

Collaborative Position and Orientation Tracking in mmWave Vehicular Systems Exploiting V2I and V2V Links

Yun Chen[†], Nuria González-Prelcic[†], Takayuki Shimizu[‡], and Chinmay Mahabal[‡]

[†] University of California San Diego, Email: {yuc216, ngprelcic}@ucsd.edu

[‡] Toyota Motor North America, Email: {takayuki.shimizu, chinmay.mahabal}@toyota.com

Abstract—In this paper, we provide a collaborative position tracking solution exploiting the parameters of vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) links to significantly improve the localization accuracy of the vehicles in the group that are in non-line-of-sight (NLOS) conditions. We assume a group size of three vehicles and unknown orientation and clock offsets among the vehicles. The vehicles with a line-of-sight (LOS) channel to the base station (BS) act as *anchors* and are localized through V2I geometric localization, while the other vehicles in the group are considered *targets* and are localized via sidelink channel parameters. An additional Kalman filter (KF) stage is considered for vehicle position tracking. The proposed methods are evaluated based on realistic simulations of the channels in urban canyon areas considering a single BS scenario, where vehicles frequently enter and exit NLOS situations. Simulation results indicate that orientation offset estimation errors remain below 0.05° , while clock offset estimation errors are under 10 ns. The average localization accuracy is 0.2 m, with errors ≤ 0.7 m for 90% of the cases.

I. INTRODUCTION

Communication at mmWave enables highly accurate user localization as a byproduct, thanks to large antenna arrays and wide bandwidths that enhance angular and delay resolution, as well as channel sparsity that facilitates geometric localization [1]–[3]. However, the performance of V2I localization degrades in NLOS conditions, since the channel parameter estimation error is larger [3]. The exploitation of sidelinks, as specified in [4], emerges as a promising option for user positioning given the high likelihood of a strong LOS component [5].

Localization leveraging sidelinks typically relies on ranging-based methods [6]—where the round trip time (RTT), the time of arrival (ToA) or the time difference of arrival (TDoA) are leveraged for distance calculations [7]—and angle based solutions—where the target can be localized at the intersection of the angular directions [4]. These techniques can be combined for more accurate localization. In a multi-vehicle scenario, collaborative positioning leverages information from multiple links to improve localization

performance. Filtering based algorithms, as proposed in [8], [9], involve collection of the observations including channel parameters [8], global navigation satellite system (GNSS) data and inertial measurement unit (IMU) data [9], and fusion with a global filter. Bayesian inference-based solutions have been widely proposed for collaborative localization [10]–[13]. In [10], ToA is estimated for ranging and fused with GNSS following a maximum-likelihood formulation. In [11], [12] the authors exploit factor graphs, treating delay and angles as observations, and estimate vehicle locations through cooperative belief propagation (BP). Fault-tolerant localization is achieved in [12] through integrity for BP. Moreover, [13] introduces an undirected graph for measurement exchanges, where the localization methods exploit least-mean-square (LMS) and conjugate gradient algorithms. Finally, solutions based on deep learning have also been developed in previous work [14], [15]. In this case, the designed neural networks, e.g., attention networks [14] or long short-term memory (LSTM) networks [15], gather the observations for predictions of the vehicle locations and the environment map.

Solutions in prior work have, however, different limitations: 1) Requirement of measurements from multiple BSs [9], [16]; 2) Neglect of clock offsets [10], [12], [13], which is a major limitation for the accuracy in the ranging process; 3) Lack of solutions to address the high complexity of time domain high-resolution channel estimation that accounts for realistic filtering effects [8]–[13]; 4) Availability of perfect channel/channel parameter distributions, or employing unrealistic synthetic channels; 5) Simplification of the localization schemes to 2D [9], [13]; 6) Neglect of the offset in the orientation of transmit and receive antenna arrays; 7) Limited accuracy (~ 4 m) provided by data driven solutions in vehicular scenarios [14], [15].

In this paper, we consider a partial coverage V2X scenario as outlined in 3GPP technical reports [4], with at least one of the communicating users in-coverage and with a connection to a BS. We first formulate the expression of the mmWave V2I and V2V channels accounting for both the unknown clock and orientation offsets. Then we propose a positioning strategy that involves localizing an anchor vehicle with a LOS to the BS through the V2I downlink as in [3], while NLOS

This material is based upon work partially supported by the National Science Foundation under grant no. 2433782 and is supported in part by funds from the federal agency and industry partners as specified in the Resilient & Intelligent NextG Systems (RINGS) program and in part by Toyota Motor North America.

targets are positioned via sidelinks. We employ the ESPRIT-D algorithm in [17] for sidelink channel estimation, which accounts for the filtering effects and yields high-accuracy channel estimates. The orientation and clock offsets are then derived based on sidelink information shared among the vehicles. Finally, the target locations are determined through geometric transformations from the anchor location. A simple yet effective KF performs a temporal filtering to improve the localization accuracy. The average accuracy of the proposed strategy is 0.2 m, with sub-meter accuracy realized for 90% of the cases.

II. SYSTEM MODEL

We consider a mmWave vehicular network, where a single BS is located at a known position, and three vehicles moving at different speeds are scheduled as a group. At a given time slot, at least one of the vehicles should have a LOS channel with the BS to act as the anchor. The vehicle/s in LOS is localized via the V2I link, while the other/s in NLOS situations will be the target(s) to be collaboratively located through V2V sidelinks. The BS and the vehicles are equipped with uniform rectangular array (URA)s to enable 3D localization.

We first formulate the expression of the channels accounting for unknown clock biases and orientation offsets among the three vehicles. The V2I channel can be written as

$$\mathbf{H}_n^{(\tau)} = \sum_{\ell=1}^L \alpha_{\ell}^{(\tau)} f_p \left(nT_s - (t_{\ell}^{(\tau)} - \Delta t_u) \right). \quad (1)$$

$$\mathbf{a} \left(\theta_{az,\ell}^{(\tau)} - \Delta \vartheta_u^{(\tau)}, \theta_{el,\ell}^{(\tau)} \right) \mathbf{a} \left(\phi_{az,\ell}^{(\tau)}, \phi_{el,\ell}^{(\tau)} \right)^*, \quad (2)$$

where n indicates the n -th tap of the channel, τ is the time frame indicator, L is the number of paths, $\alpha_{\ell}^{(\tau)}$ and $t_{\ell}^{(\tau)}$ are the complex gain and the delay of the ℓ -th path at time τ , Δt_u is the unknown clock offset between the BS and vehicle u , T_s is the sampling period, $f_p(\cdot)$ represents the pulse shaping and filtering effects, $\theta_{az,\ell}^{(\tau)}$ and $\theta_{el,\ell}^{(\tau)}$ represent the azimuth and elevation angle-of-arrival (AoA) at τ , and $\Delta \vartheta_u^{(\tau)}$ is the orientation offset varying every time frame between the array on the BS and that on vehicle u . $\phi_{az,\ell}^{(\tau)}$ and $\phi_{el,\ell}^{(\tau)}$ are defined similarly for the angle-of-departure (AoD). The array response vector is defined as $[\mathbf{a}(\omega^x, \omega^y)]_{m^x N^x + m^y} = \exp(-j\pi m^x \cos(\omega^y) \sin(\omega^x) - j\pi m^y \sin(\omega^y))$, $m^x = 0, 1, \dots, N^x - 1$, $m^y = 0, 1, \dots, N^y - 1$, considering a planar array of size $N_a = N^x \times N^y$ in the yz -plane with half-wavelength element spacing.

Similarly, the channel for V2V link from vehicle u to vehicle v is expressed as

$$\dot{\mathbf{H}}_n^{(\tau)} = \sum_{\ell=1}^{\dot{L}} \dot{\alpha}_{\ell}^{(\tau)} \dot{f}_p \left(nT_s - \left(\dot{t}_{\ell}^{(\tau)} - (\Delta t_v - \Delta t_u) \right) \right). \quad (3)$$

$$\dot{\mathbf{a}} \left(\dot{\theta}_{az,\ell}^{(\tau)} - \Delta \dot{\vartheta}_v^{(\tau)}, \dot{\theta}_{el,\ell}^{(\tau)} \right) \dot{\mathbf{a}} \left(\dot{\phi}_{az,\ell}^{(\tau)} - \Delta \dot{\vartheta}_u^{(\tau)}, \dot{\phi}_{el,\ell}^{(\tau)} \right)^*, \quad (4)$$

where the notations of the different parameters are similar to those for the V2I link, with the dot indicating the sidelink

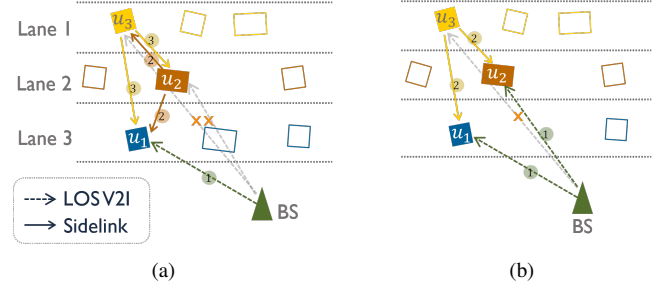


Fig. 1: Sky view of collaborative positioning considering: (a) single-anchor, and (b) multi-anchor cases, where the orientation offsets are unknown. The numbers in the circles indicate the order in the estimation of the link parameters. All the parameters are aggregated at the BS for further processing.

channel parameters. The sidelink is modeled as LOS due to the close proximity between the vehicles in a group.

During the position and channel tracking phase, the training precoders and combiners are designed based on previous channel estimates. We consider a hybrid multiple-input multiple-output (MIMO) architecture at both the BS and the vehicle. The hybrid precoder and combiner can be written as $\mathbf{F} = \mathbf{F}_{RF} \mathbf{F}_{BB}$ and $\mathbf{W} = \mathbf{W}_{RF} \mathbf{W}_{BB}$, where $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$ are the analog and digital precoders, and $\mathbf{W}_{RF}$ and $\mathbf{W}_{BB}$ are the analog and digital combiners. At time slot τ , the q -th instance of the downlink received signal at the vehicle obtained using the m -th precoder-combiner pair is

$$\mathbf{y}_{\tau,m}[q] = \mathbf{W}_{\tau,m}^* \sum_{n=0}^{N_d-1} \sqrt{P} \mathbf{H}_n^{(\tau)} \mathbf{F}_{\tau,m} \mathbf{s}[q-n] + \mathbf{W}_{\tau,m}^* \mathbf{n}_{\tau,m}[q], \quad (5)$$

where N_d is the number of channel taps, P is the transmitted power, $\mathbf{s}[q]$ is the q -th instance of the transmitted pilots, and $\mathbf{n}_{\tau,m}[q]$ represents the additive white Gaussian noise. When some of the vehicles enter the NLOS area and become the targets, the sidelink is activated. The sidelink received signals are written as in (5), with $\mathbf{y}$, $\mathbf{W}$, N_d , P , $\mathbf{H}$, $\mathbf{F}$, and $\mathbf{n}$ replaced by their dotted counterparts $\dot{\mathbf{y}}$, $\dot{\mathbf{W}}$, $\dot{N}_d$, $\dot{P}$, $\dot{\mathbf{H}}$, $\dot{\mathbf{F}}$, and $\dot{\mathbf{n}}$, indicating they correspond to the sidelink. Throughout the localization process, the anchors manage the processing and determination of positions, while the BS coordinates resource allocation and scheduling across the network. The discussion of the scheduling mechanism is outside the scope of this paper.

III. COLLABORATIVE LOCATION AND ORIENTATION TRACKING WITH V2I AND V2V LINKS

Anchors with LOS V2I links are localized based on channel geometry, as detailed in our previous work [3]. Assuming that u_1 is the anchor and the V2I downlink localization result is $\hat{\mathbf{p}}_{u_1}$, the clock offset $\Delta \hat{t}_{u_1}$ is calculated based on the distance between the BS and u_1 , and the orientation offset

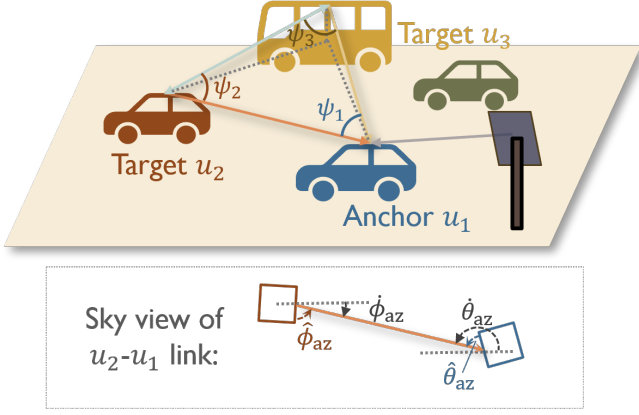


Fig. 2: Illustration of the channel geometry of sidelinks.

$\Delta\hat{\vartheta}_{u_1}$ is acquired based on the estimated AoDs of the LOS component. In this paper, we focus on tracking the NLOS targets in the group. To account for the impact of channel estimation errors, we acquire the channel parameters from the received signals using ESPRIT-D [17], which considers filtering effects in the channel model. While [17] considers the initial access scenario, in this paper we leverage historical location data here to determine the precoders and combiners that point the beams toward the transmitter (receiver) to reduce the signaling overhead. Specifically, for a V2V link from user u_i to u_k at τ , we focus on the LOS components ($\ell = 1$) and denote AoA estimates as $\hat{\theta}_{az,k}^{(\tau)} \approx \hat{\theta}_{az,k}^{(\tau)} - \Delta\vartheta_{u_k}^{(\tau)}$ and $\hat{\theta}_{el,k}^{(\tau)}$, AoD estimates as $\hat{\phi}_{az,i}^{(\tau)} \approx \hat{\phi}_{az,i}^{(\tau)} - \Delta\vartheta_{u_i}^{(\tau)}$ and $\hat{\phi}_{el,i}^{(\tau)}$, and delay estimate as $\hat{t}_{ik}^{(\tau)} \approx \hat{t}_{ik}^{(\tau)} - (\Delta t_{u_k} - \Delta t_{u_i})$, where the orientation and clock offsets are unknown. Then $\hat{\mathbf{F}}_{\text{RF}}^{(\tau+1)}$ and $\hat{\mathbf{W}}_{\text{RF}}^{(\tau+1)}$ are implemented based on the array response vectors considering $\hat{\theta}_{az,k}^{(\tau)}$, $\hat{\theta}_{el,k}^{(\tau)}$, $\hat{\phi}_{az,i}^{(\tau)}$, and $\hat{\phi}_{el,i}^{(\tau)}$, using quantized phase-shifter approximations.

The proposed collaborative positioning algorithms differ for single-anchor and multi-anchor scenarios, offering flexibility based on available information. As the following localization algorithms do not change with the time frame, the time variable τ is omitted in the notations for simplicity.

A. Single-anchor localization

Let's assume that u_1 is the anchor and u_2 and u_3 are the targets, as illustrated in Fig. 1a. The targets u_2 and u_3 take turns sending pilots to anchor u_1 and each other. When user u_2 is transmitting, both u_1 and u_3 estimate the AoA, the AoD, and the delay. The process repeats with u_3 transmitting pilots. We first extract the orientation offsets so all the angular values are computed in the global coordinate system. Then the clock offsets are estimated for ranging and positioning.

Based on the LOS path properties, illustrated in Fig. 2, the estimated angles satisfy

$$\begin{cases} \cos(\hat{\phi}_{az,i} + \Delta\vartheta_{u_i}) = -\cos(\hat{\theta}_{az,k} + \Delta\vartheta_{u_k}) \\ \sin(\hat{\phi}_{az,i} + \Delta\vartheta_{u_i}) = -\sin(\hat{\theta}_{az,k} + \Delta\vartheta_{u_k}) \end{cases} \quad (6)$$

Let $\Delta\vartheta_{ik} = \Delta\vartheta_{u_i} - \Delta\vartheta_{u_k}$, then (6) can be rewritten as

$$\begin{bmatrix} \cos(\hat{\phi}_{az,i}) & -\sin(\hat{\phi}_{az,i}) \\ \sin(\hat{\phi}_{az,i}) & \cos(\hat{\phi}_{az,i}) \end{bmatrix} \begin{bmatrix} \cos(\Delta\vartheta_{ik}) \\ \sin(\Delta\vartheta_{ik}) \end{bmatrix} = \begin{bmatrix} -\cos(\hat{\theta}_{az,k}) \\ -\sin(\hat{\theta}_{az,k}) \end{bmatrix}, \quad (7)$$

where $[\cos(\Delta\vartheta_{ik}), \sin(\Delta\vartheta_{ik})]^T$ is obtained through LS estimation, and $\Delta\vartheta_{ik}$ is derived accordingly. In the current single-anchor case, $\Delta\vartheta_{23}$, $\Delta\vartheta_{21}$, and $\Delta\vartheta_{31}$ are estimated through signal transmission from u_2 to u_3 , u_2 to u_1 , and u_3 to u_1 , respectively. Then $\Delta\vartheta_{u_2}$ and $\Delta\vartheta_{u_3}$ are determined based on the known $\Delta\vartheta_{u_1}$ estimated from the V2I channel by solving

$$\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta\vartheta_{u_2} \\ \Delta\vartheta_{u_3} \end{bmatrix} = \begin{bmatrix} \Delta\vartheta_{23} \\ \Delta\vartheta_{21} + \Delta\vartheta_{u_1} \\ \Delta\vartheta_{31} + \Delta\vartheta_{u_1} \end{bmatrix}. \quad (8)$$

Now we have all angle values in the global coordinate system.

As the targets transmit pilots to each other and the anchor, the RTT between u_2 and u_3 can be obtained as $\frac{\hat{t}_{23} + \hat{t}_{32}}{2}$, and the distance between them is derived as $\hat{d}_{23} = \frac{(\hat{t}_{23} + \hat{t}_{32}) \cdot c}{2}$, where c is the speed of light. The delay estimates $\hat{t}_{21}$ and $\hat{t}_{31}$ are obtained by the anchor u_1 exploiting the law of sine as

$$\frac{(\hat{t}_{21} + \Delta t_{u_1} - \Delta t_{u_2}) \cdot c}{\sin(\psi_3)} = \frac{(\hat{t}_{31} + \Delta t_{u_1} - \Delta t_{u_3}) \cdot c}{\sin(\psi_2)} = \frac{\hat{d}_{23}}{\sin(\psi_1)}, \quad (9)$$

where $\psi_i = \cos^{-1}(\langle \phi_{i1}, \phi_{ik} \rangle)$ for $i, k \in \{2, 3\}$, $i \neq k$, and $\psi_1 = \cos^{-1}(\langle \theta_{21}, \theta_{31} \rangle)$, with ϕ_{ik} and θ_{ik} the direction of departure (DoD) and direction of arrival (DoA) of the LOS path from u_i to u_k . Now, the clock offsets of the targets $\Delta\hat{t}_{u_i}$, $i = 2, 3$ are determined based on $\Delta\hat{t}_{u_1}$ by solving $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & \mathbf{I}_2 \\ \sin(\psi_2) & -\sin(\psi_3) \end{bmatrix}; \mathbf{b} = \begin{bmatrix} \hat{t}_{21} + \Delta\hat{t}_{u_1} - \frac{\hat{d}_{23} \sin(\psi_2)}{c \cdot \sin(\psi_1)} \\ \hat{t}_{31} + \Delta\hat{t}_{u_1} - \frac{\hat{d}_{23} \sin(\psi_3)}{c \cdot \sin(\psi_1)} \\ \sum_{i=2}^3 (-1)^i \sin(\psi_i) (\hat{t}_{i1} + \Delta\hat{t}_{u_1}) \end{bmatrix}, \quad (10)$$

and $\mathbf{x} = [\Delta t_{u_2}, \Delta t_{u_3}]^T$. Note that, though the clock offsets are assumed to be consistent for the entire trajectory, we account for the variations in each time slot due to channel estimation errors, and average the estimates over the trajectory to determine the final clock offsets. Finally, the locations of the targets are derived based on the anchor's location as

$$\hat{\mathbf{p}}_{u_i} = \hat{\mathbf{p}}_{u_1} + \theta_{i1} \cdot c(\hat{t}_{i1} + \Delta\hat{t}_{u_1} - \Delta\hat{t}_{u_i}). \quad (11)$$

B. Multi-anchor localization

We discuss now the scenario where two of the three vehicles in the group, u_1 and u_2 , serve as anchors, while u_3 is the target. As illustrated in Fig. 1b, after the anchors are independently localized from the parameters of their corresponding V2I links, u_3 sends pilots to the anchors, so they can estimate the parameters if the V2V channels. We follow the previous definitions for the channel estimates, and

(6) and (7) can be leveraged to obtain $\Delta\hat{v}_{31}$ and $\Delta\hat{v}_{32}$. Since $\Delta\hat{v}_{u_1}$ and $\Delta\hat{v}_{u_2}$ have already been computed for the localization of the anchors exploiting the V2I links, $\Delta\hat{v}_{u_3}$ can now be obtained by averaging the estimates as $\Delta\hat{v}_{u_3} = (\Delta\hat{v}_{31} + \Delta\hat{v}_{u_1} + \Delta\hat{v}_{32} + \Delta\hat{v}_{u_2})/2$. Now we derive a system of equations similar to the one in (9) to obtain the clock offsets. First, the AoAs in the global coordinate system $\hat{\theta}_{az,i} + \Delta\hat{v}_{u_i}$ and $\hat{\theta}_{el,i}$ as well as the delays $\hat{t}_{3i}$ for $i = 1, 2$ are estimated at each anchor. From these values, their associated directions θ_{31} and θ_{32} can be obtained. In addition, since $\hat{\mathbf{p}}_{u_1}$ and $\hat{\mathbf{p}}_{u_2}$ are known, the DoA for each anchor $\theta_{ik} = \hat{\mathbf{p}}_{u_i} - \hat{\mathbf{p}}_{u_k}$ can also be calculated. Next, the clock offset $\Delta\hat{t}_{u_3}$ can be computed from $\Delta\hat{t}_{u_1}$ and $\Delta\hat{t}_{u_2}$ as

$$\Delta\hat{t}_{u_3} = \frac{\sum_{i=1}^2 (-1)^2 \sin(\psi_i) (\hat{t}_{3i} - \Delta\hat{t}_{u_i})}{\sin(\psi_1) - \sin(\psi_2)}. \quad (12)$$

Finally, we obtain two estimates for $\hat{\mathbf{p}}_{u_3}$ by exploiting (11) and the anchor locations, and take the average as the final estimate for the target position.

To further increase positioning accuracy the system can now exploit the history of previous position estimates with a KF for 2D position tracking. This is implemented independently for each vehicle. For each vehicle. The state of vehicle u_i at τ is defined as $\mathbf{s}_{u_i}^{(\tau)} = [x_{u_i}^{(\tau)}, v_{x,u_i}^{(\tau)}, y_{u_i}^{(\tau)}, v_{y,u_i}^{(\tau)}]^T$, where $x_{u_i}^{(\tau)}$ and $y_{u_i}^{(\tau)}$ represent the x and y coordinates of u_i 's current location, $v_{x,u_i}^{(\tau)} \sim \mathcal{N}(\dot{v}_{x,u_i}, \sigma_{x,u_i}^2)$ and $v_{y,u_i}^{(\tau)} \sim \mathcal{N}(\dot{v}_{y,u_i}, \sigma_{y,u_i}^2)$ are the speeds along x and y axes at τ and change following Gaussian distribution [18], where $\dot{v}_{x,u_i}$ and $\dot{v}_{y,u_i}$ are the mean speeds along x and y axes, and σ_{x,u_i}^2 and σ_{y,u_i}^2 represent the covariance. The measurement vector is defined as the location estimate at τ , i.e., $\mathbf{z}_{u_i}^{(\tau)} = [\hat{\mathbf{p}}_{u_i}^{(\tau)}]_{1:2}^T$. This approach provides a high accuracy even during brief NLOS conditions among all three vehicles along the trajectory, and it uses the state transition function relying on vehicles' speeds to generate measurements.

IV. SIMULATION RESULTS

The experiments are conducted based on the realistic mmWave channels generated from *Wireless Insite*, considering an outdoor urban canyon environment within a rectangular cuboid with opposite vertices at points $[-13m, -123m, 0m]$ and $[231m, 85m, 56m]$. The three vehicles in the group can be either cars of height 1.6 m or trucks of height 3.8 m. Their speeds can take values of 20, 40, or 60 mph. The BS is located at $[120m, -21m]$ with the array height of 5m. The ray-tracing simulation operates at a carrier frequency of 73 GHz, with other parameters chosen as in [3]. We consider the ray-tracing simulation performed every 25 ms given the vehicles moving on their lanes, and we generate a total of 10 sets of the trajectories considering different starting points of the vehicles, each of which contains around 150 simulation snapshots.

The simulated communication system exploits a bandwidth of 1 GHz and operates at the carrier frequency of 73 GHz.

We assume that the vehicles and the BS are equipped with planar arrays of size $N_v = 8 \times 8$ for the vehicles and $N_B = 24 \times 24$ for the BS. A sequence of training pilots with a length $Q = 38$ is sent for channel and position tracking purposes. The precoder and combiner design at time slot τ leverages channel estimates from the preceding time slot $\tau - 1$, i.e., the analog beams pointing to the directions based on the last set of angle estimates and their adjacent beams are considered at each time slot. The transmit power at the BS is 40 dBm for V2I links, while the vehicles transmit power for V2V links is 10 dBm. A raised cosine filter with a roll-off factor of 0.6 is selected for pulse shaping.

We first analyze the performance of the sidelink channel estimation strategy, in addition to the accuracy of the orientation and clock offset estimation results. As illustrated in Fig. 3a, channel tracking with ESPRIT-D provides an angular accuracy of 0.1° for 95% of the situations. The collaborative orientation offset estimation errors are always below 0.05° . Furthermore, the algorithms guarantee a clock offset estimation error below 10 ns. The delay estimations factoring in the estimated clock offset, as shown in Fig. 3b, have an error below 10 ns.

An example of collaborative positioning results for a given trajectory is included in Fig. 4, which shows the sky view of a vehicle u_1 moving at 60 mph acting as the anchor, a truck u_2 moving at 40 mph, and another vehicle u_3 moving at 20 mph. This example primarily features single-anchor cases, with a few multi-anchor instances. Specifically, without sidelink localization, V2I links can locate u_1 , u_2 , and u_3 in 98%, 45%, and 18% of the vehicle locations, respectively. After introducing sidelink positioning, these percentages go up to 98% for all vehicles, while the remaining 2% of the measurements rely on state transition functions. The isolated points in this figure (triangles or squares) represent the results of positioning from a single snapshot, without the temporal smoothing implemented with the KF. The periods marked in grey represent the times where the V2I channel status is NLOS and collaborative positioning via sidelinks is implemented. No performance loss is observed in the gray periods which exploit sidelinks and collaborative localization with respect to V2I localization.

After analyzing the results for all the testing trajectories, we compute the cumulative distribution function (CDF) of the 2D localization error (m), which is presented in Fig. 5. Based on the accuracy of V2I LOS localization, collaborative positioning from a single snapshot provides an average error of 0.4 m for targets, and an error below 1.2 m for 90% of the cases. The accuracy further improves to 0.3 m on average, achieving sub-meter accuracy for 90% of the situations after applying the KF. Sidelink localization performs worse than V2I localization because errors from V2I are propagated when estimating the target locations based on the anchor positions. For comparison, we also include results using V2I NLOS localization [3], which requires ≥ 3 estimated channel paths, while often yielding lower accuracy and poorer localization performance, e.g., 1.2 m localization error on average.

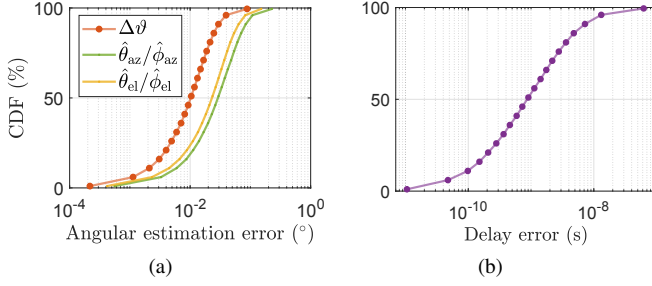


Fig. 3: Sidelink channel estimation results: (a) CDF of the angular estimation errors for 3D AoAs/AoDs and the orientation offsets; (b) CDF of the delay estimation errors factoring in the estimated clock offsets.

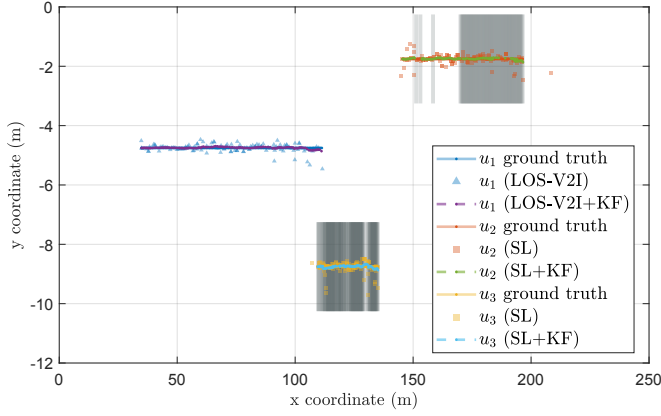


Fig. 4: An example of the sky view of the collaborative positioning and tracking results for two cars (u_1 and u_3) and a truck (u_2) moving on three lanes. The gray areas represent the NLOS occasions for u_2 and u_3 . In the example, u_1 serves as the anchor, u_2 can be either an anchor or a target, and u_3 is always the target.

Ref.	Data type	Localization scheme	Avg. accuracy (m)	90%-th error (m)
[11]	V2I/SL-AoA+ V2I/SL-ToA	Factor Graph+BP	1.2	1.5
[12]	GNSS+SL-ToA	Factor Graph+BP	2.9	3.5
[13]	GNSS+SL-AoA	Graph Laplacian with LMS	1.4	2.5
Proposed	V2I/SL-AoA+ V2I/SL-ToA	Collaborative geometric localization and KF	0.3	0.7

TABLE I: Comparison of several cooperative localization schemes in urban driving scenarios. The proposed approach clearly outperforms the baselines developed in previous work.

Table I presents a comprehensive comparison of our proposed method with various V2V cooperative localization schemes in realistic urban driving scenarios. Methods incorporating V2I and V2V communication [11] outperform GNSS based solutions [12], [13]. Compared to the other collaborative V2V/V2I solution in [11], our approach provides significantly better performance, achieving an average error reduction of 1 m and a 50% improvement in accuracy at the 90th percentile.

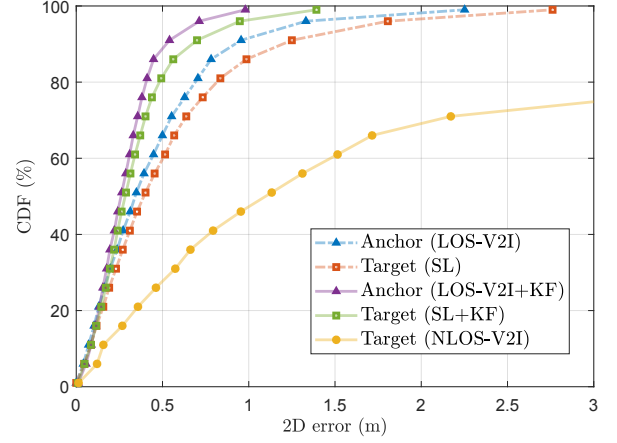


Fig. 5: CDF of the 2D error (m) with the proposed collaborative positioning, including V2I NLOS localization [3] for comparison.

V. CONCLUSION

In this paper, we considered V2X-enabled collaborative positioning accounting for unknown antenna array orientation and clock offsets. Single and multiple anchors scenarios were considered, where the transmission schedule differs depending on the available shared information. Channel estimation and tracking was performed with an algorithm that accounts for the filtering effects introduced by pulse shaping and the analog front-ends. Experimental results show that the orientation offset estimation errors are below 0.05° and the clock offset errors are below 10 ns both for 95% of the situations. Based on these parameters, location estimates for the targets, derived based on geometric transformations from a single snapshot, achieve an average accuracy of 0.3 m and 1.2 m at the 95th percentile. Applying Kalman filtering for position tracking further improves accuracy, achieving sub-meter localization in 90% of cases. For simplicity, we illustrate our approaches using a group of three vehicles as an example. However, this method can be readily adapted to scenarios involving multiple vehicles.

REFERENCES

- [1] J. Palacios, N. González-Prelcic, and C. Rusu, "Low complexity joint position and channel estimation at millimeter wave based on multidimensional orthogonal matching pursuit," in *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022, pp. 1002–1006.
- [2] M. Bayraktar, J. Palacios, N. González-Prelcic, and C. J. Zhang, "Multidimensional orthogonal matching pursuit-based RIS-aided joint localization and channel estimation at mmwave," in *2022 IEEE 23rd International Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*. IEEE, 2022, pp. 1–5.
- [3] Y. Chen, N. González-Prelcic, T. Shimizu, and H. Lu, "Learning to localize with attention: From sparse mmWave channel estimates from a single BS to high accuracy 3D location," *arXiv preprint arXiv:2307.00167*, 2023.
- [4] 3GPP, "Study on expanded and improved NR positioning," 3GPP, Tech. Rep. TR 38.859 V18.1.0 (2024-06), 2024.

- [5] S. Bartoletti, N. Decarli, B. M. Masini, C. Giovannetti, A. Zanella, A. Bazzi, and R. A. Stirling-Gallacher, "Integration of sensing and localization in V2X sidelink communications," *IEEE Communications Magazine*, vol. 62, no. 8, pp. 185–191, 2024.
- [6] B. Zhou, Z. Liu, and H. Su, "5G networks enabling cooperative autonomous vehicle localization: A survey," *IEEE Transactions on Intelligent Transportation Systems*, 2024.
- [7] Y. Ge, M. Stark, M. F. Keskin, F. Hofmann, T. Hansen, and H. Wymeersch, "V2X sidelink positioning in FR1: Scenarios, algorithms, and performance evaluation," *IEEE Journal on Selected Areas in Communications*, 2024.
- [8] Y. Lu, M. Koivisto, J. Talvitie, E. Rastorgueva-Foi, M. Valkama, and E. S. Lohan, "Cooperative positioning system for industrial IoT via mmWave device-to-device communications," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–7.
- [9] S. Li, D. Duan, J. Zhang, X. Cheng, and L. Yang, "An efficient distributed multi-vehicle cooperative tracking framework via multicast," *IEEE Internet of Things Journal*, 2023.
- [10] X. Qu, T. Liu, F. Li, and W. Tan, "Cooperative localization with the fusion of GNSS and relative range information in vehicular ad-hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 11, pp. 11 864–11 877, 2023.
- [11] W. Meng, X. Chu, Z. Lu, L. Wang, X. Wen, and M. Li, "V2V communication assisted cooperative localization for connected vehicles," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2021, pp. 1–6.
- [12] J. Xiong, Z. Xiong, X. Xie, Y. Zhuang, Y. Zheng, S. Xiong, J. W. Cheong, and A. G. Dempster, "Integrity for belief propagation-based cooperative positioning," *IEEE Transactions on Intelligent Transportation Systems*, 2024.
- [13] N. Piperigkos, A. S. Lalos, and K. Berberidis, "Graph laplacian diffusion localization of connected and automated vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12 176–12 190, 2022.
- [14] C. Chang, J. Zhang, K. Zhang, W. Zhong, X. Peng, S. Li, and L. Li, "BEV-V2X: Cooperative birds-eye-view fusion and grid occupancy prediction via V2X-based data sharing," *IEEE Transactions on Intelligent Vehicles*, 2023.
- [15] L. Zhang, H. Zhao, J. Chen, L. Li, and X. Liu, "Vehicular positioning based on GPS/IMU data fusion aided by V2X networks," *IEEE Sensors Journal*, 2024.
- [16] T. Yin, D. Zou, X. Lu, and C. Bi, "Evolutionary coalitional game-based cooperative localization in vehicular networks," *Electronics*, vol. 11, no. 4, p. 638, 2022.
- [17] Y. Chen, N. González-Prelcic, T. Shimizu, H. Lu, and C. Mahabali, "Beamspace esprite-d for joint 3d angle and delay estimation for joint localization and communication at mmwave," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 1915–1920.
- [18] R. Borsche, A. Klar, and M. Zanella, "Kinetic-controlled hydrodynamics for multilane traffic models," *Physica A: Statistical Mechanics and its Applications*, vol. 587, p. 126486, 2022.