate and test our proposed RIO system on a real-world mobile platform, shown in Figure 1, and also a public dataset ColoRadar [24]. All the experimental results are evaluated using the open-sourced EVO [31].

The mobile platform is a ground robot equipped with ARS548RDI¹, a cascaded radar kit from Continental, and an IMU². To generate ground truth poses, we employed a motion capture (MoCap) system. The extrinsic parameters between IMU and Radar are manually calibrated. We collected three distinct sequences for evaluation, each varying in difficulty: Sequence 1 involves relatively low-speed movements; Sequence 2 is collected at higher speeds compared to Sequence 1; Sequence 3 is the most challenging one, characterized by high rotation and translation speeds.

We also use the ColoRadar dataset to assess the generalizability of our methods. ColoRadar employs a single-chip Texas Instruments AWR1843BOOST-EVM, and the data is sparser compared to the cascaded radar used in our self-collected dataset. We selected one indoor sequence “12.21.2020_arpg_lab_run0” (denoted as ColoRadar 1) and one outdoor sequence “2.28.2021_outdoors_run0” (denoted as ColoRadar 2). These two sequences only provide the intensity of each scanned point, which essentially represents the relation between the received and transmitted power.

¹<https://conti-engineering.com/components/ars-548-rdi/>

²<https://www.cuav.net/en/nora-plus/>

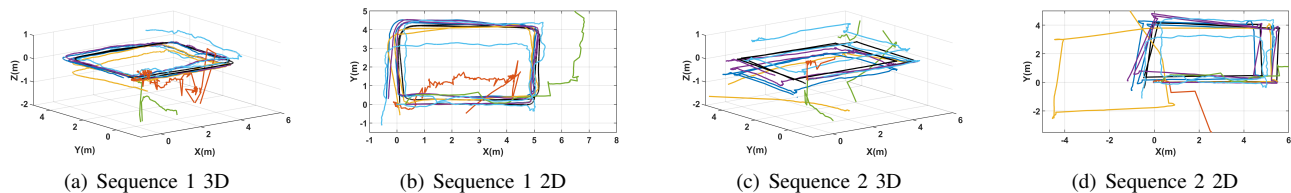


Fig. 3. Distinct colors are employed to represent different 3D and 2D (x-y) trajectories for ablation study. **Ground Truth**: obtained from the MoCap system. **Full System**: all components enabled. The other ablated variants are listed as follows: **w/o IMU Residual**, **w/o Velocity Residual**, **w/o P2P Residual**, **w/o Velocity Filter**, **w/o RCS Filter**, , and all with analyses in Section IV-C.

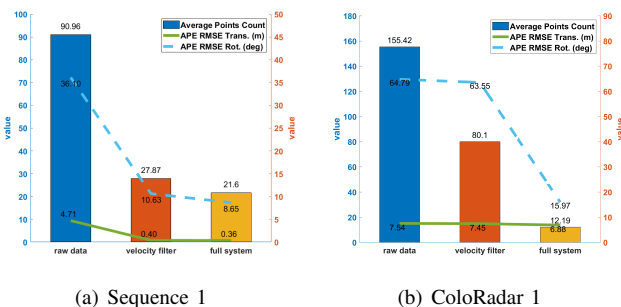


Fig. 4. Average radar point number (left Y-axis) and rotational/positional RMSE (right Y-axis) by applying physical-enhanced filters consecutively: IMU-aided velocity check scheme (denoted as velocity filter) and RCS-bounded filter (denoted as full system).

Thus we first remove the distance factor in Equation 8 and regard the remaining term as the RCS value. For ColoRadar 1, we employ the MoCap data provided by the dataset as the ground truth. As for ColoRadar 2, the ground truth is obtained from the high-precision LiDAR SLAM system due to the lack of MoCap system. These two sequences were collected using a handheld platform, and aggressive motions bring challenges to the RIO systems.

B. Less is More

In the test phase with two datasets, we record the average number of points and the Root Mean Square Error (RMSE) of absolute trajectory error (APE) by employing two filters consecutively: first, the velocity filter to obtain static points, followed by the RCS filter for P2P correspondence estimation. As depicted in Figure 4, these two physical-enhanced filters result in a reduction in the radar points utilized for state estimation. Concurrently, RMSE also experiences a decrease when using cleaner radar points, thus underscoring the effectiveness of the “less is more” principle in this study. Figure 5 presents a visualization of the point cloud map generated from the full RIO system.

C. Ablation Study

To assess each component, we performed an ablation study using self-collected Sequence 1 and Sequence 2. We test five ablated variants of our proposed RIO system, each with specific components disabled: IMU Residual, Velocity Residual, P2P Residual, Velocity Residual, Velocity Filter and RCS Filter. The trajectory results of the ablation study

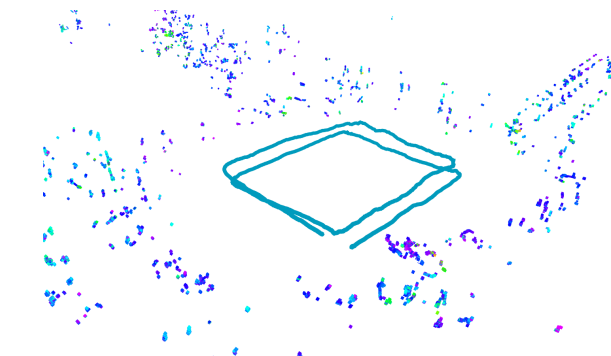


Fig. 5. Visualization of the radar point cloud map and the traveled trajectory using our proposed RIO. Darker points indicate the radar points with larger RCS values.

are visualized in Figure 3, with detailed position information plotted in Figure 6. Additionally, the statistics of the RMSE can be found in Table I.

First of all, the full system without any disabled residuals and filters achieves the best performance, meaning that each component contributes to the system. Specifically, we can conclude the following findings from the ablation study:

IMU Residual. The absence of IMU residual $r_I(\cdot)$ results in a rapid accumulation of pose errors. This underscores the fundamental role that the residual $r_I(\cdot)$ plays in the system.

Velocity Residual. If the Doppler velocity residual $r_D(\cdot)$ is disabled, there will be a performance drop for the RIO system. This demonstrates the indispensability of the Doppler property for accurate pose estimation.

P2P Residual. We find that the point-to-point residual $r_P(\cdot)$ reduces the error when the robot running at a relatively high speed. This indicates that the proposed point-to-point model could improve the estimation with aggressive motions.

Velocity Filter. If the velocity filter is disabled, more dynamic radar points will be involved in the estimation, resulting in a large performance drop for RIO. This indicates the importance of static radar points and validates the effectiveness of our IMU-aided velocity check scheme.

RCS Filter. We also disable the RCS-bounded filter for correspondence estimation, i.e., only distance-based neighbor search is used. As a result, there is a slight performance decrease. We find that its effect becomes prominent when the robot’s speed is high, as it could provide more reliable correspondences to stabilize the pose with high speed.

TABLE I
QUANTITATIVE RESULTS OF ABLATION STUDY

Seq.	method	APE RMSE		RPE RMSE	
		Trans.(m)	Rot.(°)	Trans.(m)	Rot.(°)
1	w/o IMU Res.	28.067	137.914	3.623	23.690
	w/o Vel. Res.	1.118	11.283	0.049	0.613
	w/o P2P Res.	0.391	8.231	0.030	0.642
	w/o Vel. Filter	6.441	64.394	0.384	11.617
	w/o RCS Filter	0.299	8.650	0.027	0.659
	Full System	0.294	8.194	0.027	0.599
2	w/o IMU Res.	148.652	135.747	33.517	70.960
	w/o Vel. Res.	0.798	10.589	0.064	0.898
	w/o P2P Res.	8.528	10.898	0.370	0.987
	w/o Vel. Filter	4.709	36.104	0.594	8.805
	w/o RCS Filter	0.396	10.635	0.062	0.901
	Full System	0.358	8.646	0.060	0.879

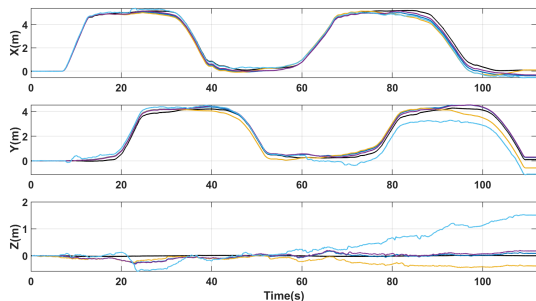


Fig. 6. Positional errors of Sequence 1 for ablation study. Distinct colors are employed to represent different ablated variants. The color scheme used is consistent with the one depicted in Figure 3.

D. Comparisons

In the following experiments, we compare our method against EKF-RIO [6], which is considered one of the best-performing open-source radar-inertial odometry solutions. In addition, we test two alternative approaches for building point correspondences and residual functions:

- Point-to-Distribution (P2D): This approach is based on the widely used laser-scan matching techniques. It takes into account the distribution of the current point in relation to the previous map and assigns weights to the distance between the point and the distribution’s center, factoring in their covariance.
- Distribution-to-Multi-Distribution (D2M) (proposed in [7]): In contrast to the P2D, this approach constructs a distribution for each current point by searching for its neighbors. It then assigns weights to the distance between the center of this distribution and the distributions of its neighbors based on their covariance.

We use the open-source version of EKF-RIO and implement the other two methods by ourselves. For the P2D and D2M, we disable our P2P matching and exclude point states during optimization. We also construct submaps to enhance their performance, following the guidelines presented in [7].

Table II presents the RMSE of all comparisons. In summary, our system outperforms other methods when using our self-collected cascaded radar data. However, in the case

TABLE II
QUANTITATIVE RESULTS OF COMPARISON STUDY

Sequence	method	APE RMSE		RPE RMSE	
		Trans.(m)	Rot.(°)	Trans.(m)	Rot.(°)
Sequence 1	EKF-RIO	0.800	15.338	0.114	0.499
	P2D	0.731	13.693	0.030	0.720
	D2M	0.565	17.007	0.030	0.640
	Ours	0.294	8.194	0.027	0.599
Sequence 2	EKF-RIO	2.805	8.835	0.392	0.866
	P2D	0.701	8.943	0.060	0.948
	D2M	83.841	8.883	2.065	0.806
Sequence 3	Ours	0.358	8.646	0.060	0.879
	EKF-RIO	68.960	66.228	2.195	2.111
	P2D	1.863	12.298	0.057	1.917
ColoRadar 1	D2M	3.180	17.717	0.274	2.408
	Ours	1.098	9.344	0.055	1.619
	EKF-RIO	5.348	24.416	0.054	6.044
	P2D	6.265	21.837	0.052	0.963
ColoRadar 2	D2M	6.212	20.254	0.057	0.640
	Ours	6.877	15.973	0.049	0.955
	EKF-RIO	10.692	14.764	0.047	0.497
	P2D	12.371	15.913	0.058	0.830
ColoRadar 2	D2M	11.076	11.069	0.054	0.652
	Ours	8.789	19.578	0.045	0.966

of the single-chip radar of the ColoRadar dataset, which generates noisier radar measurements, all methods struggle to achieve high accuracy. Specifically, EKF-RIO deals with this issue by discarding all radar measurements, which actually bring drifts in both translation and rotation. As for the P2D and D2M-based methods, these two consider the geometry information around the radar points for matching. However, compared to laser scans, the distribution may not be estimated correctly due to the noise and sparsity of radar data, thus resulting in larger errors of RIO.

Our proposed RIO filters out unreliable radar points with radar physical properties and tracks the reserved high-quality points in a sliding window. These key points reduce the estimation errors compared to other approaches, therefore our method has better performance in terms of translation since the key points are more stable in translation motion. But on the other hand, the loss of tracking of the same radar points in different radar scans still exists. The rotational motions in the ColoRadar dataset are more aggressive, and the loss of tracking may decrease the performance of methods in terms of rotation. Overall, our method performs at least as well as the aforementioned methods in cascaded millimeter wave radar, but still has some improving space in the future.

V. CONCLUSION

In this study, we design a physical-enhanced radar-inertial odometry system. The physical properties are leveraged to obtain static radar points and build robust pose estimation. The proposed RIO is validated with multiple sequences and two different radar sensor types, and also with comprehensive ablation study analysis and comparisons. Considering the natural properties of radar, there are several promising directions to improve the performance of our RIO system in the future, e.g., integrating graph neural networks for correspondences, and modeling the uncertainty models to weight the residual functions.

REFERENCES

- [1] S. Clark and G. Dissanayake, "Simultaneous localization and map building using millimeter wave radar to extract natural features," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, vol. 2. IEEE, 1999, pp. 1316–1321.
- [2] K. Harlow, H. Jang, T. D. Barfoot, A. Kim, and C. Heckman, "A new wave in robotics: Survey on recent mmwave radar applications in robotics," *arXiv preprint arXiv:2305.01135*, 2023.
- [3] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 6433–6438.
- [4] Y. Cheng, J. Su, M. Jiang, and Y. Liu, "A novel radar point cloud generation method for robot environment perception," *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3754–3773, 2022.
- [5] A. Kramer, C. Stahoviak, A. Santamaria-Navarro, A.-A. Agha-Mohammadi, and C. Heckman, "Radar-inertial ego-velocity estimation for visually degraded environments," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 5739–5746.
- [6] C. Doer and G. F. Trommer, "An ekf based approach to radar inertial odometry," in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2020, pp. 152–159.
- [7] Y. Zhuang, B. Wang, J. Huai, and M. Li, "4d iriom: 4d imaging radar inertial odometry and mapping," *IEEE Robotics and Automation Letters*, 2023.
- [8] J. Michalczyk, R. Jung, and S. Weiss, "Tightly-coupled ekf-based radar-inertial odometry," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 12 336–12 343.
- [9] J. Michalczyk, R. Jung, C. Brommer, and S. Weiss, "Multi-state tightly-coupled ekf-based radar-inertial odometry with persistent landmarks," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 4011–4017.
- [10] Y. Almalioglu, M. Turan, C. X. Lu, N. Trigoni, and A. Markham, "Milli-rio: Ego-motion estimation with low-cost millimetre-wave radar," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3314–3323, 2020.
- [11] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [12] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lid2: Fast direct lidar-inertial odometry," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.
- [13] G. Kim, Y. S. Park, Y. Cho, J. Jeong, and A. Kim, "Mulran: Multimodal range dataset for urban place recognition," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 6246–6253.
- [14] M. Sheeny, E. De Pellegrin, S. Mukherjee, A. Ahrabian, S. Wang, and A. Wallace, "Radiate: A radar dataset for automotive perception in bad weather," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1–7.
- [15] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 6045–6052.
- [16] —, "Radar-only ego-motion estimation in difficult settings via graph matching," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 298–304.
- [17] D. Adolphsson, M. Magnusson, A. Alhashimi, A. J. Lilienthal, and H. Andreasson, "Lidar-level localization with radar? the cclear approach to accurate, fast, and robust large-scale radar odometry in diverse environments," *IEEE Transactions on robotics*, vol. 39, no. 2, pp. 1476–1495, 2022.
- [18] D. Adolphsson, M. Karlsson, V. Kubelka, M. Magnusson, and H. Andreasson, "Tbv radar slam—trust but verify loop candidates," *IEEE Robotics and Automation Letters*, 2023.
- [19] D. Barnes, R. Weston, and I. Posner, "Masking by moving: Learning distraction-free radar odometry from pose information," *arXiv preprint arXiv:1909.03752*, 2019.
- [20] K. Burnett, D. J. Yoon, A. P. Schoellig, and T. D. Barfoot, "Radar odometry combining probabilistic estimation and unsupervised feature learning," *arXiv preprint arXiv:2105.14152*, 2021.
- [21] H. Yin, X. Xu, Y. Wang, and R. Xiong, "Radar-to-lidar: Heterogeneous place recognition via joint learning," *Frontiers in Robotics and AI*, vol. 8, p. 661199, 2021.
- [22] H. Yin, R. Chen, Y. Wang, and R. Xiong, "Rall: end-to-end radar localization on lidar map using differentiable measurement model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6737–6750, 2021.
- [23] Z. Hong, Y. Petillot, A. Wallace, and S. Wang, "Radarslam: A robust simultaneous localization and mapping system for all weather conditions," *The International Journal of Robotics Research*, vol. 41, no. 5, pp. 519–542, 2022.
- [24] A. Kramer, K. Harlow, C. Williams, and C. Heckman, "Coloradar: The direct 3d millimeter wave radar dataset," *The International Journal of Robotics Research*, vol. 41, no. 4, pp. 351–360, 2022.
- [25] X. Li, H. Zhang, and W. Chen, "4d radar-based pose graph slam with ego-velocity pre-integration factor," *IEEE Robotics and Automation Letters*, 2023.
- [26] J. Zhang, H. Zhuge, Z. Wu, G. Peng, M. Wen, Y. Liu, and D. Wang, "4dradar slam: A 4d imaging radar slam system for large-scale environments based on pose graph optimization," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 8333–8340.
- [27] C. X. Lu, M. R. U. Saputra, P. Zhao, Y. Almalioglu, P. P. De Gusmao, C. Chen, K. Sun, N. Trigoni, and A. Markham, "milliego: single-chip mmwave radar aided egomotion estimation via deep sensor fusion," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020, pp. 109–122.
- [28] C. Schöffmann, B. Ubezio, C. Böhm, S. Mühlbacher-Karrer, and H. Zangl, "Virtual radar: Real-time millimeter-wave radar sensor simulation for perception-driven robotics," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4704–4711, 2021.
- [29] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI'81: 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679.
- [30] S. Agarwal, K. Mierle, and T. C. S. Team, "Ceres Solver," 3 2022. [Online]. Available: <https://github.com/ceres-solver/ceres-solver>
- [31] M. Grupp, "evo: Python package for the evaluation of odometry and slam." <https://github.com/MichaelGrupp/evo>, 2017.