DoppDrive: Doppler-Driven Temporal Aggregation for Improved Radar Object Detection
Supplementary Material

Yuval Haitman[*] and Oded Bialer[*]

General Motors, Technical Center Israel

haitman@post.bgu.ac.il, oded.bialer8@gmail.com

1. Additional Qualitative Results

In this section, we provide qualitative examples supplementary to those shown in Fig. 2 of the main paper. Fig. 1 showcase an example from *LRR-Sim* dataset, while Figs. 2 and 3 present two examples from the aiMotive dataset. The layout follows the same structure as Fig. 2 in the main paper. For each example, the single-frame point cloud (without aggregation) is shown on the left, standard aggregation over 0.7s is displayed in the middle, and the results of *DoppDrive* are presented on the right. Reflection points from dynamic and static objects are plotted in blue and orange, respectively. Detected bounding boxes using SMF [5] are marked in blue, while ground truth bounding boxes are marked in pink.

In all the examples, it is evident that without aggregation, the reflection points are sparse, often with only one or two points per object, and sometimes none, leading to missed object detections. Standard aggregation results in significant scatter of reflection points from dynamic objects (blue points). In contrast, *DoppDrive* produces dense reflection points with minimal scatter, resulting in detections that are better aligned with the annotations compared to standard aggregation.

2. Computational Complexity Evaluation

We next assess the computational overhead of *DoppDrive*, implemented according to Algorithm 1 in the main paper, with $g(\theta)$ precomputed for each θ and accessed via a lookup table. Tab. 1 compares the runtime of SMF [5] object detection with *DoppDrive* aggregation, standard aggregation, and no aggregation on an NVIDIA Tesla V100 SXM2 32 GB. The results show that *DoppDrive*'s aggregation time is significantly smaller than the detection time, contributing minimal runtime overhead and making it suitable for real-time applications. *DoppDrive* slightly reduces detector

---

runtime compared to standard aggregation by aggregating fewer points due to its dynamic aggregation duration.

Aggregation Method	Aggregation Time [ms]	Detection Time [ms]	Total Time [ms]
None	0	66.4	66.4
Standard	1.4	67.8	69.2
DoppDrive	2.0	67.5	69.5

Table 1. Runtime assessment of *DoppDrive* using the SMF detector [5]. The columns, from left to right, show the point cloud aggregation methods, aggregation runtime, object detection runtime with the aggregated points, and the total runtime, calculated as the sum of aggregation and detection runtimes.

3. Impact of Point Elimination

DoppDrive limits integration duration to suppress points with high tangential dispersion, resulting in point elimination relative to fixed-duration aggregation. This subsection evaluates the elimination rate and its impact on performance.

Tab. 2 breaks down the AP gains and the average percentage of eliminated dynamic points per frame—relative to fixed-duration aggregation with radial compensation (rows 8 vs. 6 in Table 3 of the main paper)—across heading, velocity, and range bins. Elimination reaches up to 16%, increasing at short ranges, higher speeds, and more tangential headings. While the overall elimination rate is modest, it has a noticeable impact on performance, with larger gains observed at higher elimination levels.

4. Calculation of $g(\theta_k^i)$

In Eq. 10 of the main paper, we provide the expression for $g(\theta_k^i)$, which is defined as:

$$g(\theta_k^i) = \mathbb{E}_{\alpha_k^i}\left[\tan(\theta_k^i + \alpha_k^i)\right] = \int_{-\pi/2}^{\pi/2} \tan(\theta_k^i + \alpha) f_{\alpha_k^i}(\alpha)d\alpha, \tag{1}$$
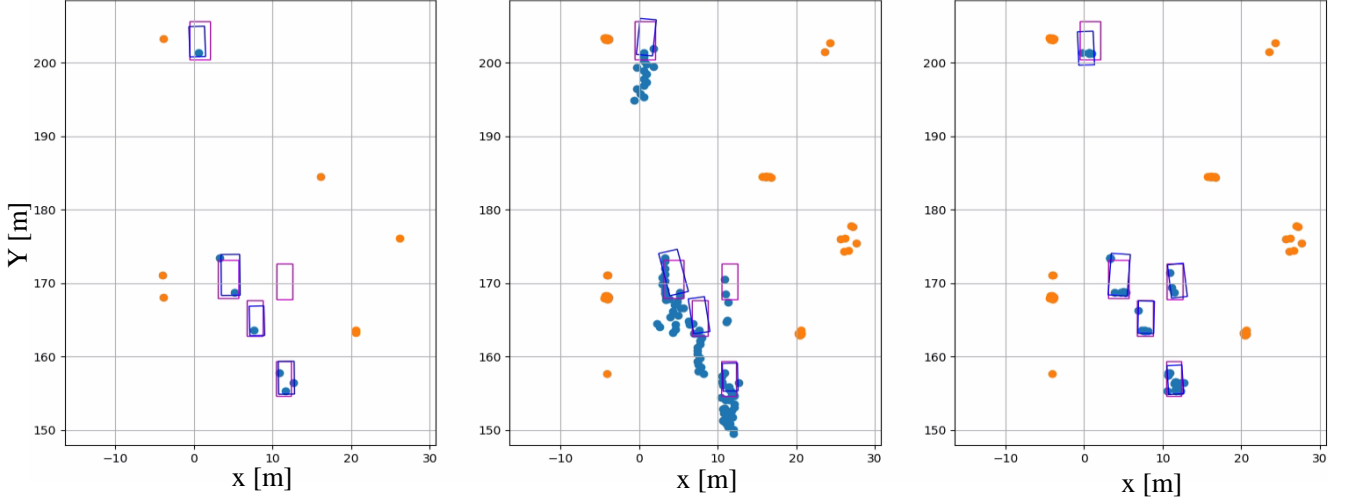
Figure 1. Qualitative example from *LRR-Sim* shown in bird's-eye view. The point cloud of dynamic objects is shown in blue, while that of static objects is shown in orange. SMF [5] detections are represented by blue bounding boxes, and ground truth bounding boxes are shown in pink. The left figure depicts results without aggregation, the middle figure shows standard aggregation with ego-motion compensation, and the right figure shows the results of *DoppDrive*.
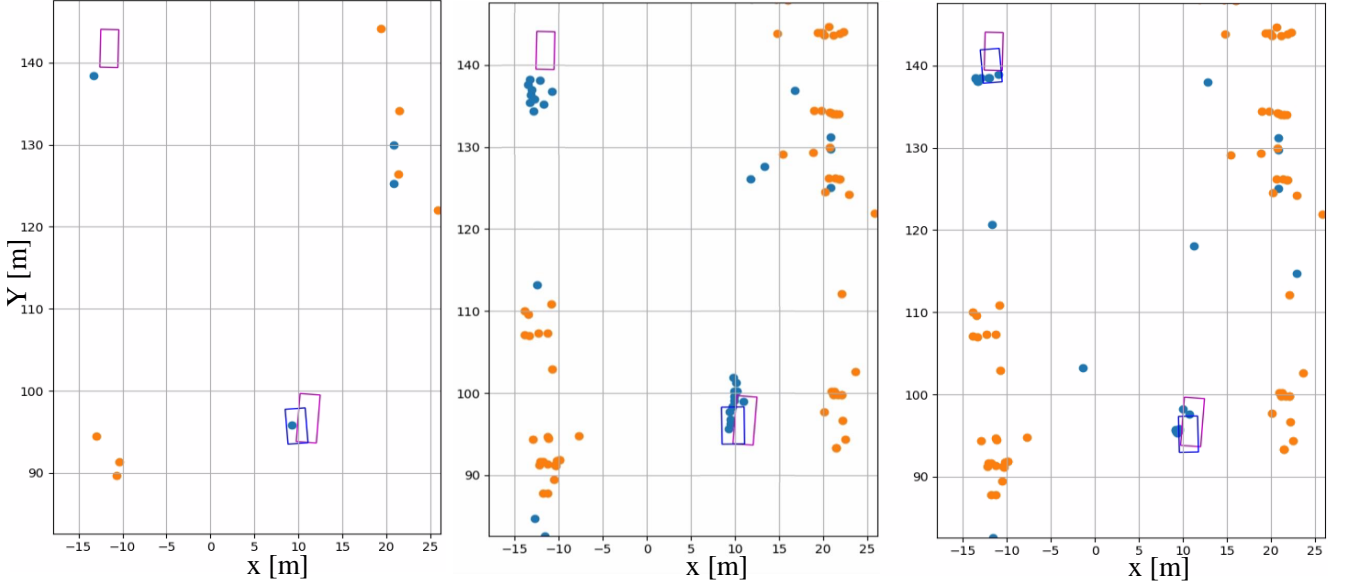

Figure 2. Qualitative example from aiMotive shown in bird's-eye view. The point cloud of dynamic objects is shown in blue, while that of static objects is shown in orange. SMF [5] detections are represented by blue bounding boxes, and ground truth bounding boxes are shown in pink. The left figure depicts results without aggregation, the middle figure shows standard aggregation with ego-motion compensation, and the right figure shows the results of *DoppDrive*.

where $f_{\alpha_k^i}(\alpha)$ represents the probability distribution function of the heading angle α. In this section, we explain how to compute $g(\theta_k^i)$ as defined by Eq. (1).

Fig. 4 shows the probability density function (PDF) of α, derived from annotations in the aiMotive dataset. The distribution is sharply concentrated around zero, with significantly smaller secondary peaks near $\pm 90°$. To model

this distribution, we fit a truncated Laplace distribution:

$$f_{\alpha_k^i}(\alpha) = \frac{1}{2b} \exp\left(-\frac{|\alpha - \mu|}{b}\right), \quad \alpha \in [-\pi/2, \pi/2], \quad (2)$$

where we set $\mu = 0$ and $b = 3.1$. The minor peaks near $\pm 90°$ were neglected in this approximation. The fitted Laplace distribution is plotted in black in Fig. 4. As
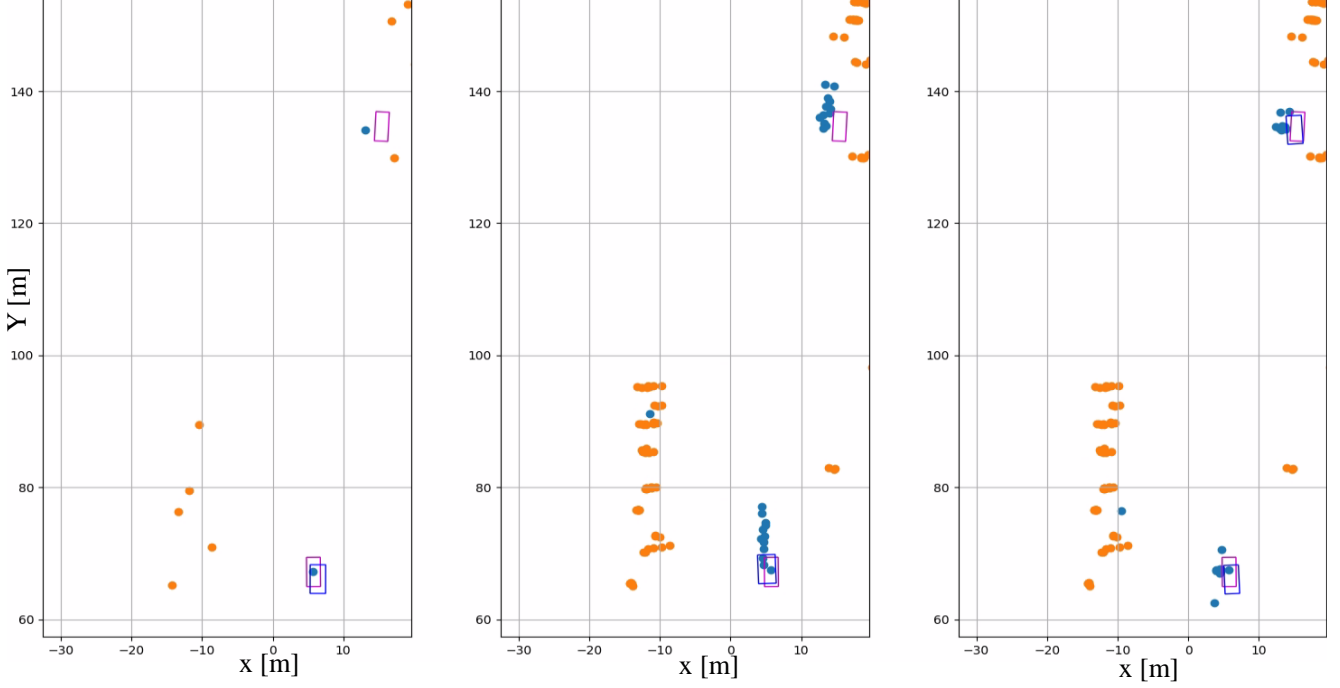
Figure 3. Qualitative example from aiMotive shown in bird's-eye view. The point cloud of dynamic objects is shown in blue, while that of static objects is shown in orange. SMF [5] detections are represented by blue bounding boxes, and ground truth bounding boxes are shown in pink. The left figure depicts results without aggregation, the middle figure shows standard aggregation with ego-motion compensation, and the right figure shows the results of *DoppDrive*.

Range $[m]$	$[0, 40]$	$[40, 80]$	$[80, 120]$	$[120, 160]$
Elim.%/AP Gain	16.6/5.6	12.4/4.6	10.9/2.9	10.5/2.7
Abs. Velocity $[m/s]$	$[0, 12]$	$[12, 24]$	$[24, 36]$	$[36, 48]$
Elim.%/AP Gain	10.6/3.7	11.1/4.3	11.8/4.9	12.6/5.3
Abs. Heading $[°]$	$[0, 30]$	$[30, 60]$	$[60, 120]$	$[120, 180]$
Elim.%/AP Gain	10.0/3.9	11.0/4.1	15.5/5.1	11.0/4.0

Table 2. Point elimination rate (blue) and AP gain (green) over radial compensation with fixed aggregation duration and no elimination, at different range, velocity, and heading intervals (aiMotive).

shown, it aligns well with the main mode of the empirical distribution.

Based on this, the integral in Eq. (1) can be evaluated numerically, either using the empirical histogram or the Laplace approximation for $f_{\alpha_k^i}(\alpha)$. For the experiments in the main paper, we used the Laplace approximation to pre-compute $g(\theta_k^i)$ over a grid of θ_k^i values and stored the results in a look-up table for efficient use in Algorithm 1.

Using the full multi-modal histogram of α from Fig. 4 to compute the integral numerically yielded a slight performance improvement. Specifically, the SMF detector [5] achieved a 0.4-point increase in Average Precision on the aiMotive dataset.
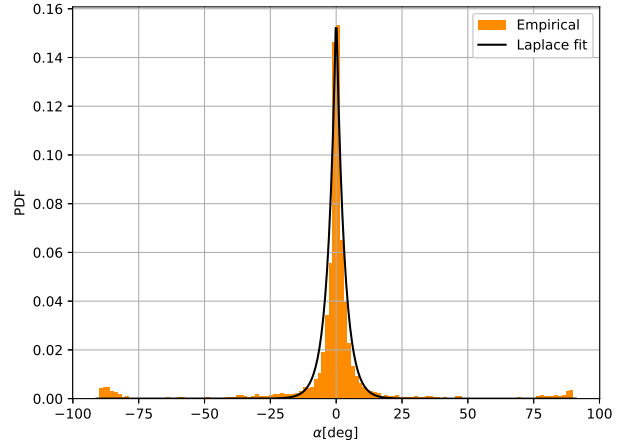

Figure 4. Probability distribution function (PDF) of reflection points' heading directions used in calculating $g(\theta_k^i)$. The empirical PDF from aiMotive dataset annotations is shown in orange, with the Laplace distribution fit plotted in black.

5. Supplementary Details on *LRR-Sim*

In this section, we provide additional details about *LRR-Sim*, the Long-Range Radar Simulation dataset introduced in Section 5 of the main paper. *LRR-Sim* focuses on high-

way scenarios with long-range vehicles of three types: 'car,' 'van,' and 'truck'. The number of vehicles as a function of range in the dataset is presented in Fig. 5, which indicating that vehicle ranges extend up to $300m$, with 17.5% of the vehicles located beyond $175m$. The distribution of vehicle types in the dataset is illustrated in Fig. 6. It is observed that 65.7% of the vehicles are of type 'car,' representing relatively small-sized vehicles, while 17.9% are medium-sized 'vans,' and 16.5% are large vehicles of type 'truck.'

The specifications of the simulated radar are detailed in Tab. 3. These specifications align with those of high-end long-range automotive radars [1–3], offering a detection range of up to $300m$ and high resolution in range, angle, and Doppler.

The simulated radar features 12 transmit and 16 receive antennas and uses a fast chirp FMCW waveform. The highway environment and vehicles were modeled using the CARLA simulation platform [4], incorporating dense reflection points from objects. Ray tracing was performed to compute interactions between the antennas and reflection points in the environment. The received signal at each antenna was the sum of the transmitted waveforms echoed by all reflection points in the scene, with adjustments for intensity and delay determined by each reflection point's intensity and distance. To obtain the radar reflection intensity spectrum in the range, Doppler, and angle dimensions, standard radar signal processing techniques were applied to the received signals. This included range FFT, Doppler FFT, and azimuth-elevation beamforming. Finally, the radar point cloud was generated by identifying spectrum peaks that exceeded a local noise threshold, using the CFAR algorithm [8]. A video demonstration of *LRR-Sim* is included our GitHub[1].
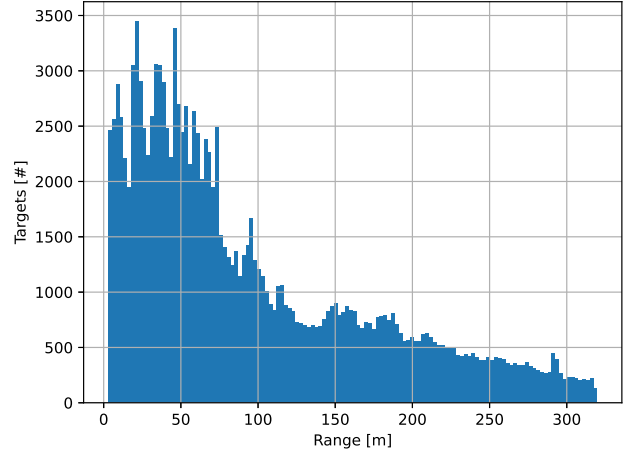
Figure 5. Number of vehicles by range in the *LRR-Sim* dataset.
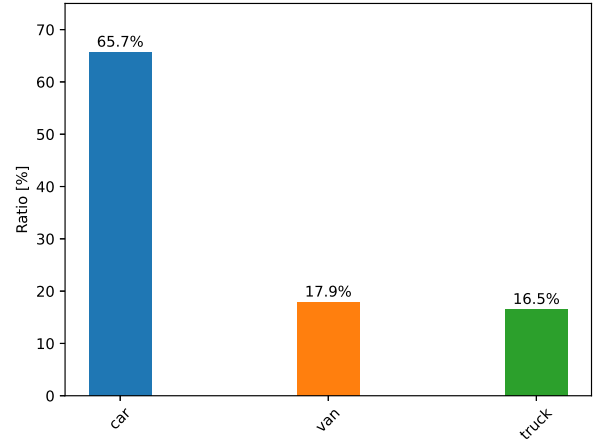
Figure 6. Distribution of vehicle types in the *LRR-Sim* dataset.

Radar parameter	
Maximal range	$300m$
Range resolution	$0.15m$
Azimuth field of view	$\pm55°$
Azimuth resolution	$1.2°$
Elevation field of view	$\pm20°$
Elevation resolution	$2°$
Doppler resolution	$0.13m/s$
Doppler range	$[-80, 30]m/s$

Table 3. Specifications of radar used in *LRR-Sim*

6. Detectors Implementation Details

For the performance evaluation of *DoppDrive* in Section 6 of the main paper, we used four detectors: Radar PillarNet (RPNet) [9], SMURF (SMF) [5], Nvradarnet (NVR) [7],

---

and K-Radar (KRD) [6]. Details of their implementation are outlined below. For all methods, ego-speed Doppler components were removed from the Doppler measurements, and the input features for each point included the x, y, z coordinates, Doppler, reflection intensity, and time index for aggregated inputs. RPNet, SMF, and NVR used a $0.25m \times 0.25m$ input grid resolution in the X-Y plane, while KRD employed a $0.4m \times 0.4m \times 0.4m$ 3D grid resolution. Three anchor sizes were used to cover small, medium, and large vehicles, each with two orientations: $0°$ and $90°$. For SMF, we used two KDE blocks with bandwidths $R1 = 1m$ and $R2 = 2m$. We apply radar point cloud augmentations, including flips, slight zoom in/out, and intensity jitter. Translations and rotations were excluded to prevent misalignment of Doppler measurements.

7. Vertical Velocity Extension

In automotive scenarios, vertical displacement over short aggregation durations is typically negligible. Therefore, in

deriving *DoppDrive*, we assume zero vertical velocity and focus solely on longitudinal and lateral motion. However, *DoppDrive* can be extended to account for vertical velocity if required. In this case, only the duration constraint in line 8 of Algorithm 1 is affected. The function g would be updated to incorporate elevation angle and vertical heading distribution, mitigating dispersion in both vertical and tangential directions.

References

[1] Continental AG. Advanced radar sensor ars540. https://www.continental-automotive.com/en/components/radars/long-range-radars/advanced-radar-sensor-ars540.html. Accessed: Nov. 10, 2024. 4

[2] ZF Friedrichshafen AG. Automated driving functions: Zf captures contract for 4d full-range radar with chinese oem, 2021. Accessed: Nov. 11, 2024.

[3] Aptiv. Radars — advanced safety. https://www.aptiv.com/en/solutions/advanced-safety/adas/radars. Accessed: Nov. 11, 2024. 4

[4] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017. 4

[5] J Liu, Q Zhao, W Xiong, T Huang, QL Han, and B Zhu SMURF. Spatial multi-representation fusion for 3d object detection with 4d imaging radar., 2023. *DOI: https://doi. org/10.1109/TIV*, pages 1–14, 2023. 1, 2, 3, 4

[6] Dong-Hee Paek, Seung-Hyun Kong, and Kevin Tirta Wijaya. Enhanced k-radar: Optimal density reduction to improve detection performance and accessibility of 4d radar tensor-based object detection. In *2023 IEEE Intelligent Vehicles Symposium (IV)*, pages 1–6. IEEE, 2023. 4

[7] Alexander Popov, Patrik Gebhardt, Ke Chen, and Ryan Oldja. Nvradarnet: Real-time radar obstacle and free space detection for autonomous driving. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6958–6964. IEEE, 2023. 4

[8] Hermann Rohling. Radar cfar thresholding in clutter and multiple target situations. *IEEE transactions on aerospace and electronic systems*, (4):608–621, 1983. 4

[9] Lianqing Zheng, Sen Li, Bin Tan, Long Yang, Sihan Chen, Libo Huang, Jie Bai, Xichan Zhu, and Zhixiong Ma. Rcfusion: Fusing 4-d radar and camera with bird's-eye view features for 3-d object detection. *IEEE Transactions on Instrumentation and Measurement*, 72:1–14, 2023. 4