CoSense: Bridging Real-Time Performance and Fine-Grained Detail in mmWave Sensing

Long Fan, *Student Member, IEEE,* Lei Xie, *Senior Member, IEEE,* Shiyuan Ma, *Student Member, IEEE,*
Jingyi Ning, *Member, IEEE,* Chuyu Wang, *Member, IEEE,* Wenhui Zhou, *Student Member, IEEE,*
and Jun Luo, *Fellow, IEEE*

***Abstract*—Millimeter-wave (mmWave) radar offers significant potential for fine-grained sensing, yet transitioning from controlled laboratory environments to dynamic real-world applications remains challenging. Existing methods face a dichotomy: real-time point clouds sacrifice crucial signal details needed for sophisticated tasks, whereas information-rich raw data sensing imposes prohibitive transmission and computation overheads, often limiting analysis to offline settings and hindering real-time viability. To this end, we present CoSense, a *real-time* edge-end collaborative sensing system built on commodity mmWave radar (end) and edge intelligence. We first introduce a novel dual-stream data acquisition mechanism via realizing radar driver-level interfaces, enabling simultaneous transmission of point clouds and raw data. To bridge the fidelity-latency trade-off, we implement an adaptive transmission strategy via firmware modifications, selectively forwarding raw data segments (corresponding to regions of interest identified in the point cloud) for detailed fine-grained analysis, while continuously delivering point clouds for low-latency coarse-grained sensing and control loops. Furthermore, we incorporate closed-loop feedback beamforming, dynamically steering the radar beam based on real-time tracking to counteract motion-induced misalignment and enhance signal fidelity. Extensive evaluations under dynamic conditions demonstrate that CoSense successfully achieves real-time fine-grained sensing with high fidelity and manageable overhead.**

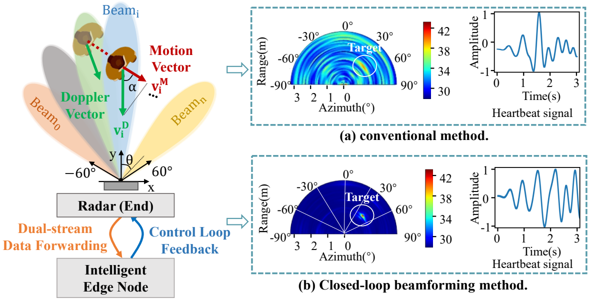
***Index Terms*—real-time sensing, fine-grained sensing, edge-end collaboration, beamforming, millimeter-wave radar**

I. INTRODUCTION

Millimeter-wave (mmWave) technology, a cornerstone for next-generation wireless systems [1]–[3], simultaneously offers potent sensing capabilities owing to its wide bandwidth. Its applications span localization [4]–[10], human activity recognition [11]–[13], speech detection [14]–[18], and vital sign monitoring [19]–[22]. However, a significant challenge remains in translating fine-grained mmWave sensing from controlled laboratory settings to reliable performance in dynamic, real-world environments. Key obstacles include handling unpredictable target mobility [14], [15], [22], operating within the limited processing resources of radar chipsets [16], managing data transmission bandwidth, and mitigating mobility-induced signal degradation. These challenges are critical yet often underestimated, posing barriers to practical deployment.

Current mmWave sensing strategies typically fall into two extremes, each posing fundamental limitations for dynamic, fine-grained applications. One approach relies on low-bandwidth *point cloud* outputs, where radar devices locally preprocess signals to generate 3D spatial information [11], [23], [24]. While this design enables real-time coarse-grained sensing in dynamic scenarios [24]–[26], it achieves efficiency by irreversibly discarding the detailed phase and waveform information crucial for fine-grained analysis (e.g., vital sign monitoring [19], [27], or gesture recognition [11], [12], [23]). Conversely, an alternative paradigm treats the radar primarily as a passive signal recorder, transmitting unprocessed, *information-rich raw data* [14]–[18] to external computing platforms such as laptops. Although this approach benefits from the abundant processing [14], [17] and storage resources available externally [15], [28], its real-time scalability is fundamentally constrained by two factors: the predominance of sensing systems designed for offline laboratory use, and the limited bandwidth between radar sensors and external processors. As a result, these two extremes expose a fundamental trade-off between signal fidelity and real-time responsiveness under conditions of mobility and bandwidth limitation.

To overcome these limitations and bridge the fidelity-latency gap, we introduce CoSense, a system explicitly designed for real-time, fine-grained mmWave sensing in dynamic scenarios. As illustrated in Figure 1, CoSense adopts a collaborative



(a) conventional method.

(b) Closed-loop beamforming method.

Fig. 1. Conceptual overview of CoSense. Left: Edge-end collaboration supports dual-stream communication and adaptive beam steering. Right: Compared to low-SNR conventional sensing (a), CoSense (b) improves signal quality via closed-loop beamforming guided by edge intelligence.

Long Fan, Lei Xie, Shiyuan Ma, Jingyi Ning, Chuyu Wang, and Wenhui Zhou are with the State Key Laboratory for Novel Software Technology, Nanjing University, China (E-mail: fanl@smail.nju.edu.cn, lxie@nju.edu.cn, mashiyuan@smail.nju.edu.cn, ningjy@nju.edu.cn, chuyu@nju.edu.cn, whzhou@smail.nju.edu.cn),

Jun Luo is with the College of Computing and Data Science, Nanyang Technological University (NTU), Singapore 639798 (E-mail: junluo@ntu.edu.sg).

This work is supported in part by Key Projects of Jiangsu Provincial Basic Research Program under Grant No. BK20243040; National Natural Science Foundation of China under Grant Nos. 92467202, 62272216, 62372224, 62502196; Basic Research Program of Jiangsu under Grant No. BK20251207; Postdoctoral Fellowship Program of CPSF under Grant No. GZB20250389; China Postdoctoral Science Foundation under Grant No. 2025M771535; and Collaborative Innovation Center of Novel Software Technology and Industrialization. Lei Xie is the corresponding author.

edge-end architecture, built upon commodity mmWave radar (End) and external intelligent computing devices (Edge). At its core lies a dual-stream data mechanism, enabled through driver-level firmware modifications, which concurrently delivers two complementary data streams: lightweight point clouds for coarse-grained sensing, and information-rich raw data for high-fidelity analysis. This synergistic design enables the edge node to leverage point clouds for efficient real-time feedback and control, while selectively processing raw data segments in regions of interest (RoIs) to support complex fine-grained algorithms, thus optimizing bandwidth usage without compromising sensing detail. Crucially, the edge node not only handles computationally intensive analyses but also exerts real-time dynamic control over radar sensing parameters, including beam steering configurations.

Realizing the potential of real-time, fine-grained mmWave sensing, particularly within edge-end collaborative systems like CoSense, necessitates overcoming three principal challenges. First, bridging the intrinsic trade-off between low-latency operation and sensing fidelity remains a core difficulty. Existing approaches are forced to compromise: either relying on point clouds that omit crucial signal details or resorting to raw data that introduces prohibitive system overhead and necessitates offline processing. Achieving both coarse-grained responsiveness and fine-grained analytical depth simultaneously remains an open problem. Second, managing the substantial data overhead associated with raw mmWave signals poses a significant bottleneck. The continuous transmission and processing of large data volumes strain both bandwidth and computational resources at the edge, severely limiting the feasibility of fine-grained sensing under real-time constraints. Third, ensuring signal quality and robustness in dynamic environments is essential yet challenging. Subtle sensing features are highly vulnerable to SNR degradation and motion-induced artifacts, such as misalignment and phase instability, issues that are further exacerbated by conventional non-adaptive sensing mechanisms (as illustrated in Figure 1(a)).

To address these challenges, we propose CoSense, an edge-end collaborative sensing framework that leverages complementary dual-stream data to bridge the fidelity-latency trade-off, thereby achieving real-time fine-grained sensing. CoSense first introduces a novel dual-stream data acquisition mechanism, enabling concurrent real-time access to both point clouds and information-rich raw data on current-generation commodity mmWave radar devices (e.g., TI xWR1843). Continuously transmitting entire high-volume raw data presents a major bottleneck for real-time operation, straining both the limited end-edge bandwidth (e.g., 78.4 MB/s of AWR1843ISK [29]) and the processing resources of the edge node. To alleviate this, we innovate in an adaptive data transmission strategy that leverages point cloud information to identify RoIs and selectively transmits only the pertinent raw data segments to the edge node, substantially reducing bandwidth usage and computational load. To maintain high signal quality and robustness against dynamic effects like target motion misalignment, CoSense employs a closed-loop beamforming mechanism (as illustrated in Figure 1(b)). Leveraging high-fidelity localization outputs, the edge node dynamically adjusts the radar's transmit/receive beam patterns via feedback control for continuous target tracking, thereby directly counteracting motion artifacts while enhancing signal fidelity through targeted energy focusing and spatial interference suppression. Finally, we implement a prototype of CoSense using a commodity mmWave radar and an intelligent edge platform, and conduct extensive experiments to evaluate its performance. In summary, our main contributions are:

- We introduce CoSense, a novel edge-end collaborative architecture, enabling concurrent dual-stream data acquisition (point cloud and raw data) and dynamic, real-time edge control over radar sensing parameters.
- We propose a firmware-level adaptive data transmission mechanism that reduces bandwidth and computational overhead by selectively forwarding task-relevant raw data segments identified through point cloud analysis.
- We design a closed-loop adaptive beamforming system that integrates high-fidelity localization (achieved through edge-fused, multi-perspective data) with real-time beam steering to enhance signal quality and mitigate motion-induced artifacts.
- We prototype and evaluate CoSense across diverse real-time, fine-grained sensing tasks, demonstrating its robustness, efficiency, and practical deployability.

The remainder of this paper is organized as follows. Section II discusses the motivation behind CoSense and examines the feasibility of applying beamforming to fine-grained sensing. Section III elaborates on the system design of CoSense. Section IV and Section V respectively explain CoSense's implementation and report the extensive evaluations on CoSense. Related works and discussions of CoSense's limitations and potentials are briefly captured in Section VII and Section VI, followed by the conclusion of our paper in Section VIII.

II. MOTIVATION AND FEASIBILITY ANALYSIS

This section establishes the motivation for the proposed CoSense system. We first analyze the limitations inherent in conventional mmWave sensing pipelines and their associated resource constraints. We then investigate the feasibility and benefits of leveraging beamforming to enhance signal quality for fine-grained sensing compared with omnidirectional approaches. Finally, we assess the viability and limitations of utilizing Doppler information derived from point clouds for real-time target tracking, establishing the need for the dynamic beam steering capabilities integrated within CoSense.

A. Limitations of Conventional mmWave Sensing Systems

Conventional mmWave sensing pipelines typically employ a sequence of operations on reflected FMCW signals, including Range Fast Fourier Transform (FFT), Doppler FFT, and Angle-of-Arrival (AoA) estimation (e.g., MUSIC [30]), to generate Range-Doppler-Angle (RDA) data cubes, from which 3D point clouds are commonly derived. While foundational, this pipeline exposes fundamental trade-offs when striving for the high resolution essential for fine-grained sensing, particularly under the real-time constraints of dynamic scenarios (e.g., target position change). However, pursuing higher resolution

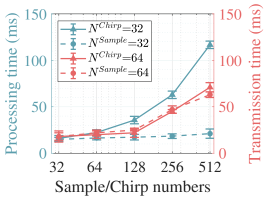
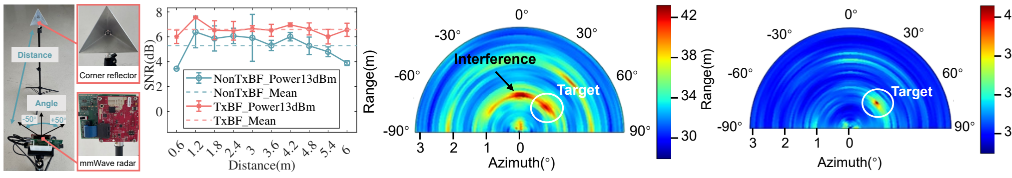


Fig. 2. Real-time analysis.



(a) Setup. (b) Impact of distance. (c) Conventional sensing method. (d) Beam-stitch sensing method.

Fig. 3. Comparison of SNR impact between conventional (omnidirectional) sensing and adaptive beamforming.

via these parameter adjustments within conventional architectures invariably encounters significant bottlenecks. This forces a compromise between two suboptimal operational modes, neither of which simultaneously satisfies the demands for low latency and high sensing fidelity: i) *On-device point cloud generation*: Generating point clouds directly on the radar chip minimizes data transmission to the edge node, seemingly favoring low-latency operation. However, configuring parameters to achieve the higher resolution often desired as a foundation for fine-grained sensing (e.g., increasing samples per chirp for better range detail) imposes a substantial computational load on the resource-constrained radar chip (e.g., DSP core). This increased processing burden for the on-chip pipeline can render real-time execution infeasible as resolution parameters scale, an effect illustrated in Figure 2 where processing time grows significantly with the number of samples. ii) *Raw data offloading*: Transmitting raw data to an external platform (edge) enables complex processing, offering the potential for high sensing fidelity. Yet, enhancing resolution by increasing parameters, such as the number of samples per chirp (for range) or the number of chirps per frame (for Doppler), results in a dramatic increase in the raw data volume that must be transmitted, as shown in Figure 2. This increase in data volume frequently saturates practical transmission links, such as universal asynchronous receiver-transmitter (UART) and low-voltage differential signaling (LVDS), creating a critical latency bottleneck that precludes real-time operation and hinders applications demanding high temporal fidelity.

Conventional systems face a trade-off: enhancing resolution either overwhelms on-chip processing limits or creates prohibitive data transmission bottlenecks. This renders the simultaneous attainment of high-fidelity sensing and low-latency responsiveness impractical; neither on-chip point-cloud generation nor complete raw-data offloading suffices. These constraints necessitate a flexible, collaborative architecture like CoSense, which strategically manages computation and data flow. Critically, our objective is not to compromise fidelity with 'loose' parameters, but rather to enable high-fidelity parameters in real-time by solving the very data overhead bottlenecks they create.

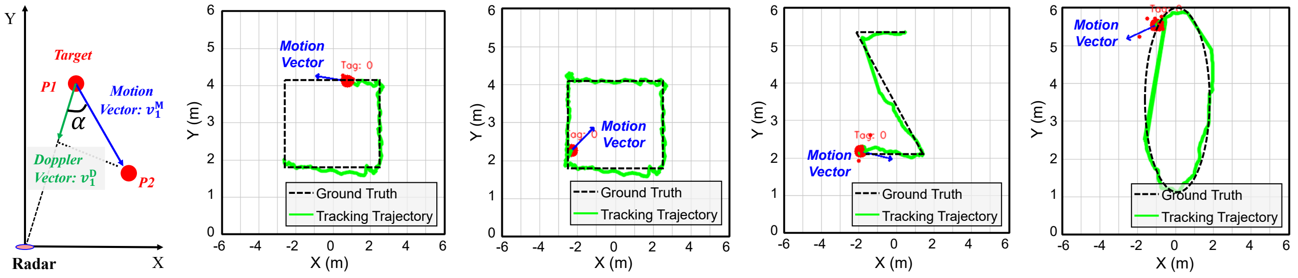
B. *Omnidirectional vs. Beam-Stitching Sensing*

Omnidirectional and multi-input multi-output (MIMO)-based sensing are widely adopted and have proven effective in dynamic scenarios, serving as a baseline in practical systems. Beyond the processing pipeline, the signal transmission strategy plays a critical role in sensing performance, especially for fine-grained tasks that are highly sensitive to signal quality. To this end, we compare conventional omnidirectional transmission with a directional beam-stitching approach, in which focused beams are sequentially transmitted over predefined spatial sectors and coherently combined to construct a spatial map. In contrast to omnidirectional sensing, which offers uniform angular coverage at the cost of reduced power density, beam-stitching concentrates energy toward specific directions. We evaluate both strategies using a TI AWR1843 mmWave radar (Figure 3(a)) and a corner reflector target (RCS $\approx 4\,\mathrm{m}^2$), with three transmit antennas operating at 13 dBm and target distances ranging from 0.6 m to 6 m.

As shown in Figure 3(b), beam-stitching consistently achieves a higher average SNR—approximately 1.3 dB above omnidirectional sensing, while also exhibiting significantly reduced SNR variability across distances, indicating more stable signal acquisition. This advantage is further illustrated by the range–angle heatmaps in Figures 3(c) and 3(d), where beam-stitching produces a clearer and stronger target response at a fixed location (1.5 m, 45°), demonstrating improved spatial focus and interference suppression. These results demonstrate that directional energy transmission (e.g., via beam-stitching) yields significantly higher and more stable SNR by concentrating energy spatially. This improved signal quality is particularly beneficial for fine-grained sensing tasks that are highly sensitive to SNR, motivating the use of dynamic, adaptive beamforming (as implemented in CoSense) as a complementary enhancement to conventional MIMO-based sensing. Under the evaluated configurations, this approach provides a measurable SNR gain of approximately 1.3 dB and significantly stabilizes signal quality under motion.

C. *Target Tracking Feasibility using Doppler Vectors*

Given the benefits of directional sensing, we next investigate the feasibility of using the Doppler vector, a standard output of mmWave processing, to guide dynamic beam steering for tracking moving targets. We experimentally evaluate the utility and inherent limitations of this approach. As illustrated in Figure 4(a), the Doppler vector ($\mathbf{v}^{\mathrm{D}}$), derived from the Doppler FFT, represents the target's velocity component along the radar's instantaneous line-of-sight (LoS). This differs from the target's true motion vector ($\mathbf{v}^{\mathrm{M}}$), which describes its velocity along the actual path of movement and is not directly measured by the radar. To assess this relationship empirically, a volunteer walks along predefined paths (Figure 4(b)) while ground truth position is recorded using an OptiTrack [31] motion capture system. The angle α between the LoS and the estimated instantaneous direction of motion (derived from historical



(a) Layout. (b) Snapshot. (c) Rectangular trajectory. (d) Z-line trajectory. (e) Elliptical trajectory.

Fig. 4. Evaluation of target tracking trajectory based solely on the Doppler vector. (a)Relationship between Doppler ($\mathbf{v}^{\mathrm{D}}$) and motion ($\mathbf{v}^{\mathrm{M}}$) vectors. (b-e) Estimated trajectory (green) vs. ground truth (dashed black) for: (b, c) Rectangular, (d) Z-line, and (e) Elliptical paths. Note tracking drift when target stops (d) and deviations on curved path (e), showing limitations of Doppler-only tracking.

ground truth positions) is used to approximate the motion vector from the measured Doppler vector via $\mathbf{v}^{\mathrm{M}} \approx \frac{\mathbf{v}^{\mathrm{D}}}{\cos(\alpha)}$.

The results highlight key limitations of relying solely on the Doppler vector for motion estimation. Rectangular path tracking (Figure 4(c)) shows better performance during straight segments but is still susceptible to errors during turns. As shown in the Z-line trajectory tracking results (Figure 4(d)), significant drift occurs in the estimated motion vector when the target stops, attributable to the poor reliability of Doppler measurements at near-zero velocities. Furthermore, during curved motion, such as the elliptical path (Figure 4(e)), the direction inferred from Doppler tends to lag or deviate from the true instantaneous direction. This arises because the $\cos(\alpha)$ estimation relies on historical position data, and the projection onto the rapidly changing LoS introduces inaccuracies. These observations demonstrate that while the Doppler vector provides useful radial velocity information, relying solely upon it for beam steering is insufficient for robust tracking, particularly during non-linear movements or stops. This underscores the need within the CoSense system for more sophisticated beamforming mechanisms to achieve reliable beam alignment.

III. SYSTEM DESIGN

This section presents the design details of the CoSense, a real-time sensing system built upon edge-end collaboration to enable fine-grained mmWave sensing in dynamic environments. As illustrated in the system overview (Figure 5), CoSense integrates three core modules specifically designed to overcome the limitations of conventional approaches and address the principal challenges outlined earlier:

- Dual-Stream Data Acquisition: The radar device (end) simultaneously generates both processed point clouds and raw data streams. These *complementary* streams are transmitted, temporally synchronized, to the edge node, providing inputs for both low-latency coarse-grained sensing (or control) and high-fidelity fine-grained analysis, directly addressing the fidelity-latency trade-off.
- Adaptive Data Transmission: Manages data overhead via a feedback-controlled adaptive transmission strategy; the edge determines precise target range bins and commands the device (end) via firmware to selectively transmit only corresponding raw data segments, drastically reducing bandwidth load and edge computation.

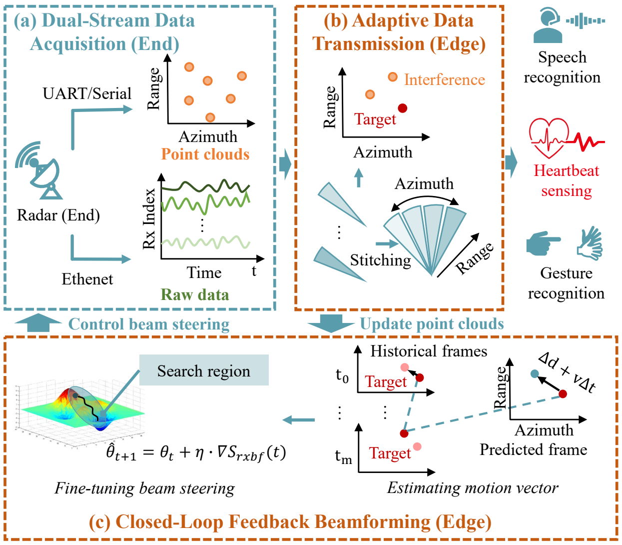


Fig. 5. CoSense system overview, illustrating core functional modules: (a) Dual-stream data acquisition providing point clouds and raw data; (b) Adaptive data transmission selectively forwarding pertinent raw data based on point-cloud RoIs to manage overhead; (c) Closed-loop adaptive beamforming utilizing localization and tracking for dynamic beam steering.

- Closed-Loop Feedback Beamforming: To ensure high signal quality and robustness to dynamic influences, CoSense establishes an adaptive beamforming sensing model and employs a closed-loop optimization strategy powered by edge intelligence, which integrates real-time motion prediction and signal quality (e.g., SNR) feedback to dynamically steer the radar beam, thereby maintaining continuous target focus and enhancing sensing fidelity.

A. Dual-Channel Data Acquisition

As discussed in Section II-A, conventional mmWave sensing approaches typically force a compromise. Systems relying solely on high-level point clouds from radar (End) enable real-time, coarse-grained sensing but discard crucial phase information required for fine-grained tasks (e.g., vital sign monitoring). Conversely, methods designed around information-rich raw data often operate offline, suitable for laboratory analysis but failing to meet the real-time processing demands of practical, dynamic deployments where access to such detailed data is also critical. To reconcile these conflicting requirements and bridge this gap, CoSense employs a dual-stream data acquisition approach, enabling the concurrent

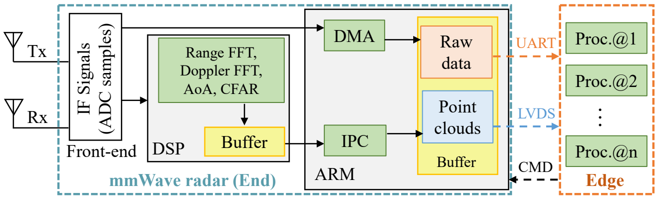


Fig. 6. CoSense end-device architecture for dual-stream data output and control. Point clouds generated by the DSP/ARM pipeline are sent via UART, while raw data captured concurrently via DMA is sent via LVDS. Control commands are received from the edge.

acquisition and utilization of both high-level point clouds and low-level information-rich raw data signals from the radar.

1) Dual-Stream Data Forwarding: The core implementation technique involves modifying the radar firmware to exploit direct memory access (DMA) [32] for intercepting raw analog-to-digital converter (ADC) samples prior to their entry into the digital signal processing (DSP) pipeline, as illustrated in Figure 6. This enables the efficient establishment of a parallel raw data stream that coexists with the standard point cloud processing pathway. Importantly, the DMA-based raw data acquisition operates independently and concurrently with the native point cloud generation process, ensuring that the latter remains unaffected. Leveraging DMA's capability for high-throughput data transfer with minimal ARM core involvement, this mechanism imposes negligible additional computational overhead on the radar processor. Each captured raw data frame is subsequently synchronized and associated with its corresponding point cloud frame, typically using embedded timestamps or frame numbers provided by the modified firmware. These paired data units are then enqueued into designated transmission buffers for forwarding. Finally, firmware enhancements allow the radar's heterogeneous communication interfaces, such as UART for low-volume point clouds and high-bandwidth channels like LVDS or Ethernet for raw data, to operate in parallel, ensuring the simultaneous and efficient transmission of both data streams to the edge.

2) Control Feedback Interface: A core advantage of CoSense is its efficient dual-stream data forwarding mechanism, which provides the edge node with concurrent access to both low-latency point clouds and information-rich raw data, enabling real-time, comprehensive analysis of the sensing environment and target states. However, passively receiving and analyzing data proves insufficient for dynamic, complex real-world scenarios, necessitating dynamic adjustment of the radar's physical parameters to achieve truly intelligent and adaptive sensing, such as precise target tracking or focusing energy on RoIs to improve fine-grained sensing SNR. To enable this capability, we modify the radar's low-level driver and firmware to establish an extensible control command interface. This interface allows the radar to receive runtime updates to key operational parameters, such as the beam steering angles required for adaptive beamforming.

During CoSense's operation, the edge node determines the optimal beam direction for the upcoming sensing frame based on the output of its target localization module (see Section III-B), and asynchronously transmits the corresponding steering command to the radar via the control interface.

Algorithm 1 Beam Steering Update Logic

```
Input: Chirp completion interrupt signal
Output: $phaseCodes\{tx_i\}$, SemaphoreCFGReady
Persistent State: $frameCounter$, $nextAngle$
procedure OnChirpInterrupt
    Increment $frameCounter$
    if IsEndOfFrame($frameCounter$) then
        Step 1: Compute phase shift codes.
        $(\mathrm{tx}_0, \mathrm{tx}_1, \mathrm{tx}_2) \leftarrow$ GenPhaseCodes($nextAngle$)
        Step 2: Build the configuration structure.
        $phaseCfg \leftarrow$ InitPhaseShiftCfg($\mathrm{tx}_0, \mathrm{tx}_1, \mathrm{tx}_2$)
        Step 3: Send configuration to the radar front end.
        ApplyTxPhaseShiftConfig($phaseCfg$)
        Step 4: Trigger notification to other task.
        SemaphoreCFGReady
        Step 5: Update the steering angle.
        $nextAngle \leftarrow$ GetNextScanAngle($nextAngle$)
        Step 6: Reset the frame counter.
        $frameCounter \leftarrow 0$
    end if
end procedure
```

Upon receiving the updated beam angle, the radar firmware preserves timing precision and uninterrupted data acquisition by deferring the parameter update: rather than modifying the current FMCW sequence immediately, it applies the update within the Chirp interrupt service routine, which is triggered after the final chirp of the current frame. As illustrated in Algorithm 1, this interrupt routine utilizes the received *nextAngle* to compute and configure the updated transmit *phaseCodes* (i.e., phase shift settings for beam steering), which are then staged or directly written to hardware registers to ensure that the new configuration takes effect at the start of the subsequent frame. This tightly integrated closed-loop feedback mechanism enables CoSense to bridge high-level edge intelligence (e.g., target prediction and beam direction optimization) with low-level radar parameter reconfiguration (e.g., real-time phase shift updates). Consequently, the radar can dynamically "focus" its sensing beam on critical targets or regions of interest, effectively compensating for target motion and substantially enhancing the robustness, SNR, and fidelity of fine-grained perception in complex, real-world environments.

B. Adaptive Data Transmission

While the dual-stream acquisition mechanism enables access to both point cloud and raw data, continuously transmitting the entire high-volume raw data stream is impractical due to bandwidth limitations and constraints on edge processing capacity (as discussed in Section II-A). To overcome these overheads, CoSense incorporates an adaptive data transmission strategy, which selectively forwards only the most relevant raw data segments to the edge node. The effectiveness of this strategy critically depends on the accurate and reliable identification of RoIs that encompass target locations. However, relying solely on coarse spatial information extracted directly from point clouds for RoI definition is often insufficient in

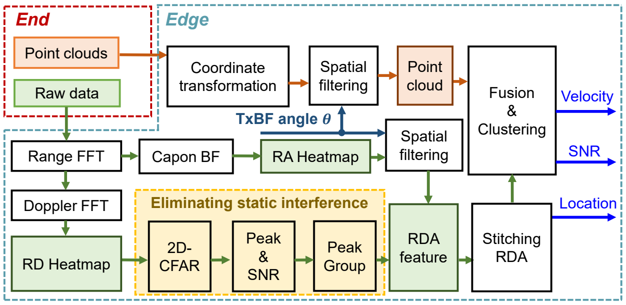
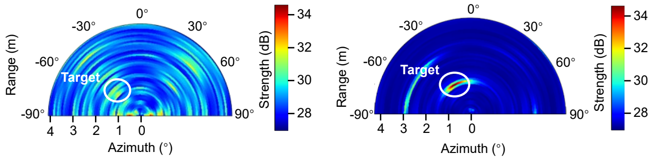


Fig. 7. Architecture of the two-phase edge-end collaboration, illustrating the cooperative processing flow between the end device (data acquisition and sensing) and the edge node (advanced signal processing and fusion).



(a) Original. (b) Beam-Stitching.

Fig. 8. Comparison of RA heatmap generation methods for positioning accuracy. (a) Conventional RA heatmap, and (b) RA heatmap enhanced with the proposed beam-stitching technique, illustrating improved spatial clarity and reduced ambiguity for more accurate localization.

practical scenarios. As discussed previously (see Section II-B), point cloud positions are prone to substantial errors caused by multipath propagation, interference, and ambient noise, particularly in complex indoor environments. An inaccurately defined RoI may result in the transmission of irrelevant raw data or, more detrimentally, the omission of crucial data associated with the actual target. To ensure robust and precise RoI identification necessary for effective selective transmission, CoSense employs a hybrid, two-stage target localization scheme prior to raw data filtering. As illustrated in Figure 7, this approach leverages both data streams provided by the dual-stream acquisition module: i) a *coarse estimation* phase that utilizes readily available, low-latency point cloud data to approximate potential target locations; and ii) a *fine-grained refinement* phase that applies beam-stitching techniques to the corresponding raw data segments, enabling high-fidelity localization via coherent data fusion at the edge.

1) Coarse Estimation: The first stage of the hybrid localization framework leverages the low-latency point cloud stream continuously received from the radar end device. Its objective is to rapidly generate initial, coarse estimates of potential target locations, thereby providing situational awareness and delineating RoIs for subsequent fine-grained refinement. Upon arrival at the edge node, the point cloud data typically undergoes preprocessing steps such as noise suppression and the removal of static clutter (e.g., reflections from stationary furniture in the environment). A clustering algorithm (e.g., DBSCAN [33]) is then applied to group points corresponding to distinct objects. For each significant cluster, a coarse position estimate (e.g., the centroid coordinates) is computed. Although these estimates may be affected by inaccuracies due to multipath effects, interference, and environmental noise (as discussed in Section II-B), they play a critical role in narrowing the search space and guiding the computationally intensive refinement process described in the following stage.

2) Fine-Grained Refinement: Upon receiving the synchronized data streams from the radar, the edge node executes parallel processing pipelines for fine-grained localization and state estimation, as depicted conceptually in Figure 7. *Range-Doppler (RD) Feature Extraction:* The raw ADC data first undergoes Range FFT and Doppler FFT operations to generate Range-Doppler (RD) heatmaps for each receiving channel. Clutter suppression techniques (e.g., 2D-CFAR) are then applied, followed by peak detection and clustering. This processing path primarily focuses on identifying dynamic targets, providing their range, velocity, and initial detection signatures. *High-Resolution RA Processing:* In parallel, for each transmitted beam angle θ_i, the corresponding raw data is processed through advanced AoA estimation algorithms (e.g., Capon beamforming [34]) following the Range FFT step. This yields a high-resolution RA heatmap per beam direction, capturing fine-grained spatial energy distribution within each sensing sector. *Beam Stitching and Spatial Filtering:* To reconstruct a holistic wide-angle spatial representation, the sequence of RA heatmaps across multiple beam directions (θ_i) is coherently combined through a beam stitching process. While this approach improves SNR compared with omnidirectional sensing (Figure 8(a)), it may introduce grating lobes and spatial ambiguities (due to the antenna spacing of commodity mmWave radar typically exceeding $\lambda/2$). To mitigate such artifacts, which could otherwise lead to "ghost" targets and degraded localization accuracy, spatial filtering techniques are applied to the stitched RA map, exploiting prior knowledge such as the beam steering sequence to suppress interference and produce a cleaner, high-resolution output (Figure 8(b)). *Data Fusion for Target State Estimation:* Simultaneously, the incoming coarse point clouds are further refined through coordinate transformation or outlier rejection. In the final fusion stage, the system integrates range-velocity information from the RD processing path with the precise spatial localization from the filtered RA map. This fusion yields accurate target state estimates, including position and velocity, with optional enhancement from refined point cloud inputs to boost robustness and tracking consistency.

To translate localization into efficient data transmission, CoSense employs a feedback mechanism where refined target location information is translated into relevant range bin indices. These indices are transmitted back to the radar end device in real-time via the control (command) interface (established as described in Section III-A). The radar firmware uses these indices to perform on-device data filtering, selectively extracting and transmitting only raw data segments (e.g., ADC samples) that correspond to target-relevant regions. This selective transmission drastically reduces bandwidth requirements. For instance, assuming a radar data structure of size $128 \times 128 \times 4$ (i.e, $N^{\text{Range}} \times N^{\text{Doppler}} \times N^{\text{Rx}}$), limiting transmission to 16 relevant range bins reduces the data volume to approximately one-eighth of the full stream. This adaptive strategy retains the most information-rich raw data while minimizing transmission overhead and computational load, thereby enhancing overall system efficiency and enabling the edge node to scale across multiple radar units.

C. Closed-Loop Feedback Beamforming

While the methods described in previous sections enable accurate target localization for fine-grained sensing, performing full-scale localization scans continuously for tracking moving targets is computationally expensive and hampers real-time responsiveness. To address this, CoSense adopts a more efficient strategy through adaptive beam steering, which dynamically adjusts the radar's transmit/receive beam direction to maintain focus on moving targets. Specifically, we propose a closed-loop adaptive beamforming mechanism designed to support efficient, robust target tracking and signal enhancement. We begin by introducing the beamforming-based sensing model and formalizing the relationship between the beam steering direction and its corresponding radar control parameters. Building upon motion prediction methods described in Section II-C, we then apply gradient descent optimization to refine the beam direction iteratively. This process continuously maximizes a signal quality metric (e.g., received SNR), thereby compensating for motion prediction errors and channel uncertainties, and enabling precise, real-time focus on the target.

1) Feedback Beamforming Model: Consider an mmWave radar system in which the transmitter employs a uniform linear array (ULA) with N elements, and the receiver is similarly equipped with a ULA with M elements. Although typical radar platforms (e.g., TI's AWR1843) feature 2D antenna layouts, their elevation aperture (e.g., two transmit antennas and one receive antenna) is severely limited by sparse vertical placement. The resulting broad vertical beamwidth renders dynamic elevation beam steering impractical for human-scale sensing, providing little to no functional gain. In this work, all experiments are conducted under near-planar deployment conditions, where the radar and targets are approximately at the same height and target motion is dominated by azimuthal (horizontal) displacement. Under these conditions, actively steering the beam in elevation is neither necessary nor effective. Therefore, we model both the transmit and receive arrays as 1D ULAs and perform beam steering exclusively in the azimuth dimension. This modeling choice accurately reflects the practical sensing geometry and hardware characteristics of our system.

Let $s^{\mathrm{Tx}} \in \mathbb{C}$ denote the complex baseband signal to be transmitted. On the transmit side, phase-controlled weights are applied to shape the beam. Let $\mathbf{u} = [u_0, u_1, \ldots, u_{N-1}]^{\mathrm{T}} \in \mathbb{C}^N$ denote the transmit beamforming vector, where u_i is the complex weight (typically a unit-magnitude phase shift, i.e., $|u_i| = 1$) applied to the i-th transmit element. The vector of signals transmitted by the array elements is thus: $\mathbf{s}^{\mathrm{Tx}} = s^{\mathrm{Tx}}\mathbf{u}$. After propagation through the environment and interaction with the targets, the signal is captured by the receive array. We represent the propagation channel between the N transmit and M receive elements with an $M \times N$ complex channel matrix $\mathbf{H}$, where the element $[\mathbf{H}]_{ki} = h_{ik}$ captures the complex gain (including path loss and phase shift) from transmit element i to receive element k. The target response is modeled as a scalar complex value h^{target}, encapsulating the radar cross-section and associated phase shift, as illustrated in Figure 9. The vector of signals received across the M elements, $\mathbf{s}^{\mathrm{Rx}} = [s_0^{\mathrm{Rx}}, \ldots, s_{M-1}^{\mathrm{Rx}}]^{\mathrm{T}} \in \mathbb{C}^M$, is then given by:

$$\mathbf{s}^{\mathrm{Rx}} = h^{\mathrm{target}}\mathbf{H}\mathbf{s}^{\mathrm{Tx}} = s^{\mathrm{Tx}}h^{\mathrm{target}}\mathbf{H}\mathbf{u}. \tag{1}$$

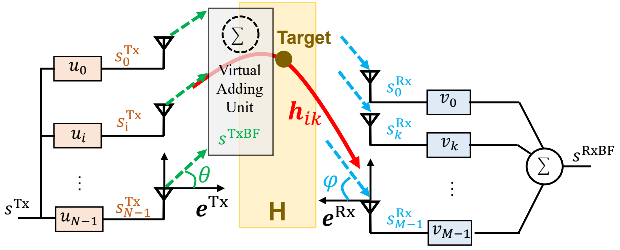


Fig. 9. Adaptive sensing model illustrating transmitter and receiver beamforming. The transmitter adjusts phases (u_i) to steer beams toward the target direction, while the receiver adjusts phases (v_k) to spatially aggregate signals, enabling accurate and dynamic sensing through Tx-Rx collaboration.

At the receiver side, beamforming coherently combines the signals received at different elements. Let $\mathbf{v} = [v_0, v_1, \ldots, v_{M-1}]^{\mathrm{T}} \in \mathbb{C}^M$ represent the receive beamforming vector, where v_k is the complex weight applied to the k-th received signal. The scalar output after receive beamforming, $\mathbf{s}^{\mathrm{RxBF}}$, is computed as the inner product:

$$s^{\mathrm{RxBF}} = \mathbf{v}^*\mathbf{s}^{\mathrm{Rx}}, \tag{2}$$

where $(\cdot)^*$ denotes the conjugate transpose. Substituting Eqn.(1) into Eqn.(2) yields the full beamformed output:

$$s^{\mathrm{RxBF}} = \mathbf{v}^*(s^{\mathrm{Tx}}h^{\mathrm{target}}\mathbf{H}\mathbf{u}) = s^{\mathrm{Tx}}h^{\mathrm{target}}\mathbf{v}^*\mathbf{H}\mathbf{u}. \tag{3}$$

This model explicitly illustrates that the beamformed received signal s^{RxBF} is jointly determined by the transmitted signal s^{Tx}, the target response h^{target}, the propagation channel matrix $\mathbf{H}$, and the selected transmit ($\mathbf{u}$) and receive ($\mathbf{v}$) beamforming vectors.

To focus the transmitted energy towards a desired direction (θ) and to selectively enhance received signal from a specific direction (ϕ), adaptive beamforming adjusts the complex weights in the transmit ($\mathbf{u}$) and receive ($\mathbf{v}$) beamforming vectors defined previously. For ULAs, these vectors are typically configured as steering vectors, aligning the array's main lobe with the intended directions of transmission and reception. Let d denote the inter-element spacing (assumed uniform and identical for Tx/Rx arrays), and λ the signal wavelength. The transmit steering vector $\mathbf{u}(\theta) \in \mathbb{C}^N$ that steers the beam toward angle θ is given by:

$$\mathbf{u}(\theta) = \frac{1}{\sqrt{N}}\left[1, e^{-j\frac{2\pi d}{\lambda}\sin(\theta)}, \ldots, e^{-j(N-1)\frac{2\pi d}{\lambda}\sin(\theta)}\right]^T, \tag{4}$$

where the normalization factor $1/\sqrt{N}$ ensures unit power, i.e., $||\mathbf{u}(\theta)||_2^2 = 1$. Similarly, the receive steering vector that directs the receiver's beam toward angle ϕ is denoted by $\mathbf{v}(\phi) \in \mathbb{C}^M$ and is constructed in the same manner. We note that this 1D steering approximation has a negligible impact on SNR and tracking accuracy in our experiments, as elevation variation remains within a small angular range ($< 5^\circ$), which is well below the elevation beamwidth supported by the hardware. Given the radar's limited elevation aperture and the near-planar deployment geometry, actively steering beams in elevation does not yield measurable performance gains. Tx–Rx collaboration in CoSense refers to closed-loop coordination between transmission and reception along the dominant motion dimension, rather than exhaustive physical beam steering in all spatial dimensions. Extending CoSense to elevation beam steering would be necessary for scenarios involving significant vertical motion or multi-floor environments, which we leave as future work.

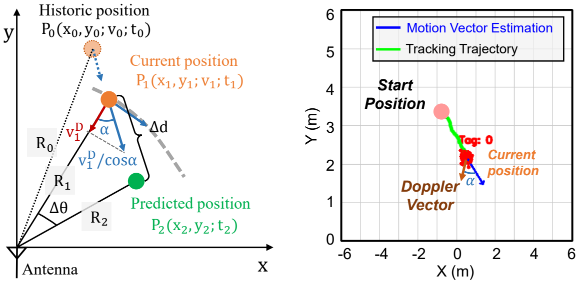
Substituting these angle-dependent steering vectors into our signal model (Eqn. (3)), the goal of beamforming is to find the optimal angles (θ^*, ϕ^*) that maximize the received signal power (or equivalently, SNR, assuming spatially uncorrelated noise). This involves maximizing the magnitude squared of the channel gain as seen through the beamforming vectors:

$$
\begin{aligned}
(\theta^*, \phi^*) &= \underset{\theta,\phi}{\operatorname{argmax}}\, |\mathbf{v}^*(\phi)\mathbf{H}\mathbf{u}(\theta) h^{\text{target}} s^{\text{Tx}}|^2, \\
\text{s.t.} &\begin{cases} |s^{\text{Tx}}|^2 \le P^{\text{Tx}} \\ -\pi/2 \le \theta, \phi \le \pi/2. \end{cases}
\end{aligned} \tag{5}
$$

Here, P^{Tx} denotes the maximum allowable transmit power of the baseband signal s^{Tx}. By iteratively adjusting the transmit angle θ and the receive angle ϕ to approach the optimal pair (θ^*, ϕ^*), the system (CoSense) can effectively concentrate energy along the desired path to the target. This beam alignment enhances the received signal quality, which is critical for high-resolution sensing and accurate target characterization in mmWave radar applications.

2) Motion Prediction and Beam Steering Refinement: The preceding signal model demonstrates that the received signal quality (s^{RxBF}) is maximized by jointly optimizing the transmit and receive steering vectors, $\mathbf{u}(\theta)$ and $\mathbf{v}(\phi)$, with respect to the channel matrix $\mathbf{H}$ and target characteristics h^{target}. Ideally, this corresponds to maximizing the effective channel gain $|\mathbf{v}^*(\phi)\mathbf{H}\mathbf{u}(\theta)|^2$, as formulated in Eqn. (5). However, explicitly solving this theoretical, $\mathbf{H}$-dependent optimization problem in real time for a moving target is impractical in dynamic sensing scenarios. Doing so would either require accurate and frequent estimation of the full channel matrix $\mathbf{H}$ or incur substantial computational overhead through exhaustive search procedures. Alternatively, repeatedly performing full localization scans to re-estimate the optimal beam direction would be prohibitively expensive in terms of latency and energy consumption. Furthermore, although the Cramér-Rao Bound (CRB) provides a useful theoretical benchmark, it is difficult to use for practical online optimization. Its reliance on specific channel models and perfect channel state information makes it impractical for highly dynamic, real-world settings.

To this end, CoSense adopts an adaptive beam-steering strategy within a closed-loop feedback framework to refine beam alignment after an initial target localization. Specifically, the system dynamically updates the transmit steering vector $\mathbf{u}(\theta)$ based on real-time feedback, enabling continuous tracking of the beam direction that maximizes the effective channel gain $|\mathbf{v}^*\mathbf{H}\mathbf{u}|^2$ without explicitly solving the underlying theoretical optimization problem. This closed-loop beam refinement mechanism improves received SNR, suppresses interference, and mitigates signal degradation caused by motion-induced beam misalignment, thereby enabling robust and efficient tracking in dynamic environments.



(a) Motion vector modeling.

(b) Estimation result.

Fig. 10. Motion vector estimation using the Doppler vector. (a) Geometric modeling illustrating key vectors and parameters. (b) Example showing the estimated motion vector output.

To support real-time beam steering, we estimate the required angular adjustment based on the predicted displacement of the target, inferred from Doppler velocity measurements. Given the target state at time t, including its position (x_t, y_t) and previous motion vector $\mathbf{v}_{t-1}^{\text{M}}$, along with the current measured radial Doppler velocity $\mathbf{v}_t^{\text{D}}$, we estimate .the predicted displacement Δd^{target} of the target over the upcoming frame interval ΔT^{f}, as illustrated in Figure 10(a):

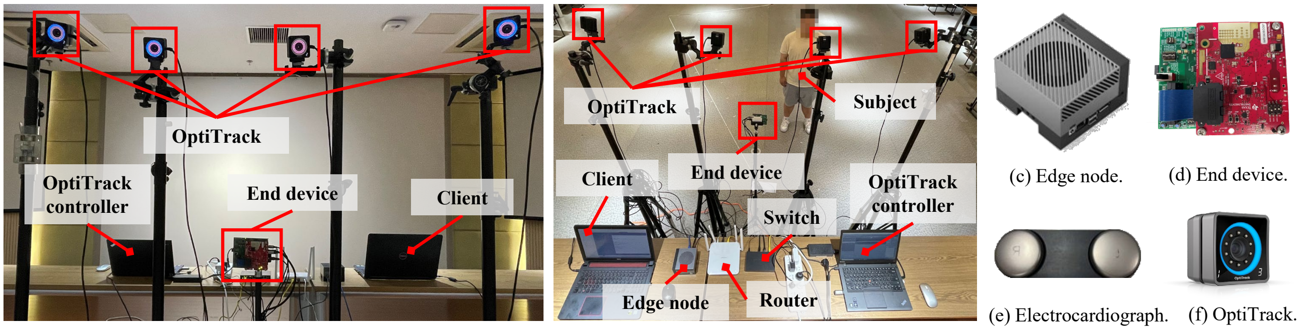
$$
\Delta d_t^{\text{target}} = \frac{|\mathbf{v}_t^{\text{D}}|}{\cos(\alpha)} \Delta T^{\text{f}}, \tag{6}
$$

where α represents the angle between the target's estimated true motion vector and the radial Doppler vector. We approximate this angle as $\alpha \approx \arccos(|\mathbf{v}_t^{\text{D}}|/|\mathbf{v}_{t-1}^{\text{M}}|)$ using the current Doppler speed and the motion speed derived from the previous state (see Figure 10(b)). Under the assumption of small angular deviations relative to the current range $R_t = \sqrt{(x_t^2 + y_t^2)}$, the predicted displacement Δd^{target} can be geometrically translated into a predicted angular offset $\Delta\theta_t$ required for beam redirection:

$$
\Delta\theta_t \approx \arcsin\left(\frac{\Delta d_t^{\text{target}}}{R_t}\right) = \arcsin\left(\frac{|\mathbf{v}_t^{\text{D}}|\Delta T^{\text{f}}}{\cos(\alpha) R_t}\right). \tag{7}
$$

This predicted offset $\Delta\theta_t$ provides a motion-informed estimate for proactively adjusting the beam direction in the next frame.

Relying solely on the predicted angular offset $\Delta\theta_t$ may result in suboptimal beam alignment due to estimation inaccuracies and complex propagation effects such as multipath. To address this, CoSense incorporates a closed-loop refinement mechanism that uses real-time SNR feedback (i.e., a practical, model-free objective) from the received signal to iteratively refine the transmit steering angle θ_t, compensating for prediction errors and channel dynamics. Instead of directly applying the potentially noisy $\Delta\theta_t$, the refinement process initiates from a predicted angle $\theta_{t+1} \leftarrow \theta_t + \Delta\theta_t$, or treats $\Delta\theta_t$ as a directional prior. It then probes a set of candidate angles $\hat{\theta}_{t+1} = \{\theta_{t+1}^0, \ldots, \theta_{t+1}^\omega, \theta_{t+1}^\Omega\}$ around this initial estimate. For each candidate, the system applies the corresponding steering vector $\mathbf{u}(\theta_{t+1}^\omega)$ (Eqn. (4)) and measures the resulting SNR of the received signal, s^{RxBF} (Eqn. (3)). This SNR feedback guides an iterative search toward the beam direction that yields the highest received signal quality. To accelerate



(a) OptiTrack motion capture system. (b) Experimental scenarios. (c) Edge node. (d) End device. (e) Electrocardiograph. (f) OptiTrack.

Fig. 11. Experiment setup: (a) OptiTrack motion capture system used for ground truth positioning and tracking. (b) Experimental scenarios for human positioning, tracking, and heartbeat sensing. (c) Edge node performs sensing strategy and controls end device beam steering. (d) End device transmits/receives signals and performs coarse-grained target positioning. (e) An ECG device used for ground truth heartbeat collection. (f) An OptiTrack camera.

convergence, techniques such as adaptive step sizing (η) or simple gradient-informed heuristics are employed. The process ultimately converges to an optimized beam steering angle θ^{opt} that maximizes real-time signal quality.

IV. PROTOTYPING & EXPERIMENT SETUP

In this section, we first elaborate on the implementation details of CoSense and then introduce the experiment setup used for its evaluation.

A. *Implementing CoSense*

CoSense is built using a commodity TI AWR1843 mmWave radar [29] as the end device, which transmits and receives FMCW signals. Raw baseband data is captured via TI's DCA1000EVM data acquisition board [35], while corresponding point cloud data is concurrently extracted through the radar's UART interface. An Nvidia Jetson Orin [36] embedded platform serves as the edge node, acquiring dual-stream data to enable collaborative signal processing. The overall hardware architecture is shown in Figure 11. Custom C/C++ driver code on the radar device configures the antenna array and accepts beam steering commands from the edge node over serial communication. Meanwhile, Python-based scripts on the edge (e.g., Jetson Orin) handle point cloud parsing, raw signal processing, and execution of real-time closed-loop beamforming algorithms for target tracking, beam steering refinement, and other applications (e.g., heartbeat monitoring). To meet real-time constraints, a multi-threaded processing pipeline is implemented on the edge node. It consists of: i) a data acquisition thread that receives radar data and transmits control commands; ii) a detection thread that performs two-stage positioning and beam parameter estimation; and iii) a fine-grained signal processing thread. Inter-thread communication is handled via shared queues to ensure synchronized data flow. A laptop client connects remotely to the edge node for monitoring and retrieving sensing outputs. Key radar signal configuration parameters are summarized in Table I.

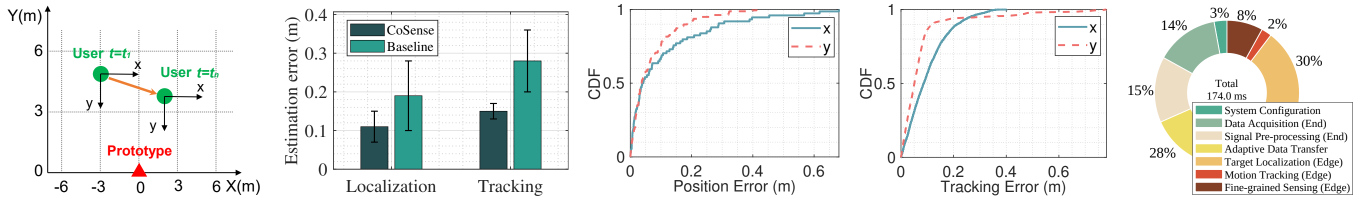
B. *Experiment Setup*

We conduct comprehensive experiments to rigorously evaluate the CoSense, encompassing both micro-benchmark tests for core functionalities and case studies (e.g., heartbeat monitoring) to assess practical applicability. Experiments are carried out in a typical indoor open-space environment measuring approximately 6 m × 6 m, furnished with standard office furniture. This setup provides a realistic level of static clutter and multipath propagation, representative of a typical office. The radar device is mounted on a tripod at a height of 1.2 m, as illustrated in Figure 11(b). The study involves 25 healthy volunteers (12 females and 13 males), aged between 20 and 40 years, with body mass indices (BMIs) ranging from 15 to 35 kg/m^2. Participants perform tasks under both static and dynamic conditions. In static trials, participants stand or sit motionless at predefined locations, with distances ranging from 1 m to 6 m and angles from $-45°$ to $+45°$, to evaluate positioning accuracy. In dynamic scenarios, participants follow predetermined walking paths, perform posture transitions (e.g., sit-to-stand), or engage in unstructured movement within a defined area to assess tracking robustness and the feasibility of heartbeat monitoring during motion. Each trial lasts approximately 60 to 120 seconds.

1) Ground Truth and Baseline: Ground truth for position and trajectory is obtained using an OptiTrack motion capture system (Figure 11(a) and (f)). Ground-truth heart rate (BPM) is measured using the ER2-ECG [37] chest strap device (Figure 11(e)). For baseline comparison, we implement a conventional on-device radar processing pipeline using the same AWR1843 hardware. The baseline operates with an identical radar configuration, frame rate, and time-division multiplexing MIMO (TDM-MIMO) scheme for spatial resolution, but does not employ closed-loop beamforming or

TABLE I
RADAR PARAMETER CONFIGURATION.

Parameter	Value
Start Frequency	77.5 GHz
ADC Sample Rate	7200 ksps
ADC Samples	128
Chirps per Frame	128
Range Resolution	0.08 m
Chirp Slope	100 MHz/μs
Ramp End Time	39 μs
Idle Time	150 μs
Doppler Sample Rate	2560 sps
Frame Periodicity	50 ms



(a) Layout. (b) Comparison with baseline. (c) Position error. (d) Tracking error. (e) Time consumption.

Fig. 12. Overall performance evaluation of CoSense. (a) Experiment setup and scene layout. (b) Comparison of positioning and tracking errors against the baseline. (c) Cumulative Distribution Function (CDF) of position estimation errors. (d) CDF of tracking estimation errors. (e) Per-module latency breakdown within the end–edge collaborative sensing pipeline.

adaptive data transmission. Its signal processing pipeline comprises standard FFT-based range-Doppler processing, CFAR-based target detection, angle-of-arrival estimation, and Kalman filter-based tracking. This design represents a widely adopted and well-established approach for dynamic radar sensing and tracking, and thus serves as a strong and representative baseline. Importantly, the baseline excludes the edge-node collaboration and closed-loop feedback mechanisms that are central to CoSense. Comparing CoSense against this TDM-MIMO baseline allows us to isolate and quantify performance gains, particularly in SNR robustness and interference resilience, directly attributable to our directional beam-steering and Tx–Rx collaborative design.

2) Performance Metrics: We employ the following evaluation metrics: Positioning/Tracking error: The Euclidean distance between the estimated and ground-truth positions in the 2D (x, y) plane, with additional analysis of deviation along each axis; Heart rate error: The absolute difference between the estimated heart rate (HR^{E}) and the ECG ground truth (HR^{G}), computed as $\mathrm{HR}^{\mathrm{Err}} = |\mathrm{HR}^{\mathrm{G}} - \mathrm{HR}^{\mathrm{E}}|$. This metric assesses the fine-grained sensing accuracy of CoSense, which is particularly challenging due to the minute (0.1–0.5 mm) chest wall displacements associated with cardiac activity. All experimental procedures were approved by the Institutional Review Board (IRB), and written informed consent was obtained from all participants.

V. EVALUATION

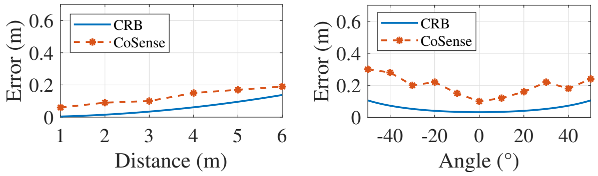
In this section, we conduct experiments to evaluate the performance of CoSense in both static and dynamic scenarios.

A. Micro-benchmark Studies

We conduct micro-benchmark evaluations to assess the real-time sensing performance of CoSense in both static and dynamic scenarios. The system exhibits robust target localization and tracking capabilities, achieving an average Euclidean positioning error of 0.11 m in static conditions and an average tracking error of 0.15 m under diverse dynamic movements. The experiment setup, including the spatial arrangement of participants relative to the prototype, is illustrated in Figure 12(a). As shown in Figure 12(b), CoSense significantly outperforms the baseline system, reducing the median positioning error by approximately 42% and the median tracking error by approximately 46%. These improvements underscore the effectiveness of CoSense 's edge-end collaborative architecture and its closed-loop feedback beamforming strategy in enhancing real-time performance.

1) Localization Performance: An analysis of static positioning accuracy, presented as a cumulative distribution function (CDF) in Figure 12(c), demonstrates the high accuracy of CoSense 's localization capability. A breakdown of error components along individual axes reveals that the median localization error is approximately 0.05 m on both the X-axis (tangential direction) and the Y-axis (radial direction). Notably, 90% of Y-axis errors fall below 0.17 m, while 90% of X-axis errors are under approximately 0.28 m. The considerably lower error along the Y-axis is primarily due to the intrinsic characteristics of FMCW radar systems, which offer superior resolution in range estimation (aligned with the Y-axis for targets near the central field of view) compared with angle estimation (X-axis accuracy, especially at wider angles).

2) CRB-based Optimality Analysis: To evaluate the "near-optimality" of CoSense, we analyze its localization performance relative to the CRB under the standard FMCW radar configuration (Table I). The CRB is derived under the assumption of unbiased estimation, additive white Gaussian noise, and ideal synchronization, yielding a fundamental lower bound on localization error; the full derivation is provided in Appendix A. Figure 13 compares CoSense's empirical localization error with the corresponding CRB over varying target distances and azimuth angles. Across both dimensions, CoSense closely tracks the CRB, maintaining a small and consistent performance gap throughout the operational range. The residual gap is primarily attributable to practical factors not captured by the idealized CRB model, including hardware non-idealities and multipath propagation. This close alignment indicates that CoSense operates near the theoretical performance limit, providing quantitative evidence of its "near-optimality" under realistic conditions. Although the CRB is derived under an idealized model with a known signal structure, the observed agreement suggests that maximizing instantaneous received SNR via closed-loop adaptation effectively drives the system toward the same fundamental limit, even without explicitly optimizing a model-based CRB objective.



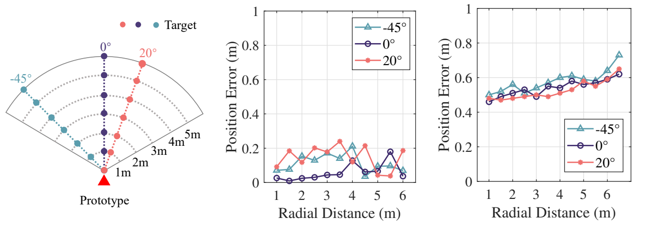
(a) Target distances. (b) Azimuth angles.

Fig. 13. Optimality analysis of CoSense via CRB benchmarking: (a) localization error across varying distances and (b) localization error across angles.

3) Tracking Performance: The dynamic tracking capabilities of CoSense are evaluated across a range of motion scenarios, with results illustrated in the CDF plot in Figure 12(d). The system consistently delivers reliable tracking performance, with axis-specific error trends mirroring those observed in the static case. The median Y-axis tracking error is approximately 0.05 m, while the median X-axis error is slightly higher at 0.08 m. At the 90th percentile, Y-axis errors remain within 0.13 m, and X-axis errors within 0.22 m. The enhanced tracking fidelity along the Y-axis stems from accurate radial velocity estimation via the Doppler effect. In contrast, tangential tracking accuracy along the X-axis is comparatively lower, as it depends on the temporal stability of AoA estimations, the predictive strength of the extended Kalman filter (EKF), and the real-time adaptability of the beam steering mechanism.

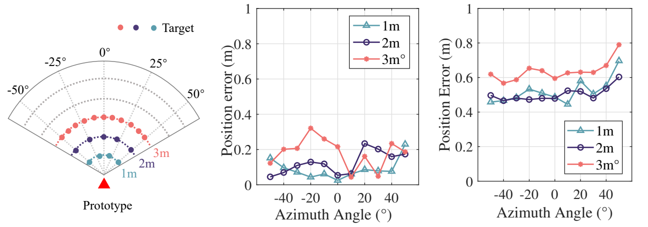
4) Time Consumption Analysis: We evaluate the real-time performance of CoSense by breaking down its end-to-end processing pipeline, as illustrated in Figure 12(e). Under the standard radar configuration (see Table I) with a 50 ms frame period, the cumulative processing latency for a single frame is approximately 174 ms. Despite this per-frame latency, CoSense sustains real-time operation through a pipelined edge–end collaborative architecture, where sensing, processing, and data transmission proceed concurrently across successive frames. In particular, the two most time-consuming stages, adaptive data transfer for end-to-edge (≈50 ms) and target localization at the edge node (≈52 ms), are carefully optimized to match the radar frame period. This alignment enables efficient pipeline parallelism: while the edge node performs fine-grained analysis on frame N, the end device simultaneously acquires and pre-processes frame $N+1$. As a result, CoSense maintains a stable throughput of 20 FPS, with the pipeline masking cumulative latency and preventing frame backlogs or drops. Although the system supports per-frame localization, practical deployments typically adopt a hybrid design that combines an initial localization step with subsequent multi-frame tracking to improve robustness and computational efficiency. The latency breakdown in Figure 12(e) reflects the core sensing, localization, tracking, and data forwarding modules shared across applications. Application-specific processing (e.g., heartbeat estimation) introduces only marginal overhead and does not affect the system-level execution pipeline. Overall, while the cumulative per-frame latency exceeds the radar frame period, system throughput is fundamentally determined by the frame rate rather than by the sum of individual module latencies. The pipelined execution model ensures that end-to-end latency does not translate into throughput degradation.

5) Impact of Distance and Azimuth Angle on Localization: We further examine how target distance and azimuth angle affect CoSense's localization accuracy under two controlled experimental settings. **Impact of distance:** Volunteers are positioned at radial distances from 1 m to 6 m along three fixed azimuth angles: −45°, 0° (boresight), and +20°. The experiment setup and results are shown in Figure 14. CoSense achieves a consistently low average localization error of 0.11 m under static conditions. As illustrated in Figure 14(b), accuracy is most stable along the boresight, where errors remain below 0.1 m across most tested distances. In contrast, the baseline (Figure 14(c)) exhibits substantially higher, more variable errors that increase rapidly with distance. While CoSense preserves localization stability at longer ranges via closed-loop beamforming, the baseline suffers pronounced performance degradation, particularly at wider angles. This degradation primarily stems from reduced array gain away from the boresight and compounded path loss at increased distances. **Impact of angle:** We next evaluate angular robustness by sweeping the azimuth angle from −50° to +50° at fixed distances of 1 m, 2 m, and 3 m. As shown in Figure 15(b), CoSense maintains a low average error of 0.13 m across this wide angular span. Consistent with radar theory, localization precision peaks near the boresight and gradually degrades toward the edges of the field of view. In contrast, the baseline (Figure 15(c)) is highly sensitive to wide-angle and long-range conditions (e.g., 3 m at ±50°), where localization accuracy deteriorates significantly. Overall, these results demonstrate that although CoSense achieves its highest accuracy in front-facing configurations, it maintains reliable, stable localization across a broad operational envelope. By effectively mitigating SNR loss and beam misalignment, CoSense demonstrates strong practicality for real-world deployments with diverse spatial geometries.



(a) Layout. (b) CoSense. (c) Baseline.

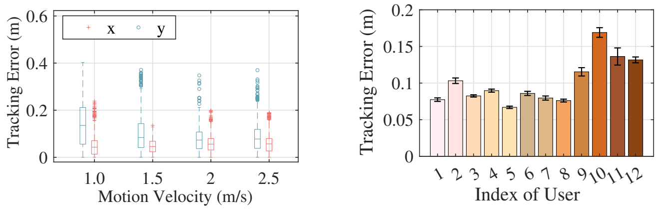
Fig. 14. Comparative analysis of localization error across varying distances. (a) Experiment layout, (b) localization error of CoSense, and (c) localization error of baseline.



(a) Layout. (b) CoSense. (c) Baseline.

Fig. 15. Comparative analysis of localization error across varying angles. (a) Experiment layout, (b) localization error of CoSense, and (c) localization error of baseline.

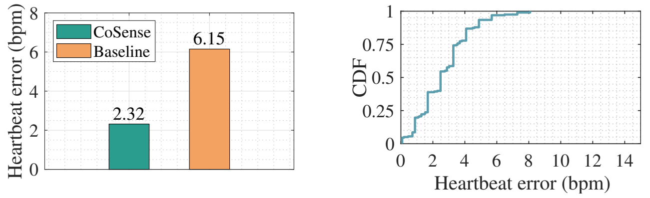
6) Impact of Velocity and User on Tracking: To further CoSense's tracking robustness, we evaluate its sensitivity to target velocity and individual user characteristics. For velocity analysis, volunteers walked randomly (Figure 12(a)) at controlled speeds of 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s. As presented in Figure 16(a), tracking accuracy remains stable across this range; the median error and interquartile range show minimal variation, and the average tracking error remains below 0.13 m or all speeds. This confirms CoSense's high-fidelity tracking for typical human motion, including walking and light jogging. To evaluate user variability, we conduct a



(a) Impact of velocity.

(b) Impact of individuals.

Fig. 16. Impact of velocity (a) and user variability (b) on tracking error.



(a) Comparison with baseline.

(b) Heartbeat error.

Fig. 17. Heartbeat monitoring performance: (a) Comparison of average error between CoSense and baseline. (b) CDF of heartbeat error for CoSense.

separate experiment involving 12 volunteers (6 males and 6 females), each instructed to walk at a target speed of 1.5 m/s within a 4 × 4 m area in front of the radar. Figure 16(b) illustrates the average tracking error for each participant. The results demonstrate consistent tracking performance across users, with an overall average error of just 0.098 m. Most users exhibit average errors between 0.07 m and 0.13 m, although slight deviations are observed, such as user 10, who shows a higher error of approximately 0.17 m, potentially due to variations in gait or deviations from the target walking speed. These results underscore CoSense's robustness across a range of user dynamics, validating its applicability in real-world environments with diverse and fast-moving subjects.

7) Ablation Study of CoSense Components: To validate the contributions of CoSense's key modules, we conduct an ablation study with results presented in Table II. Disabling closed-loop beamforming (i.e., w/o Beamforming) substantially degrades sensing fidelity: the mean tracking error increases from 0.15 m to 0.28 m, corresponding to an 87% degradation. In contrast, adaptive data transmission (i.e., w/o Adaptive Transmission) primarily governs system efficiency. Removing this module increases the required data rate by 8.1×, from 9.7 MB/s to 78.4 MB/s, confirming its critical role in alleviating the data bottleneck and enabling real-time operation. We further validate the dual-stream design by replacing the lightweight point-cloud stream with a conventional FFT-based coarse sensing pipeline (i.e., Replace Point Cloud). This substitution proves highly inefficient, simultaneously increasing the tracking error to 0.26 m and inflating the bandwidth demand to 58.2 MB/s. Overall, these results demonstrate that closed-loop beamforming is indispensable for sensing fidelity, while adaptive transmission is essential for efficiency; their joint integration underpins CoSense 's performance.

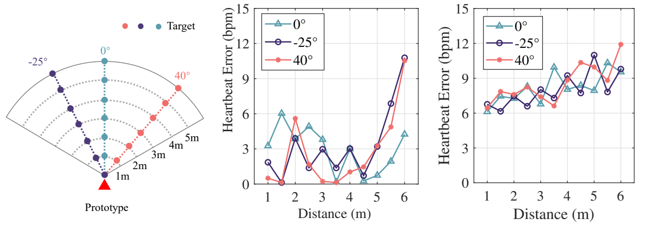
B. Case Study: Heartbeat Monitoring

This case study examines CoSense's ability to perform fine-grained physiological sensing by monitoring human heartbeats, leveraging its high-precision localization and tracking capabilities. Figure 17(a) presents a direct comparison of average heartbeat estimation errors between CoSense and the conventional baseline system (defined in Section IV-B-1). CoSense achieves a significantly lower average error of 2.32 bpm, whereas the baseline exhibits a much higher error of 6.15 bpm, underscoring the advantages of CoSense 's advanced signal processing pipeline for extracting subtle physiological signatures. The CDF of heartbeat errors across all participants and conditions is shown in Figure 17(b). The median error is approximately 2 bpm, indicating that 50% of all measurements fall within this error bound. Furthermore, 75% of the results lie within a 4 bpm error margin, reflecting the system's strong overall accuracy. Notably, while heartbeat monitoring under static conditions yields the lowest error (average ≈1.8 bpm), CoSense maintains reliable performance during movement, achieving an average dynamic error of approximately 3.36 bpm. This resilience in dynamic scenarios is critical for continuous health monitoring in the real world.

TABLE II
ABLATION STUDY OF COSENSE COMPONENTS

Configuration	Mean Position Error (m)	Tracking Error (m)	Bandwidth (MB/s)
Full CoSense	0.11	0.15	9.7
w/o Beamforming	0.18	0.28	9.6
w/o Adaptive Transmission	0.12	0.17	78.4
Replace Point Cloud with FFT-based Coarse Sensing	0.17	0.26	58.2

1) Impact of Distance on Heartbeat Monitoring: We evaluate the effective operational range of fine-grained physiological sensing by examining heartbeat monitoring accuracy as a function of radial distance. In this experiment, 11 participants are positioned at distances from 1 m to 6 m along three azimuth angles ($-25°$, $0°$, and $40°$), as shown in Figure 18(a). As illustrated in Figure 18(b), CoSense maintains high accuracy within a 5 m radius, with heartbeat estimation errors typically below 6 bpm across all tested angles. Although coarse localization and tracking remain reliable beyond this range (e.g., localization error < 0.2 m), heartbeat accuracy degrades sharply at 6 m, where errors exceed 9 bpm for all azimuths. This abrupt performance drop reflects the stringent SNR requirements of physiological sensing: heartbeat-induced micro-motions generate extremely weak phase perturbations whose effective power decays rapidly with distance (proportional to R^4), eventually

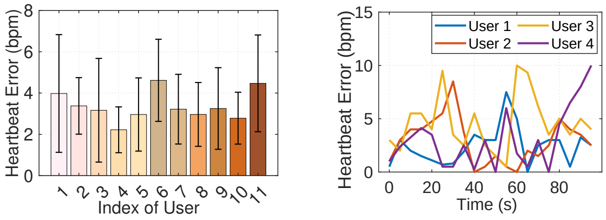


(a) Layout.

(b) CoSense.

(c) Baseline.

Fig. 18. Comparative analysis of heartbeat error across varying distances. (a) Experiment layout illustrating user locations, (b) heartbeat error of CoSense, and (c) heartbeat error of baseline.



(a) Impact of user. (b) Impact of duration time.

Fig. 19. Impact of user and duration time on heartbeat monitoring.

falling below the detection threshold beyond 5 m. In contrast, the baseline system (Figure 18(c)) incurs substantially larger errors throughout the entire operating range. Even at short distances (1–3 m), baseline errors frequently fluctuate between 6 bpm and 9 bpm, lacking the stability needed for reliable heartbeat monitoring. This limitation arises from the baseline's omnidirectional transmission, which fails to provide sufficient array gain to resolve subtle chest wall displacements on the order of 0.1–0.5 mm. By comparison, CoSense employs closed-loop beamforming to sustain a high-gain, target-focused link, whereas the baseline's dispersed energy is far more vulnerable to noise and multipath interference.

2) Impact of User Variability and Monitoring Duration on Heartbeat Accuracy: To further evaluate CoSense 's generalization capability, we analyze performance across users and examine system stability over time. First, we investigate inter-user variability. Figure 19(a) presents the average heartbeat estimation error for each of the participants. Despite minor differences due to individual physiological factors or movement artifacts, the overall average error is 3.21 bpm, demonstrating CoSense's consistent and reliable performance across a diverse user set. Furthermore, we assessed the system's stability during continuous monitoring. The results, illustrated in Figure 19(b), show that while instantaneous heartbeat error naturally fluctuates during the monitoring period, the average error remains stable, calculated at 3.24 bpm over the 90 seconds. The average variance during this period is 5.59 bpm, quantifying the degree of fluctuation around the mean. These findings suggest that while short-term readings exhibit some variability, averaging over a sufficient duration provides a reliable heart rate estimate. This highlights the importance of considering monitoring duration when evaluating the performance of heartbeat detection systems, as it influences noise averaging and overall estimate stability.

VI. DISCUSSION

This section elucidates the practical implications, inherent limitations, and prospective future directions stemming from the design and evaluation of CoSense. Our evaluations, conducted in a typical indoor office with furniture and including dynamic, unstructured user motion, demonstrate CoSense's robustness beyond simple static or controlled settings. Despite these promising results, several limitations merit consideration.

a) Limitations: Despite the promising results demonstrated by CoSense, several limitations merit consideration. First, the current evaluation is limited to controlled indoor laboratory environments. As such, the system's performance in more complex real-world scenarios, such as larger indoor spaces with diverse architectural layouts, increased clutter, heterogeneous materials, or outdoor environments, remains unexplored. These settings may introduce stronger multipath propagation and dynamic background interference, potentially affecting the reliability of tracking and sensing. In addition, our current implementation adopts azimuth-only (1D) beam steering, which is well-suited for near-planar sensing scenarios but does not explicitly handle targets with significant elevation changes. Applications involving pronounced vertical motion or multi-floor environments would require extending the framework to support full 2D beamforming and elevation-aware tracking. While our design, which relies on high-resolution processing of raw data rather than just point clouds, is intended to be more robust to such effects, a rigorous evaluation in cluttered environments is a critical part of future work. Second, while CoSense includes mechanisms for target detection, the present study focuses on single-target tracking and physiological monitoring. Extending the system to support robust multi-target tracking, particularly when targets interact or occlude one another, poses challenges related to data association, real-time interference suppression, and adaptive beam allocation. Addressing these challenges is critical for deploying the system in crowded, dynamic environments involving multiple users or agents.

b) Future work: Several avenues exist to advance the CoSense framework and broaden its application scope. Now that our evaluation has validated the core platform's ability to robustly improve accuracy while preserving low latency, a key direction involves extending CoSense to additional fine-grained sensing modalities, such as gesture recognition, gait characterization, and potentially even non-contact speech sensing. These applications could benefit from the system's ability to capture subtle motion signatures. Moreover, exploring target orientation or rotational motion estimation represents an exciting trajectory. This may be achieved by simultaneously tracking multiple anatomical landmarks or by incorporating advanced features such as polarimetric sensing, assuming hardware support. Further, integrating intelligent beam control mechanisms, such as dynamic or context-aware beam steering and allocation, could enhance the system's responsiveness and efficiency in complex environments. Finally, future work may explore the development of collaborative distributed sensing systems, wherein multiple CoSense-enabled nodes share contextual and spatial information. Such cooperation could facilitate synchronized multi-point tracking and improve robustness in larger-scale deployments.

VII. RELATED WORK

In this section, we review related work on mmWave sensing, focusing on two primary directions: advanced sensing methods and beamforming technologies.

a) mmWave Sensing Methods: Current mmWave sensing approaches can be broadly categorized into two types based on their data processing paradigm: *offline* and *online* sensing. *Offline* sensing uses mmWave radar to collect data, which is later processed with sophisticated methods to extract features and fine-grained sensing tasks, such as vital signs monitoring [19]–[22], [27], [38], voice sensing [14], [17], [18], and facial recognition [28], [39], [40]. These methods often rely on powerful external computing platforms to extract nuanced features, thereby achieving high sensing accuracy. However, in dynamic real-world scenarios where the target moves [19], [39], maintaining continuous phase coherence becomes challenging, which is critical for extracting fine-grained motion features. Moreover, the high processing latency of complex algorithms makes them unsuitable for real-time applications like continuous vital sign monitoring [21] or real-time speech/motion recognition [18], [39]. By contrast, *online* sensing acquires real-time data and transmits extracted features or point cloud data to edge devices or clients for immediate processing. This strategy supports coarse-grained applications like gesture [11], [12], [23] and human action recognition [41]. However, many of these systems still rely primarily on point-cloud-level data [42] or heavily downsampled [43] representations to meet real-time constraints, thereby sacrificing the rich, fine-grained information available in the raw data. To address these limitations, an emerging direction is adaptive configuration radar sensing [44], [45], where the system dynamically tunes radar or processing parameters (e.g., frame rate, resolution) to match the application's needs or environmental conditions [1], [46]. While these systems adapt sensing or processing parameters, they still face the fundamental limitations of commercial devices: limited on-chip computing capabilities and constrained data transmission bandwidth. They typically fail to address the underlying raw-data bottleneck because they lack mechanisms for concurrent dual-stream acquisition (i.e., low-latency point clouds and high-fidelity raw data) or for firmware-level adaptive filtering before transmission.

b) Beamforming Technology: Beamforming plays a critical role in mmWave sensing and communication systems, with key strategies including analog beamforming [47]–[49], digital beamforming [50], [51], and hybrid beamforming [46], [52]–[54]. Analog beamforming is widely deployed in practical systems due to its simplicity and power efficiency. It has been proven effective for enhancing signal coverage and quality [55], improving spectral efficiency [47], and increasing throughput and link reliability [49]. Digital beamforming, by contrast, enables more flexible control over individual antennas and supports advanced features such as massive MIMO [56] and adaptive beam steering [50], [57]. Despite its advantages in energy and spectrum efficiency, its high cost, increased complexity, and sensitivity to hardware impairments limit its adoption in lightweight or resource-constrained systems. Hybrid beamforming attempts to strike a balance between analog and digital approaches, providing a cost-effective solution for mmWave communication and sensing. It has been applied to optimize communication [53], find optimal sensing directions [46], and enhance system performance [52]. Trebsen [46] introduces a transmitter-receiver collaborative beamforming technique for fine-grained sensing. However, it primarily focuses on static users and lacks adaptability to dynamic scenarios where targets may move across an extensive spatial range.

Unlike previous methods that rely on offline processing or lack adaptability to dynamic conditions, CoSense offers a real-time end-edge collaborative sensing system integrating embedded beam control and adaptive signal processing.

VIII. CONCLUSION

This paper presented CoSense, an edge-end collaborative system enabling real-time, fine-grained mmWave sensing in dynamic environments by bridging the conventional fidelity-latency gap. CoSense synergistically combines dual-stream data acquisition, adaptive data transmission informed by hybrid localization, and closed-loop adaptive beamforming to manage overhead and enhance signal quality against dynamic effects. Experimental evaluations validate the effectiveness of CoSense, achieving a tracking latency of less than 50 ms. Under dynamic conditions, it attains an average positioning error of 0.11 m, a tracking error of 0.15 m, and a heartbeat monitoring error of 3.36 bpm, demonstrating its potential as a robust platform for adaptive sensing in real-world scenarios.

REFERENCES

[1] Z. Gao, Z. Wan, D. Zheng, S. Tan, C. Masouros, D. W. K. Ng, and S. Chen, "Integrated Sensing and Communication with mmWave Massive MIMO: A Compressed Sampling Perspective," *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, pp. 1745–1762, 2022.

[2] J. Zhang, R. Xi, Y. He, Y. Sun, X. Guo, W. Wang, X. Na, Y. Liu, Z. Shi, and T. Gu, "A Survey of mmWave-based Human Sensing: Technology, Platforms and Applications," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2052–2087, 2023.

[3] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter Wave Communication: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1616–1653, 2018.

[4] D. Zhang, X. Zhang, Y. Xie, F. Zhang, X. Wang, Y. Li, and D. Zhang, "LoCal: An Automatic Location Attribute Calibration Approach for Large-Scale Deployment of mmWave-based Sensing Systems," *Proc. of the ACM IMWUT*, vol. 7, no. 4, 2024.

[5] T. Wei and X. Zhang, "mTrack: High-Precision Passive Tracking Using Millimeter Wave Radios," in *Proc. of the 21st ACM MobiCom*, 2015, pp. 117–129.

[6] C. Wu, F. Zhang, B. Wang, and K. J. Ray Liu, "mmTrack: Passive Multi-Person Localization Using Commodity Millimeter Wave Radio," in *Proc. of the 39th IEEE INFOCOM*, 2020, pp. 2400–2409.

[7] H. Liu, X. Liu, X. Xie, X. Tong, and K. Li, "PmTrack: Enabling Personalized mmWave-based Human Tracking," *Proc. of the ACM IMWUT*, vol. 7, no. 4, 2024.

[8] Z. Chen, G. Zhu, S. Wang, Y. Xu, J. Xiong, J. Zhao, J. Luo, and X. Wang, "M^3: Multipath Assisted Wi-Fi Localization with a Single Access Point," *IEEE Transactions on Mobile Computing*, vol. 20, no. 2, pp. 588–602, 2019.

[9] E. Soltanaghaei, A. Prabhakara, A. Balanuta, M. Anderson, J. M. Rabaey, S. Kumar, and A. Rowe, "Millimetro: mmWave Retro-Reflective Tags for Accurate, Long Range Localization," in *Proc. of the 27th ACM MobiCom*, 2021, pp. 69–82.

[10] S. Ma, L. Xie, C. Wang, Y. Bu, L. Fan, J. Ning, Q. Guo, B. Ye, and S. Lu, "Multi-modal based 3d localization via the channel adjustment led-tag," *IEEE Transactions on Mobile Computing*, 2025.

[11] H. Liu, Y. Wang, A. Zhou, H. He, W. Wang, K. Wang, P. Pan, Y. Lu, L. Liu, and H. Ma, "Real-time Arm Gesture Recognition in Smart Home Scenarios via Millimeter Wave Sensing," *Proc. of the ACM IMWUT*, vol. 4, no. 4, 2020.

[12] D. Salami, R. Hasibi, S. Palipana, P. Popovski, T. Michoel, and S. Sigg, "Tesla-Rapture: A Lightweight Gesture Recognition System From mmWave Radar Sparse Point Clouds," *IEEE Transactions on Mobile Computing*, vol. 22, no. 8, pp. 4946–4960, 2023.

[13] C. Jin, X. Meng, X. Li, J. Wang, M. Pan, and Y. Fang, "Rodar: Robust Gesture Recognition Based on mmWave Radar Under Human Activity Interference," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 11 735–11 749, 2024.

[14] T. Liu, M. Gao, F. Lin, C. Wang, Z. Ba, J. Han, W. Xu, and K. Ren, "Wavoice: A Noise-resistant Multi-modal Speech Recognition System Fusing mmWave and Audio Signals," in *Proc. of the 19th ACM SenSys*, 2021, p. 97–110.

[15] L. Fan, L. Xie, X. Lu, Y. Li, C. Wang, and S. Lu, "mmMIC: Multi-modal Speech Recognition based on mmWave Radar," in *Proc. of the 42nd IEEE INFOCOM*, 2023, pp. 1–10.

[16] J. Zhang, Y. Zhou, R. Xi, S. Li, J. Guo, and Y. He, "AmbiEar: mmWave Based Voice Recognition in NLoS Scenarios," *Proc. of the ACM IMWUT*, 2022.

[17] C. Kaiyan, S. Leming, Z. Yuanqing, X. Fu, and H. Jinsong, "Talk2Radar: Talking to mmWave Radars via Smartphone Speaker," in *Proc. of the 43rd IEEE INFOCOM*, 2024, pp. 1–10.

[18] Y. Feng, K. Zhang, C. Wang, L. Xie, J. Ning, and S. Chen, "mmEavesdropper: Signal Augmentation-based Directional Eavesdropping with mmWave Radar," in *Proc. of the 42nd IEEE INFOCOM*, 2023, pp. 1–10.

[19] C. Xu, H. Li, Z. Li, H. Zhang, A. S. Rathore, X. Chen, K. Wang, M.-c. Huang, and W. Xu, "CardiacWave: A MmWave-Based Scheme of Non-Contact and High-Definition Heart Activity Computing," *Proc. of the ACM IMWUT*, vol. 5, no. 3, 2021.

[20] Z. Shi, T. Gu, Y. Zhang, and X. Zhang, "mmBP: Contact-Free Millimetre-Wave Radar Based Approach to Blood Pressure Measurement," in *Proc. of the 20th ACM SenSys*, 2023, p. 667–681.

[21] Z. Chang, F. Zhang, X. Ma, P. Wang, W. Chen, D. Zhang, B. Jouaber, and D. Zhang, "mmECare: Enabling Fine-grained Vital Sign Monitoring for Emergency Care with Handheld MmWave Radars," *Proc. of the ACM IMWUT*, vol. 8, no. 4, 2024.

[22] Y. Cao, S. Zhang, F. Li, Z. Chen, and J. Luo, "hBP-Fi: Contactless Blood Pressure Monitoring via Deep-Analyzed Hemodynamics," in *Proc. of the 43rd IEEE INFOCOM*, 2024, pp. 1211–1220.

[23] H. Wei, Z. Li, A. D. Galvan, Z. Su, X. Zhang, K. Pahlavan, and E. T. Solovey, "IndexPen: Two-Finger Text Input with Millimeter-Wave Radar," *Proc. of the ACM IMWUT*, vol. 6, no. 2, 2022.

[24] H. Xue, Y. Ju, C. Miao, Y. Wang, S. Wang, A. Zhang, and L. Su, "mmMesh: Towards 3D Real-time Dynamic Human Mesh Construction Using Millimeter-wave," in *Proc. of the 19th ACM MobiSys*, 2021, p. 269–282.

[25] Y. Liu, J. Zhang, Y. Chen, W. Wang, S. Yang, X. Na, Y. Sun, and Y. He, "Real-Time Continuous Activity Recognition With a Commercial mmWave Radar," *IEEE Transactions on Mobile Computing*, vol. 24, pp. 1684–1698, 2024.

[26] Y. Wang, Z. Wang, J. A. Zhang, H. Zhang, and M. Xu, "Vital Sign Monitoring in Dynamic Environment via mmWave Radar and Camera Fusion," *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 4163–4180, 2023.

[27] X. Xu, J. Yu, C. Ma, Y. Ren, H. Liu, Y. Zhu, Y.-C. Chen, and F. Tang, "mmECG: Monitoring Human Cardiac Cycle in Driving Environments Leveraging Millimeter Wave," in *Proc. of the 41st IEEE INFOCOM*, 2022, pp. 90–99.

[28] Y. Li, C. Wang, L. Xie, Q. Jin, L. Fan, J. Ning, and S. Lu, "Facial Landmark Detection Based on High Precision Spatial Sampling via Millimeter-wave Radar," *Proc. of the ACM IMWUT*, vol. 8, no. 4, pp. 1–26, 2024.

[29] T. Instruments, "AWR1843 Single-chip 76-GHz to 81-GHz Automotive Radar Sensor Evaluation Module," 2024, https://www.ti.com/tool/AWR1843BOOST.

[30] B. Friedlander, "A Sensitivity Analysis of the MUSIC Algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1740–1751, 1990.

[31] OptiTrack., "OptiTrack," 2024, https://optitrack.com.

[32] S. Rixner, W. J. Dally, U. J. Kapasi, P. Mattson, and J. D. Owens, "Memory Access Scheduling," *ACM SIGARCH Computer Architecture News*, vol. 28, no. 2, pp. 128–138, 2000.

[33] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *Proc. in the 2nd ACCM KDD*, vol. 96, no. 34, 1996, pp. 226–231.

[34] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, 2003.

[35] T. Instruments, "DCA1000 Evaluation Module for Real-time Data Capture and Streaming," 2024, https://www.ti.com/tool/DCA1000EVM?keyMatch=DCA1000EVM.

[36] NVIDIA, "NVIDIA Jetson Orin," 2024, https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/.

[37] Viatom, "Handheld and Wearable ECG Monitor ER2," 2023, https://www.viatomtech.com/er2.

[38] Z. Chen, T. Zheng, C. Cai, and J. Luo, "MoVi-Fi: Motion-robust Vital Signs Waveform Recovery via Deep Interpreted RF Sensing," in *Proc. of the 27th ACM MobiCom*, 2021, pp. 392–405.

[39] W. Xu, W. Song, J. Liu, Y. Liu, X. Cui, Y. Zheng, J. Han, X. Wang, and K. Ren, "Mask Does Not Matter: Anti-spoofing Face Authentication Using mmWave without On-site Registration," in *Proc. of the 28th ACM MobiCom*, 2022, pp. 310–323.

[40] X. Zhang, Y. Zhang, Z. Shi, and T. Gu, "mmFER: Millimetre-wave Radar-based Facial Expression Recognition for Multimedia IoT Applications," in *Proc. of the 29th ACM MobiCom*, 2023, pp. 1–15.

[41] D. Cao, R. Liu, H. Li, S. Wang, W. Jiang, and C. X. Lu, "Cross Vision-RF Gait Re-identification with Low-cost RGB-D Cameras and mmWave Radars," *Proc. of the ACM IMWUT*, vol. 6, no. 3, pp. 1–25, 2022.

[42] H. Yang, D. Zhang, X. Zhang, J. Xiong, Z. Fan, W. Ning, W. Chen, F. Zhang, Z. Han, and D. Zhang, "From spatial domain to temporal domain: Unleashing the capability of cfar for mmwave point cloud generation," *Proc. of the ACM on IMWUT*, vol. 9, no. 2, pp. 1–29, 2025.

[43] X. Liu, H. Liu, J. Zhang, X. Xie, and K. Li, "Multi-user behavioral privacy filtering for mmwave radar sensing," *IEEE Transactions on Mobile Computing*, 2025.

[44] M. Alouzi, H. Yanikomeroglu, and G. K. Kurt, "Adaptive phase shifters for hybrid beamforming in mm-wave systems," *IEEE Transactions on Wireless Communications*, 2024.

[45] J. Tian, S. Li, C. He, and W. Zhu, "Wideband transmissive programmable metasurface for adaptive millimeter-wave beamforming," *Laser & Photonics Reviews*, vol. 19, no. 3, p. 2401333, 2025.

[46] L. Fan, L. Xie, W. Zhou, C. Wang, Y. Bu, and S. Lu, "Beamforming for Sensing: Hybrid Beamforming based on Transmitter-Receiver Collaboration for Millimeter-Wave Sensing," *Proc. of the ACM IMWUT*, vol. 8, no. 2, 2024.

[47] I. K. Jain, R. Reddy Vennam, R. Subbaraman, and D. Bharadia, "mmFlexible: Flexible Directional Frequency Multiplexing for Multi-user mmWave Networks," in *Proc. of the 42nd IEEE INFOCOM*, 2023, pp. 1–10.

[48] Z. Gao, Z. Qi, and T. Chen, "Mambas: Maneuvering Analog Multi-User Beamforming using an Array of Subarrays in mmWave Networks," in *Proc. of the 30th ACM MobiCom*, 2024, p. 694–708.

[49] I. K. Jain, R. Subbaraman, and D. Bharadia, "Two Beams are Better Than One: Towards Reliable and High Throughput mmWave Links," in *Proc. of the 35th ACM SIGCOMM*, 2021, p. 488–502.

[50] M. Hashemi, A. Sabharwal, C. Emre Koksal, and N. B. Shroff, "Efficient Beam Alignment in Millimeter Wave Systems Using Contextual Bandits," in *Proc. of the 37th IEEE INFOCOM*, 2018, p. 2393–2401.

[51] F. Sohrabi, Z. Chen, and W. Yu, "Deep Active Learning Approach to Adaptive Beamforming for mmWave Initial Alignment," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2347–2360, 2021.

[52] J. Chen, W. Feng, J. Xing, P. Yang, G. E. Sobelman, D. Lin, and S. Li, "Hybrid Beamforming/Combining for Millimeter Wave MIMO: A Machine Learning Approach," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 353–11 368, 2020.

[53] R. Rajamäki, S. P. Chepuri, and V. Koivunen, "Hybrid Beamforming for Active Sensing Using Sparse Arrays," *IEEE Transactions on Signal Processing*, vol. 68, pp. 6402–6417, 2020.

[54] M. Xu, Y. He, X. Li, J. Hu, Z. Chen, F. Xiao, and J. Luo, "Beamforming made Malicious: Manipulating Wi-Fi Traffic via Beamforming Feedback Forgery," in *Proc. of the 30th ACM MobiCom*, 2024, pp. 908–922.

[55] I. P. Roberts, S. Vishwanath, and J. G. Andrews, "LoneSTAR: Analog Beamforming Codebooks for Full-Duplex Millimeter Wave Systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 5754–5769, 2023.

[56] G. Zhu, K. Huang, V. K. Lau, B. Xia, X. Li, and S. Zhang, "Hybrid beamforming via the Kronecker decomposition for the millimeter-wave massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2097–2114, 2017.

[57] K. Zheng, W. Zhao, T. Woodford, R. Zhao, X. Zhang, and Y. Hua, "Enhancing Mmwave Radar Sensing Using a Phased-mimo Architecture," in *Proc. of the 22nd ACM MobiSys*, 2024, pp. 56–69.

Long Fan received his Ph.D. from the School of Computer Science at Nanjing University, China, in 2025. His research centers on mmWave sensing systems and ubiquitous computing, with a particular interest in advancing mobile and pervasive sensing technologies.

Lei Xie received his B.S. and Ph.D. degrees in Computer Science from Nanjing University, China, in 2004 and 2010, respectively. He is currently a Professor in the School of Computer Science at Nanjing University. He has published over 150 papers in IEEE Transactions on Mobile Computing, IEEE Transactions on Networking, IEEE Transactions on Parallel and Distributed Systems, ACM Transactions on Sensor Networks, ACM MobiCom, UbiComp, MobiHoc, IEEE INFOCOM, ICNP, ICDCS, etc.

Shiyuan Ma is a Ph.D. candidate in the School of Computer Science at Nanjing University, China. His research interests include RFID sensing systems and backscatter communication.

Jingyi Ning received her Ph.D. degree in computer science from Nanjing University, China in 2024. She is currently an assistant researcher in the School of Computer Science at Nanjing University. Her research interests include Intelligent Sensing, Mobile Computing, and Multi-Modal Sensing.

Chuyu Wang received his Ph.D. degree in computer science from Nanjing University, China in 2018. He is currently an assistant professor in the School of Computer Science at Nanjing University. His research interests include RFID systems, Software-Defined Radio, activity sensing, indoor localization, etc.

Wenhui Zhou is a Ph.D. candidate in the School of Computer Science at Nanjing University. His research centers on cloud-edge collaborative video analytics systems, with a particular interest in edge computing and edge intelligence .

Jun Luo (Fellow, IEEE) received his BS and MS degrees in Electrical Engineering from Tsinghua University, China, and the Ph.D. degree in Computer Science from EPFL (Swiss Federal Institute of Technology in Lausanne), Lausanne, Switzerland. From 2006 to 2008, he has worked as a postdoctoral research fellow in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. In 2008, he joined the faculty of Nanyang Technological University in Singapore, where he is currently an Associate Professor. His research interests include mobile and pervasive computing, wireless networking, machine learning and computer vision, applied operations research, as well as security. More information can be found at http://www3.ntu.edu.sg/home/junluo.

Appendix A
Derivation of the CRB for CoSense

In this appendix, we derive the Cramér-Rao Bound (CRB) for range and azimuth angle estimation under a Rician fading model, which captures both the deterministic line-of-sight (LoS) target reflection and stochastic non-line-of-sight (NLoS) multipath interference commonly encountered in dynamic sensing environments.

Signal Model: The propagation channel $\mathbf{H} \in \mathbb{C}^{M\times N}$ between N transmit and M receive elements follows a Rician distribution:

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\mathbf{H}_{\mathrm{LoS}} + \sqrt{\frac{1}{K+1}}\mathbf{H}_{\mathrm{NLoS}}, \tag{A.1}$$

where K denotes the Rician factor (in our experiment, $k = 15$.). The deterministic LoS component corresponding to a dominant point target at range r and azimuth θ_0 is given by

$$\mathbf{H}_{\mathrm{LoS}}(r,\theta_0) = e^{-j\psi(r)}\mathbf{a}_{\mathrm{Rx}}(\theta_0)\mathbf{a}^*_{\mathrm{Tx}}(\theta_0), \tag{A.2}$$

where $(\cdot)^*$ denotes the Hermitian transpose, $\mathbf{a}_{\mathrm{Tx}}(\cdot)$ and $\mathbf{a}_{\mathrm{Rx}}(\cdot)$ are the transmit and receive array response vectors, respectively, and

$$\psi(r) = \frac{4\pi r}{\lambda} \tag{A.3}$$

denotes the round-trip propagation phase at carrier wavelength λ. The NLoS component $\mathbf{H}_{\mathrm{NLoS}}$ has independent, identically distributed complex Gaussian entries, modeled as $\mathcal{CN}(0,1)$, to represent diffuse multipath reflections.

After transmit and receive beamforming using steering vectors $\mathbf{u}(\theta)$ and $\mathbf{v}(\phi)$, the received baseband signal can be written as

$$y = s^{\mathrm{Tx}}h^{\mathrm{target}}\mathbf{v}^*(\phi)\mathbf{H}\mathbf{u}(\theta) + n, \tag{A.4}$$

where $n \sim \mathcal{CN}(0,\sigma^2)$ is additive white Gaussian noise. When the beams are perfectly aligned with the target ($\theta = \phi = \theta_0$), the observation y follows a complex Gaussian distribution $y \sim \mathcal{CN}(\mu, \Sigma_y)$.

Under a dominant-LoS regime ($K \gg 1$), the mean of y is

$$\mu(r,\theta_0) = \sqrt{\frac{K}{K+1}}\, s^{\mathrm{Tx}}h^{\mathrm{target}}\, e^{-j\psi(r)}\,\sqrt{MN}, \tag{A.5}$$

where we assume unit-norm beamformers matched to the steering vectors, $\mathbf{v}^*\mathbf{a}_{\mathrm{Rx}} = \sqrt{M}$, $\mathbf{a}^*_{\mathrm{Tx}}\mathbf{u} = \sqrt{N}$. The effective variance Σ_y accounts for both thermal noise and the residual NLoS-induced fluctuation:

$$\Sigma_y = \frac{|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2}{K+1}MN + \sigma^2. \tag{A.6}$$

We assume that Σ_y is independent of the estimation parameters (r,θ_0), which is reasonable under far-field and dominant-LoS assumptions. Consequently, the Fisher information matrix for the parameter vector $\boldsymbol{\eta} = [r,\theta_0]^T$ reduces to

$$\mathbf{J}(\boldsymbol{\eta}) = \frac{2}{\Sigma_y}\Re\left\{\left(\frac{\partial\mu}{\partial\boldsymbol{\eta}}\right)^*\left(\frac{\partial\mu}{\partial\boldsymbol{\eta}}\right)\right\}. \tag{A.7}$$

CRB for Range Estimation: The derivative of the mean signal with respect to range r is

$$\frac{\partial\mu}{\partial r} = -j\frac{\partial\psi(r)}{\partial r}\mu = -j\frac{4\pi}{\lambda}\mu. \tag{A.8}$$

Accordingly, the corresponding Fisher information matrix element is

$$J_{rr} = \frac{2}{\Sigma_y}\frac{K}{K+1}|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2\left(\frac{\partial\psi(r)}{\partial r}\right)^2 MN. \tag{A.9}$$

The CRB for range estimation is therefore given by

$$\boxed{\mathrm{CRB}_r = \frac{\frac{|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2}{K+1}MN + \sigma^2}{2\frac{K}{K+1}|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2\left(\frac{\partial\psi(r)}{\partial r}\right)^2 MN}}. \tag{A.10}$$

As $K \to \infty$, this expression naturally reduces to the classical CRB for an ideal point target in additive white Gaussian noise.

CRB for Angle Estimation: The sensitivity of the received signal to the azimuth angle θ_0 is determined by the angular gradients of the transmit and receive array response vectors. Under perfect beam alignment, i.e., $\theta = \phi = \theta_0$, the corresponding Fisher information matrix element associated with θ_0 can be expressed as

$$J_{\theta\theta} = \frac{2}{\Sigma_y}\frac{K}{K+1}|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2\mathcal{G}(\theta_0), \tag{A.11}$$

where $\mathcal{G}(\theta_0)$ denotes the array angular sensitivity gain factor, which jointly characterizes the contribution of the transmit and receive arrays to angle estimation accuracy, and is defined as:

$$\begin{aligned}\mathcal{G}(\theta_0) &= \|\partial_\theta\mathbf{a}_{\mathrm{Rx}}(\theta_0)\|_2^2\cdot\|\mathbf{a}_{\mathrm{Tx}}(\theta_0)\|_2^2\\ &\quad+\|\partial_\theta\mathbf{a}_{\mathrm{Tx}}(\theta_0)\|_2^2\cdot\|\mathbf{a}_{\mathrm{Rx}}(\theta_0)\|_2^2.\end{aligned} \tag{A.12}$$

Accordingly, the CRB for azimuth angle estimation is given by

$$\boxed{\mathrm{CRB}_\theta = \frac{\frac{|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2}{K+1}MN + \sigma^2}{2\frac{K}{K+1}|s^{\mathrm{Tx}}h^{\mathrm{target}}|^2\mathcal{G}(\theta_0)}}. \tag{A.13}$$

The explicit dependence on the Rician factor K reveals the degradation of estimation accuracy in cluttered environments, where NLoS components effectively act as an additional noise source. This derivation theoretically justifies CoSense's design principle: by adaptively aligning transmit and receive beams to maximize the effective array gain MN, the system directly minimizes the CRBs for both range and angle estimation.