

# Implicit Cooperative Positioning in Vehicular Networks

Gloria Soatti, Monica Nicoli, Nil Garcia, Benoît Denis, Ronald Raulefs and Henk Wymeersch

**Abstract**—Absolute positioning of vehicles is based on Global Navigation Satellite Systems (GNSS) combined with on-board sensors and high-resolution maps. In Cooperative Intelligent Transportation Systems (C-ITS), the positioning performance can be augmented by means of vehicular networks that enable vehicles to share location-related information. This paper presents an Implicit Cooperative Positioning (ICP) algorithm that exploits the Vehicle-to-Vehicle (V2V) connectivity in an innovative manner, avoiding the use of explicit V2V measurements such as ranging. In the ICP approach, vehicles jointly localize non-cooperative physical features (such as people, traffic lights or inactive cars) in the surrounding areas, and use them as common noisy reference points to refine their location estimates. Information on sensed features are fused through V2V links by a consensus procedure, nested within a message passing algorithm, to enhance the vehicle localization accuracy. As positioning does not rely on explicit ranging information between vehicles, the proposed ICP method is amenable to implementation with off-the-shelf vehicular communication hardware. The localization algorithm is validated in different traffic scenarios, including a crossroad area with heterogeneous conditions in terms of feature density and V2V connectivity, as well as a real urban area by using Simulation of Urban MObility (SUMO) for traffic data generation. Performance results show that the proposed ICP method can significantly improve the vehicle location accuracy compared to the stand-alone GNSS, especially in harsh environments, such as in urban canyons, where the GNSS signal is highly degraded or denied.

**Index Terms**—Cooperative positioning, vehicular networks, distributed tracking, message passing, consensus algorithms, intelligent transportation systems (ITS).

## I. INTRODUCTION

Intelligent Transportation Systems (ITSs) are becoming a crucial component of our society and precise vehicle positioning is playing a key role in it, such as for assisted or autonomous driving, fleet management, and road safety [1].

This research was supported in part by “COPPLAR CampusShuttle cooperative perception & planning platform”, funded under Strategic Vehicle Research and Innovation Grant No. 2015-04849, by the Horizon2020 project HIGHTS (High precision positioning for cooperative ITS applications) MG-3.5a-2014-636537 and by the project MIE (Mobilità Intelligente Ecosostenibile) CTN01\_00034\_594122 funded by the Italian Ministry of Education, University and Research (MIUR) within the framework Cluster Tecnologico Nazionale “Tecnologie per le Smart Communities”.

G. Soatti and M. Nicoli are with the Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, 20133 Milano, Italy (e-mail: gloria.soatti@polimi.it; monica.nicoli@polimi.it). N. Garcia and H. Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden (e-mail: nilg@chalmers.se; henkw@chalmers.se). B. Denis is with the Laboratory of Electronics and Information Technology (LETI), French Atomic Energy Commission (CEA), 38054 Grenoble Cedex 9, France (email: benoit.denis@cea.fr). R. Raulefs is with the Institute of Communications and Navigation, German Aerospace Center (DLR), 82234 Wessling, Germany (e-mail: ronald.raulefs@dlr.de).

Global Navigation Satellite Systems (GNSSs), e.g. Global Positioning System (GPS) or Galileo, have been widely used in ITS. Standard GNSS provide an accuracy of 5-10 meters in open sky areas, i.e., when there is a direct line-of-sight between the vehicle receiver and satellites [2]. Augmented by differential corrections and/or multi-constellation receivers, they can achieve meter-level accuracy in ideal operating conditions, or even centimeter level in Real Time Kinematics (RTK) variants [3]. However, RTK is still subject to long and unpredictable convergence times when cold starting. Moreover, in urban canyons, the availability of both standard and advanced GNSS systems is limited by adverse local environmental conditions, as GNSS signals can be substantially degraded or mostly blocked. A possible way to cope with the degradation of GNSS signals is to perform graph-based Simultaneous Localization And Mapping (SLAM) [4]. Vehicles localize themselves by building maps of the surrounding environment and fusing the available GNSS information into the mapping process. However, in these approaches vehicles are considered as autonomous entities and not connected with each other, while cooperation through vehicular networking could provide significant benefits, particularly in the context of cooperative autonomous driving applications [5,6].

In recent years, there has been growing interest on Cooperative ITS (C-ITS) [7], where vehicles are able to share information with other vehicles and/or the network infrastructure (e.g., base stations or Road Side Units, RSUs) through respectively Vehicle-to-Vehicle (V2V) and/or Vehicle-to-Infrastructure (V2I) communications. Currently, Vehicle-to-anything (V2X) communications can be achieved by using the available IEEE 802.11p technology, which is a basis of the Dedicated Short Range Communication (DSRC) [8] and ITS-G5 standards [9], in the USA and Europe, respectively. Even more recently, cooperative communications have been considered to improve the positioning availability, integrity and continuity in the specific vehicular context [10]–[18]. In this context, RSUs can be used for performing vehicle self-localization as they provide wireless connectivity to passing vehicles [19]–[21]. In a first stage, a vehicle retrieves the number and locations of the nearby RSUs [19,20], and then positions itself by measuring two-way time-of-arrivals [21]. However, such solutions require populating roads with many RSUs, and thus their use is limited in the short term. Instead, Cooperative Positioning (CP) approaches [22]–[24], which benefit from mobile-to-mobile interactions (i.e., in terms of both measurements and exchanged positional information), have been increasingly adopted. They enable to mitigate the shortcomings of GNSS by incorporating additional GNSS-

independent information into the positioning problem. Most of these cooperative localization approaches are applied in a more global hybrid data fusion framework [25]. The idea is for instance to combine GNSS measurements with auxiliary information, such as ranging measurements, reference points and digital road maps. Several solutions have thus been proposed based on the measurement of inter-vehicle distances [10]–[18]. For example, the works in [10,12,26] propose CP algorithms based on V2V radio ranging techniques, assisted with road map and/or vehicle kinematics information. In [16], the authors develop a strategy for selecting the best V2V links to be used for the purpose of CP. Another CP approach is proposed in [17], which relies on Radio-Frequency IDentification (RFID) tags installed along the street to compute the GNSS bias for positioning correction. A CP variant worth mentioning is also cooperative map matching [18] where GNSS raw measurements and precise road-map information are exchanged between the vehicles to further mitigate GNSS errors. Finally, a consensus-based method has been proposed in [24] to enable the localization of an entire fleet of entities (i.e., vehicles, pedestrians or any objects), at each member of the fleet, by means of local ranging measurements and repeated device-to-device iterations.

Overall, although the current state-of-the-art approaches have been shown to effectively improve the vehicle positioning accuracy, they either rely on high-complexity techniques, or require dedicated hardware or large-scale infrastructure. In addition, most of them need the vehicle to extract explicit range measurements (e.g., round trip time, time of flight or Received Signal Strength – RSS) from the radio signals exchanged with neighboring mobile entities. Such measurements tend to be of low quality (e.g., RSS measurements) or are incompatible with IEEE 802.11p (e.g., time-based range measurements that rely on unicast). On the one hand, dedicated wireless ranging technologies (e.g., Impulse Radio - Ultra Wideband [15]) indeed require specific acquisition schemes and handshake protocols. Accordingly, they might induce extra latency, as well as additional coordination or synchronization constraints (e.g., local scheduling of ranging packets within a pseudo-coordinated time division access), which are deemed challenging in highly scalable ad-hoc contexts. On the other hand, as IEEE 802.11p relies on broadcast transmissions, accurate explicit measurements from V2V communications are not available. Consequently, the above methods cannot be straightforwardly applied into the current vehicular context.

Asides, vehicles are nowadays equipped with more and more perceptual sensors to detect objects and physical obstacles in their close vicinity. These sensors have been mostly intended for forward collision warning, lane change assistance, automatic park control, autonomous driving and more recently, high-definition cartography [27]. Commercially available Radio Detection And Ranging (RADAR) devices devoted to automotive applications, which operate in millimeter frequency bands, can typically measure relative distances and azimuth angles with respect to passive targets within accuracies of a few decimeters and a few tenths of degree respectively, at ranges up to 250 m and with refresh periods lower than 80 ms. LIght Detection And Ranging (LIDAR) devices, including

rotating 2D laser scanners, can detect points in the plane within accuracies of a few centimeters and a few hundredths of degree, at ranges up to 200 m and with refresh periods lower than 50 ms. Finally, camera-based systems relying on pixel analysis can also achieve centimeter-level ranging accuracy at shorter ranges on the order of 10 m and with refresh rates up to 60 frames per second. Combining range and angle information thus enables unambiguous 2D relative positioning. Even if the three standalone technologies cited above are also subject to limitations (e.g., rainy/foggy/snowy weathers, optical/radio obstructions, cost per unit for laser scanners, etc.), they generally claim much better spatial resolution than active wireless technologies.

*Original Contributions:* In this paper, a new *implicit* cooperative positioning (ICP) technique is proposed to improve the GNSS-based vehicle positioning by sharing information and enabling cooperation amongst vehicles through V2V communication links, while making use of on-board sensing devices rather than active wireless technologies enabling explicit V2V measurements. In particular, an innovative distributed processing framework is proposed where a set of non-cooperative features (e.g., people, traffic lights, trees, etc.) are used as common noisy reference points that are cooperatively localized by the vehicles and *implicitly* used to enhance the vehicle location accuracy. A distributed Gaussian Message Passing (GMP) algorithm is designed to solve the positioning problem, integrating a Kalman filter to track the vehicle dynamics based on the on-board GNSS measurements. Vehicles gather noisy observations about the Vehicle-to-Feature (V2F) relative locations by their on-board equipment (e.g., RADAR, camera-based detector, etc.), where perfect association between vehicle measurements and features is assumed as available. Then, they reach a *consensus* on the features' absolute locations by engaging in a distributed cooperative estimation, which implicitly reflects on a more accurate vehicle positioning. A particle-based implementation of message passing combined with consensus has been proposed in [23], though not specifically for vehicular scenarios; this approach however requires high computational load and communication overhead. In this paper, we propose a distributed Kalman filtering approach which enables a more efficient implementation. Some preliminary results have been presented in [28] for a static environment with fixed features. Here the analysis is extended to include velocity information of both features and vehicles, in order to handle highly dynamic scenarios with mobile features. The ICP method is derived analytically, together with the related performance bounds for the ideal case of all-to-all connectivity, and it is validated in a realistic road scenario. In particular, the overall solution is first validated by simulation in a simulated crossroad scenario for varying levels of V2V/V2F connectivity. Then, the assessment of the proposed ICP algorithm is carried out in a real urban area of Bologna city (Italy) with traffic generated by using Simulation of Urban MObility (SUMO) for mixed environment conditions. Numerical results show that the proposed approach is able to significantly increase the GNSS-based vehicle location accuracy, especially in urban areas with high density of features and cooperative vehicles, compensating the performance degradation that is typically

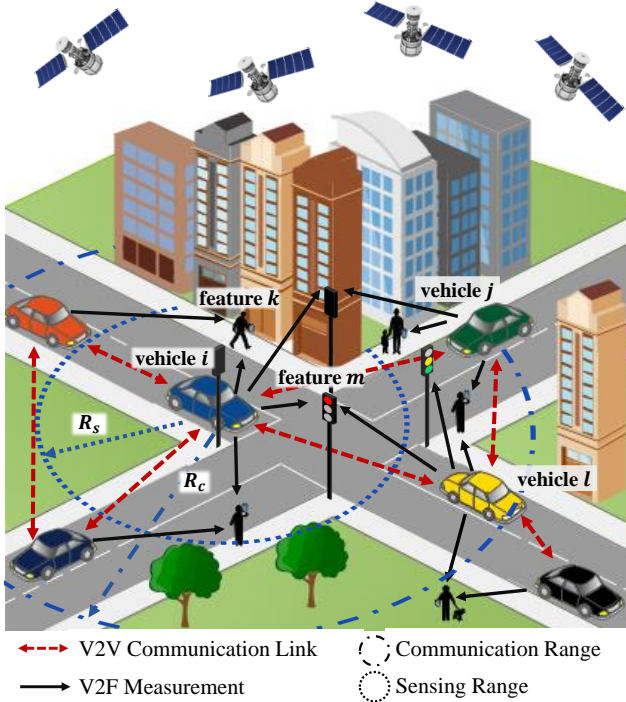


Fig. 1. Vehicular network with cooperative vehicles and non-cooperative features.

observed in these areas due to multipath and non-line-of-sight.

## II. PROBLEM FORMULATION AND SYSTEM MODEL

Consider a set of  $N_v$  interconnected vehicles  $\mathcal{V} = \{1, \dots, N_v\}$ , deployed over a two-dimensional space as exemplified in Fig. 1. Each vehicle  $i \in \mathcal{V}$  has state  $\mathbf{x}_{i,t}^{(V)} = [\mathbf{p}_{i,t}^{(V)\top}, \mathbf{v}_{i,t}^{(V)\top}]^\top \in \mathbb{R}^{4 \times 1}$ , defined as the joint set of the position  $\mathbf{p}_{i,t}^{(V)} = [p_{xi,t}, p_{yi,t}]^\top \in \mathbb{R}^{2 \times 1}$  and the velocity  $\mathbf{v}_{i,t}^{(V)} = [v_{xi,t}, v_{yi,t}]^\top \in \mathbb{R}^{2 \times 1}$ . The state evolves over the time  $t$  according to the dynamic model [29]:

$$\mathbf{x}_{i,t}^{(V)} = \mathbf{A}\mathbf{x}_{i,t-1}^{(V)} + \mathbf{B}\mathbf{a}_{i,t-1} + \mathbf{w}_{i,t-1}^{(V)}, \quad (1)$$

with

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_2 & T_s \mathbf{I}_2 \\ \mathbf{0}_{2 \times 2} & \mathbf{I}_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \frac{T_s^2}{2} \mathbf{I}_2 \\ T_s \mathbf{I}_2 \end{bmatrix}, \quad (2)$$

where  $\mathbf{A}$  denotes the transition matrix,  $\mathbf{B}$  the matrix relating the vehicle state to the acceleration input  $\mathbf{a}_{i,t-1} \in \mathbb{R}^{2 \times 1}$  here assumed as known (e.g., from an accelerometer),  $T_s$  is the sampling interval and  $\mathbf{w}_{i,t-1}^{(V)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{i,t-1}^{(V)})$  the zero-mean Gaussian driving noise.

The vehicular network is modelled as a time-varying connected undirected graph,  $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$ , with vertices  $\mathcal{V}$  representing the vehicles and edges  $\mathcal{E}_t$  the V2V communication links. Assuming the communication range equal to  $R_c$  at each vehicle, the edge set is  $\mathcal{E}_t = \{(i, j) \in \mathcal{V} \times \mathcal{V} : \|\mathbf{p}_{i,t}^{(V)} - \mathbf{p}_{j,t}^{(V)}\| \leq R_c\}$ , with vehicles  $i$  and  $j$  connected if and only if their distance is lower or equal to  $R_c$  (as illustrated by dashed red links in Fig. 1) and with  $\|\cdot\|$  denoting the Frobenius norm. The set of neighbors that directly communicates with vehicle  $i \in \mathcal{V}$  is denoted as  $\mathcal{J}_{i,t} = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}_t\}$ , its cardinality (i.e.,

the vehicle degree) as  $d_{i,t} = |\mathcal{J}_{i,t}|$  and the maximum degree (over all vehicles) as  $\Delta_t = \max d_{i,t}$ .

As illustrated in Fig. 1, the scenario involves also a set  $\mathcal{F} = \{1, \dots, N_f\}$  of  $N_f$  features (e.g., people, traffic lights, trees, etc.), either static or mobile, which are non-cooperative entities sensed by the vehicle's on-board equipment (e.g., RADAR, LIDAR, camera-based detector, etc.) and used as common noisy reference points for cooperative localization. Note that features are defined as non-cooperative passive objects since they cannot communicate with each other, or with vehicles, and do not perform any computations and measurements. At time  $t$ , the state of feature  $k$  is  $\mathbf{x}_{k,t}^{(F)} = [\mathbf{p}_{k,t}^{(F)\top}, \mathbf{v}_{k,t}^{(F)\top}]^\top \in \mathbb{R}^{4 \times 1}$ , defined as the joint set of the position  $\mathbf{p}_{k,t}^{(F)} = [p_{xk,t}, p_{yk,t}]^\top \in \mathbb{R}^{2 \times 1}$  and the velocity  $\mathbf{v}_{k,t}^{(F)} = [v_{xk,t}, v_{yk,t}]^\top \in \mathbb{R}^{2 \times 1}$ . The feature state evolves according to the first order Markov model:

$$\mathbf{x}_{k,t}^{(F)} = \mathbf{A}\mathbf{x}_{k,t-1}^{(F)} + \mathbf{w}_{k,t-1}^{(F)}, \quad (3)$$

with zero-mean Gaussian driving noise  $\mathbf{w}_{k,t-1}^{(F)} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k,t-1}^{(F)})$ . Due to the limited sensing range  $R_s$ , each vehicle  $i$  senses only a subset of all features given by  $\mathcal{F}_{i,t} = \{k \in \mathcal{F} : \|\mathbf{p}_{k,t}^{(F)} - \mathbf{p}_{i,t}^{(V)}\| \leq R_s\} \subseteq \mathcal{F}$  (see Fig. 1). Note that in this analysis the focus is on passive objects, but the framework can be easily generalized to handle also non-cooperative active features, equipped with transmitting devices or even combined with a subset of active cooperative units, sometimes denoted as beaconing anchors.

At time instant  $t$ , the measurements available at each vehicle for localization are the GNSS location fix and the V2F relative location observations gathered for all the features within the sensing range. The GNSS measurement of vehicle  $i$  state is:

$$\mathbf{p}_{i,t}^{(\text{GNSS})} = \mathbf{p}_{i,t}^{(V)} + \mathbf{n}_{i,t}^{(\text{GNSS})} = \mathbf{P}\mathbf{x}_{i,t}^{(V)} + \mathbf{n}_{i,t}^{(\text{GNSS})}, \quad (4)$$

where  $\mathbf{P} = [\mathbf{I}_2 \ \mathbf{0}_{2 \times 2}]$  is the  $2 \times 4$  matrix selecting the vehicle position, and  $\mathbf{n}_{i,t}^{(\text{GNSS})} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{i,t}^{(\text{GNSS})})$  is the GNSS measurement error [10, 13]. The relative location measurement  $\mathbf{p}_{i,k,t}^{(\text{V2F})} \in \mathbb{R}^{2 \times 1}$  made by vehicle  $i$  to its nearby feature  $k \in \mathcal{F}_{i,t}$  is modelled as:

$$\mathbf{p}_{i,k,t}^{(\text{V2F})} = q(\mathbf{P}\mathbf{x}_{k,t}^{(F)} - \mathbf{P}\mathbf{x}_{i,t}^{(V)}) + \mathbf{n}_{i,k,t}^{(\text{V2F})} = q(\boldsymbol{\delta}_{i,k,t}) + \mathbf{n}_{i,k,t}^{(\text{V2F})}, \quad (5)$$

where  $\mathbf{n}_{i,k,t}^{(\text{V2F})} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{i,k,t}^{(\text{V2F})})$  is the measurement uncertainty and the deterministic function  $q(\boldsymbol{\delta}_{i,k,t})$  models the relation of the observation to the V2F relative position  $\boldsymbol{\delta}_{i,k,t} = \mathbf{p}_{k,t}^{(F)} - \mathbf{p}_{i,t}^{(V)}$ . Depending on the class of on-board sensors detecting the feature, the measurement may represent the V2F range  $q(\boldsymbol{\delta}_{i,k,t}) = \|\boldsymbol{\delta}_{i,k,t}\|$ , the angle  $q(\boldsymbol{\delta}_{i,k,t}) = \angle(\boldsymbol{\delta}_{i,k,t})$ , or the relative position  $q(\boldsymbol{\delta}_{i,k,t}) = \boldsymbol{\delta}_{i,k,t}$ . In this study, we focus on the latest case assuming that both range and angle measurements are available from on-board vehicle RADAR equipment. This choice allows a more compact information representation and deeper insight into the performance behavior (see Sec. III). Nevertheless, the framework is general enough to include other classes of measurements. The measurement errors  $\mathbf{n}_{i,t}^{(\text{GNSS})}$  and  $\mathbf{n}_{i,k,t}^{(\text{V2F})}$  are assumed to be mutually independent [22], and also independent over vehicles, features and time. Note that in

this work perfect association between V2F measurements and features is considered as available at each vehicle. A discussion on the association problem can be found in Sec. VI-A.

### III. CENTRALIZED ICP METHOD

In this section, we first define the overall set of GNSS and V2F measurements available at each vehicle and gathered by a fusion center for processing (see Sec. III-A). Then, we present the centralized ICP approach for the joint estimation of vehicles' and features' state information  $\theta_t = [\mathbf{x}_t^{(V)T}, \mathbf{x}_t^{(F)T}]^T \in \mathbb{R}^{4(N_v+N_f) \times 1}$ , with  $\mathbf{x}_t^{(V)} = [\mathbf{x}_{i,t}^{(V)}]_{i \in \mathcal{V}} \in \mathbb{R}^{4N_v \times 1}$  and  $\mathbf{x}_t^{(F)} = [\mathbf{x}_{k,t}^{(F)}]_{k \in \mathcal{F}} \in \mathbb{R}^{4N_f \times 1}$  being respectively the aggregate vehicle and feature states at time  $t$  (see Sec. III-B).

#### A. Data Fusion Model

Let  $\rho_t^{(\text{GNSS})} = [\rho_{i,t}^{(\text{GNSS})}]_{i \in \mathcal{V}} \in \mathbb{R}^{2N_v \times 1}$  and  $\rho_t^{(\text{V2F})} = [\rho_{i,k,t}^{(\text{V2F})}]_{i \in \mathcal{V}, k \in \mathcal{F}_{i,t}} \in \mathbb{R}^{2M \times 1}$  be the related GNSS and V2F noisy observations collected by the vehicles at time  $t$ . The V2F measurements are conveniently re-indexed as  $\rho_{m,t}^{(\text{V2F})} = \rho_{i_m,k_m,t}^{(\text{V2F})}$ , with  $m \in \mathcal{M} = \{1, \dots, M\}$  univocally identifying the measurement made by vehicle  $i_m \in \mathcal{V}$  to feature  $k_m \in \mathcal{F}_{i_m,t}$  and  $M = \sum_{i=1}^{N_v} |\mathcal{F}_{i,t}|$  denoting the total number of V2F measurements. In the centralized ICP approach, all the observations from vehicles are gathered by a fusion center to synchronously estimate the overall dynamic state  $\theta_t$ . The augmented measurement model is then:

$$\rho_t = \begin{bmatrix} \rho_t^{(\text{GNSS})} \\ \rho_t^{(\text{V2F})} \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{\mathbf{P}} & \mathbf{0}_{2N_v \times 2N_f} \\ \mathbf{M}_v & \mathbf{M}_f \end{bmatrix}}_{\mathbf{H}_t} \theta_t + \underbrace{\begin{bmatrix} \mathbf{n}_t^{(\text{GNSS})} \\ \mathbf{n}_t^{(\text{V2F})} \end{bmatrix}}_{\mathbf{n}_t}, \quad (6)$$

where  $\mathbf{H}_t$  is the matrix of the known regressors, with  $\tilde{\mathbf{P}} = \mathbf{I}_{N_v} \otimes \mathbf{P}$  and  $\otimes$  denoting the Kronecker product. The matrix  $\mathbf{M}_v = [\mathbf{M}_{m,i}]$  is block-partitioned into  $M \times N_v$  blocks of dimensions  $2 \times 4$  defined as:  $\mathbf{M}_{m,i} = -\mathbf{P}$  if  $i = i_m$ ,  $\mathbf{M}_{m,i} = \mathbf{0}$  otherwise. Similarly, the matrix  $\mathbf{M}_f = [\mathbf{M}_{m,k}]$  is block-partitioned into  $M \times N_f$  blocks of dimensions  $2 \times 4$  defined as:  $\mathbf{M}_{m,k} = \mathbf{P}$  if  $k = k_m$ ,  $\mathbf{M}_{m,k} = \mathbf{0}$  otherwise. The Gaussian vector  $\mathbf{n}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$  aggregates the GNSS and V2F measurement errors,  $\mathbf{n}_t^{(\text{GNSS})} = [\mathbf{n}_{i,t}^{(\text{GNSS})}]_{i \in \mathcal{V}}$  and  $\mathbf{n}_t^{(\text{V2F})} = [\mathbf{n}_{m,t}^{(\text{V2F})}]_{m \in \mathcal{M}}$  respectively, and it has covariance defined as  $\mathbf{R}_t = \text{blockdiag}(\mathbf{R}_t^{(\text{GNSS})}, \mathbf{R}_t^{(\text{V2F})})$ , with combined GNSS and V2F covariance matrices as  $\mathbf{R}_t^{(\text{GNSS})} = \text{blockdiag}(\mathbf{R}_{1,t}^{(\text{GNSS})}, \dots, \mathbf{R}_{N_v,t}^{(\text{GNSS})})$  and  $\mathbf{R}_t^{(\text{V2F})} = \text{blockdiag}(\mathbf{R}_{1,t}^{(\text{V2F})}, \dots, \mathbf{R}_{M,t}^{(\text{V2F})})$ , respectively. Note that the V2F sensing errors  $\mathbf{n}_{m,t}^{(\text{V2F})} = \mathbf{n}_{i_m,k_m,t}^{(\text{V2F})}$  and the related covariance matrices  $\mathbf{R}_{m,t}^{(\text{V2F})} = \mathbf{R}_{i_m,k_m,t}^{(\text{V2F})}$  are re-indexed as previously discussed according to the V2F measurement numbering.

#### B. Centralized Bayesian Tracking

Based on Bayesian filtering [30], the Minimum Mean Square Error (MMSE) estimate of  $\theta_t$ , given all measurements  $\rho_{1:t} = \{\rho_1, \dots, \rho_t\}$  up to time  $t$ , can be computed as:

$$\hat{\theta}_{t|t} = \begin{bmatrix} \hat{\mathbf{x}}_{t|t}^{(V)} \\ \hat{\mathbf{x}}_{t|t}^{(F)} \end{bmatrix} = \int \theta_t p(\theta_t | \rho_{1:t}) d\theta_t, \quad (7)$$

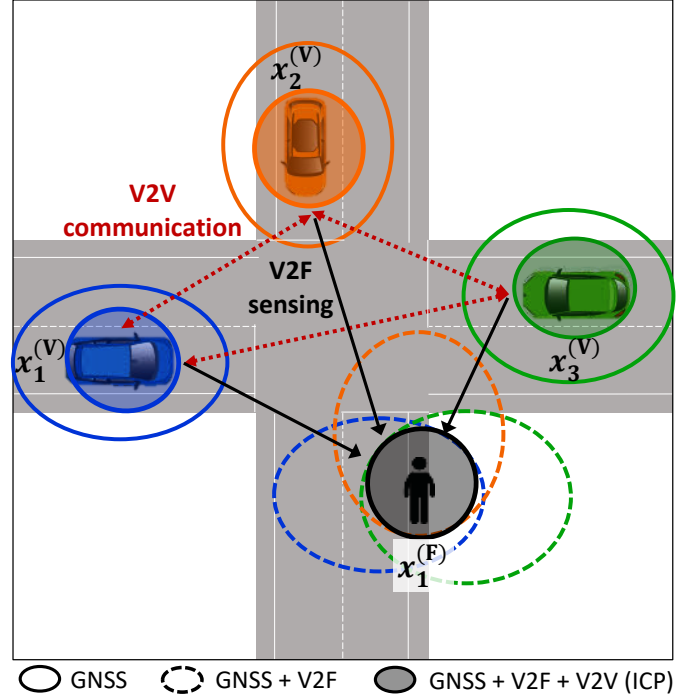


Fig. 2. ICP example with three vehicles and one passive feature. The location accuracy for the vehicles (solid colored contours) and the feature (dashed colored contours), based on stand-alone GNSS and V2F sensing, is represented as 1- $\sigma$  error ellipse. The accuracy obtained by ICP (colored ellipse) is significantly higher thanks to the cooperative localization of the jointly sensed feature through V2V links.

from the posterior probability density function (pdf):

$$p(\theta_t | \rho_{1:t}) \propto p(\rho_t | \theta_t) \int p(\theta_t | \theta_{t-1}) p(\theta_{t-1} | \rho_{1:t-1}) d\theta_{t-1}, \quad (8)$$

where  $\propto$  stands for proportionality. Here, the likelihood function is  $p(\rho_t | \theta_t) = p(\rho_t^{(\text{GNSS})} | \mathbf{x}_t^{(V)}) p(\rho_t^{(\text{V2F})} | \theta_t)$  as GNSS and V2F measurements are conditionally independent, while the transition pdf  $p(\theta_t | \theta_{t-1}) = p(\mathbf{x}_t^{(V)} | \mathbf{x}_{t-1}^{(V)}) p(\mathbf{x}_t^{(F)} | \mathbf{x}_{t-1}^{(F)})$  follows from the vehicle and feature mobility models respectively in (1) and (3). Based on these assumptions, and recalling that models are Gaussian and linear, the centralized ICP estimate reduces to the Kalman filter [30]:

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + \mathbf{C}_{t|t} \mathbf{H}_t^T \mathbf{R}_t^{-1} (\rho_t - \mathbf{H}_t \hat{\theta}_{t|t-1}), \quad (9)$$

where  $\hat{\theta}_{t|t-1}$  is the predicted state at time  $t$  for all vehicles and features, given the measurements collected up to time  $t-1$ :

$$\hat{\theta}_{t|t-1} = \begin{bmatrix} \hat{\mathbf{x}}_{t|t-1}^{(V)} \\ \hat{\mathbf{x}}_{t|t-1}^{(F)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}} \hat{\mathbf{x}}_{t-1|t-1}^{(V)} + \tilde{\mathbf{B}} \mathbf{a}_{t-1} \\ \tilde{\mathbf{A}} \hat{\mathbf{x}}_{t-1|t-1}^{(F)} \end{bmatrix}, \quad (10)$$

with  $\tilde{\mathbf{A}} = \mathbf{I}_{N_v} \otimes \mathbf{A}$ ,  $\tilde{\mathbf{B}} = \mathbf{I}_{N_v} \otimes \mathbf{B}$ ,  $\mathbf{a}_{t-1} = [\mathbf{a}_{i,t-1}]_{i \in \mathcal{V}}$  according to (1) and  $\tilde{\mathbf{A}} = \mathbf{I}_{N_f} \otimes \mathbf{A}$  based on (3). The covariance of the centralized ICP estimate is:

$$\mathbf{C}_{t|t} = \text{Cov}(\hat{\theta}_{t|t}) = (\mathbf{C}_{t|t-1}^{-1} + \mathbf{H}_t^T \mathbf{R}_t^{-1} \mathbf{H}_t)^{-1}, \quad (12)$$

where  $\mathbf{C}_{t|t-1}$  is the covariance of the prediction:

$$\mathbf{C}_{t|t-1} = \text{Cov}(\hat{\theta}_{t|t-1}) = \text{blockdiag}(\mathbf{C}_{t|t-1}^{(V)}, \mathbf{C}_{t|t-1}^{(F)}), \quad (13)$$

with sub-matrix  $\mathbf{C}_{t|t-1}^{(V)} = \text{blockdiag}(\mathbf{C}_{1,t|t-1}^{(V)}, \dots, \mathbf{C}_{N_v,t|t-1}^{(V)})$  collecting the prior covariances  $\mathbf{C}_{i,t|t-1}^{(V)} = \mathbf{A}\mathbf{C}_{i,t-1|t-1}^{(V)}\mathbf{A}^\top + \mathbf{Q}_{i,t-1}^{(V)}$  for all vehicles  $i \in \mathcal{V}$ , and sub-matrix  $\mathbf{C}_{t|t-1}^{(F)} = \text{blockdiag}(\mathbf{C}_{1,t|t-1}^{(F)}, \dots, \mathbf{C}_{N_f,t|t-1}^{(F)})$  gathering the prior covariances  $\mathbf{C}_{k,t|t-1}^{(F)} = \mathbf{A}\mathbf{C}_{k,t-1|t-1}^{(F)}\mathbf{A}^\top + \mathbf{Q}_{k,t-1}^{(F)}$  for all features  $k \in \mathcal{F}$ .

#### IV. DISTRIBUTED ICP METHOD

The centralized ICP solution is not practical for large-scale networks: not only does a central computing unit constitute a single point of failure, the central solution has a computational complexity that scales cubically in the number of vehicles and features. For this reason, here we propose a distributed solution based on a combination of GMP and consensus algorithms.

The distributed method enables the sequential evaluation, at each vehicle  $i \in \mathcal{V}$ , of the marginal posterior pdfs  $p(\mathbf{x}_{i,t}^{(V)} | \boldsymbol{\rho}_{1:t})$  and  $p(\mathbf{x}_{k,t}^{(F)} | \boldsymbol{\rho}_{1:t})$ , for all features  $k \in \mathcal{F}$ . However, the GMP implementation is complicated by the fact that features are passive objects and therefore they are not actively involved in the estimation process. This means that each vehicle  $i$  has to calculate not only its own belief but also all features' beliefs using only communication with neighboring vehicles. To address this challenge, we propose a novel consensus-based GMP method that enables the cooperation between vehicles for the distributed evaluation of all features' beliefs.

An example of the proposed approach, and the related benefits, is in Fig. 2 for a scenario with one feature jointly sensed by three vehicles. The figure shows the localization accuracy drawn from the local beliefs when vehicles rely only on their own GNSS (solid colored contours) and V2F measurements (dashed colored contours). On the other hand, in the ICP approach vehicles engage in a V2V cooperative localization of the feature and reach a consensus on the feature location (black ellipse). This implicitly reflects on a significant improvement of vehicle position accuracies (colored ellipses).

##### A. Gaussian Message Passing Algorithm

The posterior pdf (8) can be computed by marginalizing the joint posterior pdf  $p(\boldsymbol{\theta}_{1:t} | \boldsymbol{\rho}_{1:t})$ , being  $\boldsymbol{\theta}_{1:t} = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t\}$  all vehicle/feature states up to time  $t$ . Taking into account the conditional independence of the measurements and of the vehicle/feature mobility models, the joint posterior pdf  $p(\boldsymbol{\theta}_{1:t} | \boldsymbol{\rho}_{1:t})$  can be factorized over vehicles and features as in (11) at bottom of the page. In order to derive the GMP, we first encode (11) as a Factor Graph (FG) [31] and then derive the Sum-Product Algorithm (SPA) message passing rules. The FG of the factorization in (11) for one time step is depicted in Fig. 3, where the state of each vehicle and feature is shown as a circle, while the factors are shown as squares.

For visualization purposes and to simplify the notation, we introduce  $h_i \triangleq p(\mathbf{x}_{i,t}^{(V)} | \mathbf{x}_{i,t-1}^{(V)})$  (i.e., vehicle state-transition pdf),  $z_k \triangleq p(\mathbf{x}_{k,t}^{(F)} | \mathbf{x}_{k,t-1}^{(F)})$  (i.e., feature state-transition pdf),  $s_i \triangleq p(\boldsymbol{\rho}_{i,t}^{(\text{GNSS})} | \mathbf{x}_{i,t}^{(V)})$  (i.e., the GNSS likelihood) and  $g_{ik} \triangleq p(\boldsymbol{\rho}_{i,k,t}^{(\text{V2F})} | \mathbf{x}_{k,t}^{(F)}, \mathbf{x}_{i,t}^{(V)})$  (i.e., the V2F measurement likelihood).

If the prior distributions of the vehicles and features are Gaussian, if all measurements and state transition models are linear in the state and have independent Gaussian noise, it can be shown that all the messages in the FG are Gaussian [32]. Hence, the SPA reverts to GMP, which has several benefits in terms of complexity and convergence [33]–[35]. In [34], authors proved that if belief propagation converges in case of loopy graphs, then the posterior marginal belief mean vector converges to the optimal centralized estimate. Note that if the linearity and Gaussianity conditions are not fulfilled, particle-based approaches can be used, though these generally incur a significant computational and communication cost.

At each time  $t$ , the GMP scheme provides approximate marginal posteriors, which are represented by the beliefs of the vehicles' and features' states, respectively  $b_{i,t}(\mathbf{x}_{i,t}^{(V)}) \approx p(\mathbf{x}_{i,t}^{(V)} | \boldsymbol{\rho}_{1:t})$  and  $b_{k,t}(\mathbf{x}_{k,t}^{(F)}) \approx p(\mathbf{x}_{k,t}^{(F)} | \boldsymbol{\rho}_{1:t})$ . Moreover, as the considered FG has cycles, the GMP algorithm becomes iterative. Hence, the beliefs of the vehicle node  $i \in \mathcal{V}$  and feature node  $k \in \mathcal{F}$  at GMP iteration  $n = 1, \dots, N_{\text{mp}}$  are:

$$b_{k,t}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto m_{z_k \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)}) \prod_{i \in \mathcal{V}_{k,t}} m_{g_{ik} \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)}), \quad (14)$$

$$b_{i,t}^{(n)}(\mathbf{x}_{i,t}^{(V)}) \propto m_{h_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)}) m_{s_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)}) \prod_{k \in \mathcal{F}_{i,t}} m_{g_{ik} \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)}), \quad (15)$$

with  $\mathcal{V}_{k,t}$  being the set of vehicles that acquire measurements of feature  $k$  (assuming that the belief of feature  $k \in \mathcal{F}$  in (14) is reset to a uniform distribution if the feature is not observed by any vehicle  $i \in \mathcal{V}$ ).

Note that the product of  $L$  Gaussian pdfs over the same vector is also Gaussian (though not normalized) [31]:

$$\prod_{\ell=1}^L \mathcal{N}(\boldsymbol{\mu}_\ell, \mathbf{C}_\ell) \propto \mathcal{N}(\tilde{\boldsymbol{\mu}}, \tilde{\mathbf{C}}), \quad (16)$$

with covariance  $\tilde{\mathbf{C}} = (\sum_{\ell=1}^L \mathbf{C}_\ell^{-1})^{-1}$  and mean  $\tilde{\boldsymbol{\mu}} = \tilde{\mathbf{C}} \cdot (\sum_{\ell=1}^L \mathbf{C}_\ell^{-1} \boldsymbol{\mu}_\ell)$ . This observation plays a key role for the evaluation of the beliefs in (14)–(15), computed as follows.

- *Feature prediction message*: The predicted state of feature  $k$  is represented by the message:

$$\begin{aligned} m_{z_k \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)}) &= \int p(\mathbf{x}_{k,t}^{(F)} | \mathbf{x}_{k,t-1}^{(F)}) b_{k,t-1}^{(N_{\text{mp}})}(\mathbf{x}_{k,t-1}^{(F)}) d\mathbf{x}_{k,t-1}^{(F)} \\ &= \mathcal{N}(\mathbf{A}\boldsymbol{\mu}_{x_{k,t-1}}^{(N_{\text{mp}})}, \mathbf{Q}_{k,t-1}^{(F)} + \mathbf{A}\mathbf{C}_{x_{k,t-1}}^{(N_{\text{mp}})}\mathbf{A}^\top), \end{aligned} \quad (17)$$

$$p(\boldsymbol{\theta}_{1:t} | \boldsymbol{\rho}_{1:t}) \propto \left( \prod_{i \in \mathcal{V}} p(\mathbf{x}_{i,0}^{(V)}) \prod_{t'=1}^t p(\mathbf{x}_{i,t'}^{(V)} | \mathbf{x}_{i,t'-1}^{(V)}) p(\boldsymbol{\rho}_{i,t'}^{(\text{GNSS})} | \mathbf{x}_{i,t'}) \prod_{k \in \mathcal{F}_{i,t}} p(\boldsymbol{\rho}_{i,k,t'}^{(\text{V2F})} | \mathbf{x}_{i,t'}, \mathbf{x}_{k,t'}) \right) \left( \prod_{k \in \mathcal{F}} p(\mathbf{x}_{k,0}^{(F)}) \prod_{t'=1}^t p(\mathbf{x}_{k,t'}^{(F)} | \mathbf{x}_{k,t'-1}^{(F)}) \right) \quad (11)$$



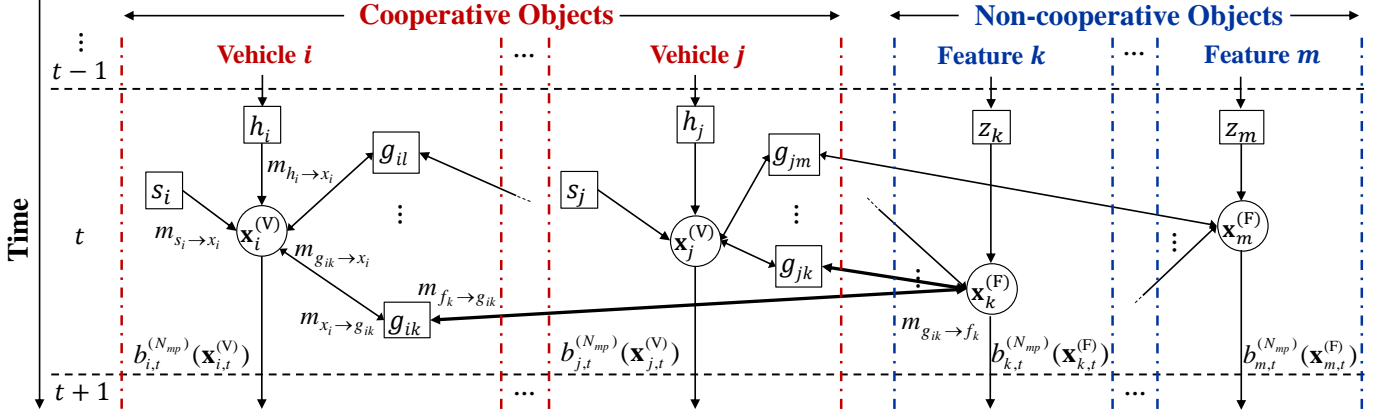


Fig. 3. FG of the joint posterior pdf (11) for one time step, showing the states of two vehicles  $i, j \in \mathcal{V}$  and two features  $k, m \in \mathcal{F}$ . Vehicles (delimited by red dashed-dot lines) are active objects that cooperate through V2V links. Features (within blue dotted lines) are passive objects, not actively involved in the GMP, that are estimated through consensus by vehicles. As V2V cooperation is implicitly performed through jointly sensed features, in the FG the subgraphs associated to different vehicles are only connected by means of V2F measurements, e.g., by the thickest black arrows connecting vehicles  $i$  and  $j$  through feature  $k$ .

in which  $\boldsymbol{\mu}_{x_k, t-1}^{(Nmp)}$  and  $\mathbf{C}_{x_k, t-1}^{(Nmp)}$  are the mean and covariance of the feature belief at previous time instant, i.e.,  $b_{k, t-1}^{(Nmp)}(\mathbf{x}_{k, t-1}^{(F)}) = \mathcal{N}(\boldsymbol{\mu}_{x_k, t-1}^{(Nmp)}, \mathbf{C}_{x_k, t-1}^{(Nmp)})$  (similar notation will be used for other beliefs and messages).

- **Vehicle prediction message:** The predicted state of vehicle  $i$  is represented by the message:

$$\begin{aligned} m_{h_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)}) &= \int p(\mathbf{x}_{i, t}^{(V)} | \mathbf{x}_{i, t-1}^{(V)}) b_{i, t-1}^{(Nmp)}(\mathbf{x}_{i, t-1}^{(V)}) d\mathbf{x}_{i, t-1}^{(V)} \\ &= \mathcal{N}(\mathbf{B}\mathbf{a}_{i, t-1} + \mathbf{A}\boldsymbol{\mu}_{x_i, t-1}^{(Nmp)}, \mathbf{Q}_{i, t-1}^{(V)} + \mathbf{A}\mathbf{C}_{x_i, t-1}^{(Nmp)}\mathbf{A}^T), \end{aligned} \quad (18)$$

in which  $\boldsymbol{\mu}_{x_i, t-1}^{(Nmp)}$  and  $\mathbf{C}_{x_i, t-1}^{(Nmp)}$  are the mean and covariance of the vehicle belief at previous time instant, i.e.,  $b_{i, t-1}^{(Nmp)}(\mathbf{x}_{i, t-1}^{(V)}) = \mathcal{N}(\boldsymbol{\mu}_{x_i, t-1}^{(Nmp)}, \mathbf{C}_{x_i, t-1}^{(Nmp)})$ .

- **GNSS message:** The message  $m_{s_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)})$  is obtained according to the  $i$ th GNSS measurement (4) and is a degenerate Gaussian with infinite variance in the velocity domain as the GNSS device provides only position information. Thereby, recalling that  $\mathbf{p}_{i, t}^{(V)} = \mathbf{P}\mathbf{x}_{i, t}^{(V)}$  and  $\mathbf{P}^\dagger = \mathbf{P}^T$ , the parameters  $\boldsymbol{\mu}_{s_i \rightarrow x_i}$  and  $\mathbf{C}_{s_i \rightarrow x_i}$  fulfill the following relations:  $\mathbf{C}_{s_i \rightarrow x_i}^{-1} = \mathbf{P}^T \mathbf{R}_{i, t}^{(GNSS)^{-1}} \mathbf{P}$  and  $\mathbf{C}_{s_i \rightarrow x_i}^{-1} \boldsymbol{\mu}_{s_i \rightarrow x_i} = \mathbf{C}_{s_i \rightarrow x_i}^{-1} \mathbf{P}^T \boldsymbol{\rho}_{i, t}^{(GNSS)}$ . These parameters will be considered for the computation of the message below according to (16).
- **Message from vehicle  $i$  to feature  $k$ :** The outgoing message at iteration  $n$  from vehicle state  $\mathbf{x}_i^{(V)}$  to factor  $g_{ik}$  is obtained as:

$$\begin{aligned} m_{x_i \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{i, t}^{(V)}) &\propto m_{h_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)}) m_{s_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)}) \\ &\quad \times \prod_{m \in \mathcal{F}_{i, t} \setminus \{k\}} m_{g_{im} \rightarrow x_i}^{(n-1)}(\mathbf{x}_{i, t}^{(V)}), \end{aligned} \quad (19)$$

in which  $m_{g_{im} \rightarrow x_i}^{(n-1)}(\mathbf{x}_{i, t}^{(V)})$  is the incoming message to vehicle computed as explained hereinafter. At BP iteration  $n = 1$ , the outgoing message is set to  $m_{x_i \rightarrow g_{ik}}^{(1)}(\mathbf{x}_{i, t}^{(V)}) \propto m_{h_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)}) m_{s_i \rightarrow x_i}(\mathbf{x}_{i, t}^{(V)})$ . The message is again Gaussian, i.e.,  $m_{x_i \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{i, t}^{(V)}) =$

$\mathcal{N}(\boldsymbol{\mu}_{x_i \rightarrow g_{ik}}^{(n)}, \mathbf{C}_{x_i \rightarrow g_{ik}}^{(n)})$ , and it is then used to obtain the incoming message from factor  $g_{ik}$  to feature state  $\mathbf{x}_{k, t}^{(F)}$  as follows:

$$m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k, t}^{(F)}) = \int p(\boldsymbol{\rho}_{i, k, t}^{(V2F)} | \mathbf{x}_{k, t}^{(F)}, \mathbf{x}_{i, t}^{(V)}) m_{x_i \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{i, t}^{(V)}) d\mathbf{x}_{i, t}^{(V)}. \quad (20)$$

Note that the above message is a degenerate density with marginal pdf in the position domain Gaussian with mean  $\boldsymbol{\rho}_{i, k, t}^{(V2F)} + \mathbf{P}\boldsymbol{\mu}_{x_i \rightarrow g_{ik}}^{(n)}$  and covariance  $\mathbf{R}_{i, k, t}^{(V2F)} + \mathbf{P}\mathbf{C}_{x_i \rightarrow g_{ik}}^{(n)}\mathbf{P}^T$ . No information is available in the velocity domain (as variance over velocity tend to infinite). Thus, only the following parameters can be calculated:  $(\mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)})^{-1} = \mathbf{P}^T (\mathbf{R}_{i, k, t}^{(V2F)} + \mathbf{P}\mathbf{C}_{x_i \rightarrow g_{ik}}^{(n)}\mathbf{P}^T)^{-1} \mathbf{P}$  and  $(\mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)})^{-1} \boldsymbol{\mu}_{g_{ik} \rightarrow x_k}^{(n)} = (\mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)})^{-1} \mathbf{P}^T (\boldsymbol{\rho}_{i, k, t}^{(V2F)} + \mathbf{P}\boldsymbol{\mu}_{x_i \rightarrow g_{ik}}^{(n)})$ . These parameters are sufficient for the evaluation of the following message according to (16).

- **Message from feature  $k$  to vehicle  $i$ :** The outgoing message from feature state  $\mathbf{x}_{k, t}^{(F)}$  to factor  $g_{ik}$  is:

$$m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k, t}^{(F)}) \propto m_{z_k \rightarrow x_k}(\mathbf{x}_{k, t}^{(F)}) \prod_{l \in \mathcal{V}_{k, t} \setminus \{i\}} m_{g_{lk} \rightarrow x_k}^{(n)}(\mathbf{x}_{k, t}^{(F)}), \quad (21)$$

in which  $m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k, t}^{(F)}) = \mathcal{N}(\boldsymbol{\mu}_{x_k \rightarrow g_{ik}}^{(n)}, \mathbf{C}_{x_k \rightarrow g_{ik}}^{(n)})$ . Note that if vehicle  $i$  is the only vehicle that observes feature  $k$ , the message is equal to the belief computed at the previous time  $t - 1$ . Now, the incoming message from factor  $g_{ik}$  to vehicle state  $\mathbf{x}_{i, t}^{(V)}$  is given by:

$$m_{g_{ik} \rightarrow x_i}^{(n)}(\mathbf{x}_{i, t}^{(V)}) = \int p(\boldsymbol{\rho}_{i, k, t}^{(V2F)} | \mathbf{x}_{k, t}^{(F)}, \mathbf{x}_{i, t}^{(V)}) m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k, t}^{(F)}) d\mathbf{x}_{k, t}^{(F)}. \quad (22)$$

Here, similarly to (20) we get that:  $(\mathbf{C}_{g_{ik} \rightarrow x_i}^{(n)})^{-1} = \mathbf{P}^T (\mathbf{R}_{i, k, t}^{(V2F)} + \mathbf{P}\mathbf{C}_{x_k \rightarrow g_{ik}}^{(n)}\mathbf{P}^T)^{-1} \mathbf{P}$  and  $(\mathbf{C}_{g_{ik} \rightarrow x_i}^{(n)})^{-1} \boldsymbol{\mu}_{g_{ik} \rightarrow x_i}^{(n)} = (\mathbf{C}_{g_{ik} \rightarrow x_i}^{(n)})^{-1} \mathbf{P}^T (-\boldsymbol{\rho}_{i, k, t}^{(V2F)} + \mathbf{P}\boldsymbol{\mu}_{x_k \rightarrow g_{ik}}^{(n)})$ .

Note that the beliefs of features and vehicles (14)–(15), as well as the outgoing messages (19) and (21), are all Gaussians with

mean and covariance that can be evaluated from the related incoming pdfs based on (16).

Distributed implementation of the above GMP scheme requires all feature and vehicle nodes to make local computations, and exchange messages with neighbors. However, features are non-cooperative passive nodes that cannot make computations, neither can they communicate with vehicles. To enable fully distributed location estimation under these conditions, in the following we propose an average consensus algorithm [36] that is nested into the GMP in such a way to allow vehicles to evaluate the features' beliefs without their cooperation, by using only V2V broadcast communications.

### B. Consensus-based Evaluation of the Feature Beliefs and Outgoing Messages

The product of measurement messages at feature  $k$

$$u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \triangleq \prod_{i \in \mathcal{V}_{k,t}} m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}), \quad (23)$$

is needed for the evaluation of the feature belief (14) and the outgoing message (21), which can be conveniently rewritten as, respectively:

$$b_{k,t}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto m_{z_k \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)}) \cdot u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}), \quad (24)$$

$$m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto m_{z_k \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)}) \cdot \frac{u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})}{m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})}. \quad (25)$$

Unfortunately, as features are passive objects, they are not actively involved in the GMP and they cannot merge the incoming messages into (23). Messages can not even be calculated by the vehicles individually; a cooperation between them is needed using V2V communication links. Cooperation enables each vehicle to evaluate all features' beliefs, even in case of no measurement between vehicle and feature.

Considering that all vehicles know the number of features in the network, the message  $u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  from (23) can be expressed as a product over all vehicles:

$$u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}) = \prod_{i \in \mathcal{V}} m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}), \quad (26)$$

where each measurement message  $m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  is defined according to (20) if  $i \in \mathcal{V}_{k,t}$ , while for  $i \notin \mathcal{V}_{k,t}$  it is set as a Gaussian pdf with covariance matrix tending to infinity.

Now, since  $u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto \mathcal{N}(\boldsymbol{\mu}_{u_{x_k}}^{(n)}, \mathbf{C}_{u_{x_k}}^{(n)})$  according to (16), both the mean  $\boldsymbol{\mu}_{u_{x_k}}^{(n)}$  and the covariance  $\mathbf{C}_{u_{x_k}}^{(n)}$  can be expressed in terms of arithmetic average. Therefore, we propose to employ the average consensus approach [36], based on successive refinements of local estimates at vehicles and information exchange between neighbors, to cooperatively determine the first two moments of  $u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$ , as detailed in the following.

- *Computation of  $\mathbf{C}_{u_{x_k}}^{(n)}$* : We introduce a consensus variable  $\Phi_{i,x_k}^{(n,r)}$  for each value of  $i, k$  and  $n$ , that is initialized at consensus iteration  $r = 0$  as:

$$\Phi_{i,x_k}^{(n,0)} = \left( \mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)} \right)^{-1}, \quad (27)$$

and subsequently updated according to the rule:

$$\Phi_{i,x_k}^{(n,r+1)} = \Phi_{i,x_k}^{(n,r)} + \epsilon \sum_{j \in \mathcal{J}_{i,t}} \left( \Phi_{j,x_k}^{(n,r)} - \Phi_{i,x_k}^{(n,r)} \right). \quad (28)$$

We recall that  $\mathcal{J}_{i,t}$  is the set of neighboring vehicles, while the step-size  $0 < \epsilon < 1/\Delta_t$  is chosen to ensure convergence [36] to the average  $1/N_v \sum_i \Phi_{i,x_k}^{(n,0)}$ . Hence, after  $N_{\text{con}}$  consensus iterations we get:

$$\Phi_{i,x_k}^{(n,N_{\text{con}})} \approx \frac{1}{N_v} \sum_{i \in \mathcal{V}} \left( \mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)} \right)^{-1}, \quad (29)$$

from which we easily find that:

$$\mathbf{C}_{u_{x_k}}^{(n)} \approx \left( N_v \Phi_{i,x_k}^{(n,N_{\text{con}})} \right)^{-1}. \quad (30)$$

- *Computation of  $\boldsymbol{\mu}_{u_{x_k}}^{(n)}$* : We again introduce a consensus variable  $\tilde{\Phi}_{i,x_k}^{(n,r)}$  for each value of  $i, k$  and  $n$ , initialized as:

$$\tilde{\Phi}_{i,x_k}^{(n,0)} = \left( \mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)} \right)^{-1} \boldsymbol{\mu}_{g_{ik} \rightarrow x_k}^{(n)}, \quad (31)$$

and refined at iteration  $r$  according to the same rule as in (28). Using the same reasoning, we find that:

$$\boldsymbol{\mu}_{u_{x_k}}^{(n)} = N_v \mathbf{C}_{u_{x_k}}^{(n)} \tilde{\Phi}_{i,x_k}^{(n,N_{\text{con}})}, \quad (32)$$

in which  $\mathbf{C}_{u_{x_k}}^{(n)}$  was obtained through (30).

Once an agreement is reached and (30), (32) are computed, each vehicle  $i$  can evaluate the approximate marginal posterior pdf of the feature  $k$  at  $n$ th GMP iteration,  $b_{k,t}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto \mathcal{N}(\boldsymbol{\mu}_{x_k,t}^{(n)}, \mathbf{C}_{x_k,t}^{(n)})$ , from (24), with mean and covariance computed from the moments of  $u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  and  $m_{z_k \rightarrow x_k}(\mathbf{x}_{k,t}^{(F)})$  according to (16). Next, the message  $m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k,t}^{(F)}) \propto \mathcal{N}(\boldsymbol{\mu}_{x_k \rightarrow g_{ik}}^{(n)}, \mathbf{C}_{x_k \rightarrow g_{ik}}^{(n)})$  is obtained from (25), where the mean and the covariance are given by:

$$\begin{aligned} \boldsymbol{\mu}_{x_k \rightarrow g_{ik}}^{(n)} &= \mathbf{C}_{x_k \rightarrow g_{ik}}^{(n)} \left( \left( \mathbf{C}_{x_k,t}^{(n)} \right)^{-1} \boldsymbol{\mu}_{x_k,t}^{(n)} - \left( \mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)} \right)^{-1} \boldsymbol{\mu}_{g_{ik} \rightarrow x_k}^{(n)} \right), \\ \mathbf{C}_{x_k \rightarrow g_{ik}}^{(n)} &= \left( \left( \mathbf{C}_{x_k,t}^{(n)} \right)^{-1} - \left( \mathbf{C}_{g_{ik} \rightarrow x_k}^{(n)} \right)^{-1} \right)^{-1}. \end{aligned} \quad (33)$$

The proposed method is summarized in Algorithm 1, where the average consensus approach is nested into the GMP scheme discussed in Sec. IV-A.

## V. ICP PERFORMANCE ANALYSIS

### A. Fundamental limits

In the following, we derive a lower bound to the cooperative localization accuracy that is reached if and only if the  $N_v$  vehicles and the  $N_f$  features are in full sensing and communication view (i.e., for all-to-all connectivity). To evaluate the vehicle positioning accuracy, we first derive the overall Fisher Information Matrix (FIM),  $\mathbf{F}$ , for the joint vehicle-feature state  $\boldsymbol{\theta}_t$  and we then compute the submatrix of the FIM inverse,  $\mathbf{C}_{t|t} = \mathbf{F}^{-1}$ , related to the single vehicle location.

---

**Algorithm 1** Consensus-based Gaussian Message Passing
 

---

```

1: At time  $t = 0$  Initialization:
2: vehicles  $i \in \mathcal{V}$  in parallel
3: initialize non-informative prior on vehicle  $p(\mathbf{x}_{i,0}^{(V)})$ 
4: initialize non-informative prior on feature  $p(\mathbf{x}_{k,0}^{(F)})$ ,  $\forall k \in \mathcal{F}$ 
5: end parallel
6: for  $t = 1 \rightarrow T$  do (time slot index)
7: vehicles  $i \in \mathcal{V}$  in parallel
8:   compute prediction messages  $m_{h_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)})$  according to
   the state-transition pdf as (18)
9:   compute the message  $m_{s_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)})$  based on GNSS
   estimated position as in (4)
10:  compute the initial outgoing message as
    $m_{x_i \rightarrow g_{ik}}^{(1)}(\mathbf{x}_{i,t}^{(V)}) = m_{h_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)}) m_{s_i \rightarrow x_i}(\mathbf{x}_{i,t}^{(V)})$ 
11: end parallel
12: for  $n = 1 \rightarrow N_{mp}$  do (GMP iteration index)
13:  vehicles  $i \in \mathcal{V}$  in parallel
14:   for  $k \in \mathcal{F}$  do
15:    evaluate feature  $k$  measurement message
     $m_{g_{ik} \rightarrow x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  according to (20)
16:    compute the measurement message product  $u_{x_k}^{(n)}(\mathbf{x}_{k,t}^{(F)})$ 
    as (30), (32) by applying consensus algorithm (28)
17:    update feature  $k$  belief  $b_{k,t}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  as (24)
18:    compute feature  $k$  outgoing message
     $m_{x_k \rightarrow g_{ik}}^{(n)}(\mathbf{x}_{k,t}^{(F)})$  as (25) by (33)
19:   end for
20:   compute vehicle  $i$  incoming message  $m_{g_{ik} \rightarrow x_i}^{(n)}(\mathbf{x}_{i,t}^{(V)})$  as (22)
21:   update vehicle  $i$  belief  $b_{i,t}^{(n)}(\mathbf{x}_{i,t}^{(V)})$  as (15)
22:   compute vehicle  $i$  outgoing message  $m_{x_i \rightarrow g_{ik}}^{(n+1)}(\mathbf{x}_{i,t}^{(V)})$  as (19)
   by applying (16)
23:  end parallel
24: end for
25: end for

```

---

The total FIM can be obtained from (12), taking into account  $\mathbf{H}_t$  and the block-diagonal structure of  $\mathbf{R}_t$ . After some algebraic manipulations, we get:

$$\mathbf{F} = \begin{bmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{G} \end{bmatrix}, \quad (34)$$

where the  $4N_v \times 4N_v$  matrix  $\mathbf{D} = \text{blockdiag}(\mathbf{D}_1, \dots, \mathbf{D}_{N_v})$  has submatrices  $\mathbf{D}_i \in \mathbb{R}^{4 \times 4}$ ,  $i \in \mathcal{V}$ , given by:

$$\mathbf{D}_i = \mathbf{C}_{i,t|t-1}^{(V)-1} + \mathbf{P}^T \left( \mathbf{R}_{i,k,t}^{(\text{GNSS})-1} + \sum_{k \in \mathcal{F}_{i,t}} \mathbf{R}_{i,k,t}^{(\text{V2F})-1} \right) \mathbf{P}. \quad (35)$$

The  $4N_v \times 4N_f$  matrix  $\mathbf{E} = [\mathbf{E}_{ik}]$  is partitioned into blocks  $\mathbf{E}_{ik} \in \mathbb{R}^{4 \times 4}$ ,  $i \in \mathcal{V}$ ,  $k \in \mathcal{F}$ , defined as:

$$\mathbf{E}_{ik} = \begin{cases} -\mathbf{P}^T \mathbf{R}_{i,k,t}^{(\text{V2F})-1} \mathbf{P}, & \text{if } k \in \mathcal{F}_{i,t} \\ \mathbf{0}, & \text{otherwise} \end{cases}. \quad (36)$$

Moreover, the  $4N_f \times 4N_f$  matrix  $\mathbf{G} = \text{blockdiag}(\mathbf{G}_1, \dots, \mathbf{G}_{N_f})$  is built from the submatrices  $\mathbf{G}_k \in \mathbb{R}^{4 \times 4}$ ,  $\forall k \in \mathcal{F}$ , such that:

$$\mathbf{G}_k = \mathbf{C}_{k,t|t-1}^{(F)-1} + \mathbf{P}^T \sum_{i \in \mathcal{V}_{k,t}} \mathbf{R}_{i,k,t}^{(\text{V2F})-1} \mathbf{P}. \quad (37)$$

To simplify the analysis, in the following we assume the prior covariance matrices for vehicle  $i$  and feature  $k$  as, respectively,  $\mathbf{C}_{i,t|t-1}^{(V)} = \text{blockdiag}(\sigma_{p,\text{pr}}^{(V)2} \mathbf{I}_2, \sigma_{v,\text{pr}}^{(V)2} \mathbf{I}_2)$  and

$\mathbf{C}_{k,t|t-1}^{(F)} = \text{blockdiag}(\sigma_{p,\text{pr}}^{(F)2} \mathbf{I}_2, \sigma_{v,\text{pr}}^{(F)2} \mathbf{I}_2)$ . In addition, we assume the GNSS and V2F measurements as i.i.d. in each subset, with covariance matrices  $\mathbf{R}_{i,t}^{(\text{GNSS})} = \sigma_{\text{GNSS}}^2 \mathbf{I}_2$  and  $\mathbf{R}_{i,k,t}^{(\text{V2F})} = \sigma_{\text{V2F}}^2 \mathbf{I}_2$ . In this case, the FIM submatrices in (34) reduce to  $\mathbf{D} = \mathbf{I}_{N_v} \otimes \text{blockdiag}(\alpha_p^{(V)} \mathbf{I}_2, \alpha_v^{(V)} \mathbf{I}_2)$ , with  $\alpha_p^{(V)} = N_f / \sigma_{\text{V2F}}^2 + 1 / \sigma_{p,\text{pr}}^{(V)2} + 1 / \sigma_{\text{GNSS}}^2$  and  $\alpha_v^{(V)} = 1 / \sigma_{v,\text{pr}}^{(V)2}$ ,  $\mathbf{G} = \mathbf{I}_{N_f} \otimes \text{blockdiag}(\alpha_p^{(F)} \mathbf{I}_2, \alpha_v^{(F)} \mathbf{I}_2)$ , with  $\alpha_p^{(F)} = N_v / \sigma_{\text{V2F}}^2 + 1 / \sigma_{p,\text{pr}}^{(F)2}$  and  $\alpha_v^{(F)} = 1 / \sigma_{v,\text{pr}}^{(F)2}$ , and  $\mathbf{E} = -1 / \sigma_{\text{V2F}}^2 \mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P}$ , where  $\mathbf{1}_{N_v \times N_f}$  is an  $N_v \times N_f$  matrix of all ones.

The equivalent FIM (EFIM) for the vehicles' states is given by the Schur complement [37]:

$$\begin{aligned} \mathbf{F}^{(V)} &= \mathbf{D} - \mathbf{E} \mathbf{G}^{-1} \mathbf{E}^T \\ &= \mathbf{D} - \beta (\mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P}) (\mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P})^T \\ &= \mathbf{I}_{N_v} \otimes \begin{bmatrix} \alpha_p^{(V)} \mathbf{I}_2 & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \alpha_v^{(V)} \mathbf{I}_2 \end{bmatrix} - \tilde{\beta} \mathbf{1}_{N_v \times N_v} \otimes \mathbf{P}^T \mathbf{P} \end{aligned} \quad (38)$$

with  $\beta = 1 / (\alpha_p^{(F)} \sigma_{\text{V2F}}^4)$ ,  $\tilde{\beta} = N_f \beta$  and where we made use of  $(\mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P}) (\mathbf{I}_{N_f} \otimes \text{blockdiag}(1 / \alpha_p^{(F)} \mathbf{I}_2, 1 / \alpha_v^{(F)} \mathbf{I}_2)) = 1 / \alpha_p^{(F)} (\mathbf{1}_{N_v \times N_v} \otimes \mathbf{P}^T \mathbf{P})$  and  $(\mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P}) (\mathbf{1}_{N_v \times N_f} \otimes \mathbf{P}^T \mathbf{P})^T = N_f (\mathbf{1}_{N_v \times N_v} \otimes \mathbf{P}^T \mathbf{P})$ . The inverse of the EFIM,  $\mathbf{C}_{t|t}^{(V)} = \mathbf{F}^{(V)-1}$ , represents the lower bound on the mean square error (MSE) matrix of the vehicles' position estimates and it is of the form:

$$\mathbf{C}_{t|t}^{(V)} = \mathbf{I}_{N_v} \otimes \begin{bmatrix} \alpha_p^{(V)-1} \mathbf{I}_2 & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \alpha_v^{(V)-1} \mathbf{I}_2 \end{bmatrix} + \eta \mathbf{1}_{N_v \times N_v} \otimes \mathbf{P}^T \mathbf{P}, \quad (39)$$

where a simple association yields to  $\eta = (1 / \alpha_p^{(V)}) / (\alpha_p^{(V)} / \tilde{\beta} - N_v)$ . Hence, the posterior covariance matrix of the position-velocity estimate for any vehicle  $i$  is:

$$\mathbf{C}_{i,t|t}^{(V)} = \begin{bmatrix} \sigma_{p,\text{post}}^{(V)2} \mathbf{I}_2 & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \sigma_{v,\text{post}}^{(V)2} \mathbf{I}_2 \end{bmatrix} = \begin{bmatrix} (\alpha_p^{(V)-1} + \eta) \mathbf{I}_2 & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \alpha_v^{(V)-1} \mathbf{I}_2 \end{bmatrix}, \quad (40)$$

where  $\alpha_p^{(V)-1}$  is the expected uncertainty if the features behaved as anchors, i.e., their locations were perfectly known and thus  $\sigma_{p,\text{pr}}^{(F)} = 0$ .

Focusing on vehicle position only, after some manipulations we get:

$$\sigma_{p,\text{post}}^{(V)2} = \frac{1}{\alpha_p^{(V)}} \left( 1 + \frac{N_f / \sigma_{\text{V2F}}^2}{N_v / \sigma_{p,\text{pr}}^{(V)2} + N_v / \sigma_{\text{GNSS}}^2 + \alpha_p^{(V)} \sigma_{\text{V2F}}^2 / \sigma_{p,\text{pr}}^{(F)2}} \right). \quad (41)$$

### B. Performance Scaling

Based on the result in (41), we consider the following limiting cases (see Appendix for derivation):

1) *Large number of vehicles*,  $N_v \rightarrow \infty$ :

$$\sigma_{p,\text{post}}^{(V)2} \rightarrow \frac{1}{\alpha_p^{(V)}} = \frac{1}{N_f / \sigma_{\text{V2F}}^2 + 1 / \sigma_{p,\text{pr}}^{(V)2} + 1 / \sigma_{\text{GNSS}}^2}, \quad (42)$$

which is the accuracy reached when all features act as anchors. Since each feature is observed by an infinite number of vehicles, its location becomes perfectly known.

2) *Large number of features*,  $N_f \rightarrow \infty$ :

$$\sigma_{p,\text{post}}^{(V)2} \rightarrow 0. \quad (43)$$



In this case all vehicles' locations become certain, as features behave like anchors, even if still 'virtual' (i.e., possibly uncertain to some extent), provided that there are many of them.

- 3) *Small V2F measurement variance*,  $\sigma_{V2F}^2 \rightarrow 0$ :

$$\sigma_{p,\text{post}}^{(V)^2} \rightarrow \frac{1}{N_v/\sigma_{p,\text{pr}}^{(V)^2} + N_v/\sigma_{\text{GNSS}}^2 + N_f/\sigma_{p,\text{pr}}^{(F)^2}}, \quad (44)$$

so the performance reaches a limiting value. When also the features' locations are perfectly known, i.e.,  $\sigma_{p,\text{pr}}^{(F)^2} \rightarrow 0$ , we get  $\sigma_{p,\text{post}}^{(V)^2} \rightarrow 0$ . It follows that good measurements *and* good prior feature information are required to have good positioning, when there are not many features, as intuitively expected.

- 4) *Small feature prior uncertainty*,  $\sigma_{p,\text{pr}}^{(F)^2} \rightarrow 0$ :

$$\sigma_{p,\text{post}}^{(V)^2} \rightarrow \frac{1}{\alpha_p^{(V)}}, \quad (45)$$

which means that features are like true anchors.

- 5) *Small vehicle prior uncertainty*,  $\sigma_{p,\text{pr}}^{(V)^2} \rightarrow 0$ , or *GNSS position uncertainty*,  $\sigma_{\text{GNSS}}^2 \rightarrow 0$ :

$$\sigma_{p,\text{post}}^{(V)^2} \rightarrow 0. \quad (46)$$

In this case cooperation is not worth, as stand-alone positioning at each vehicle is enough accurate.

## VI. IMPLEMENTATION ASPECTS

In this section, we comment on implementation aspects related to data association, V2V communication, complexity, measurement synchronization and V2V latency.

### A. Cooperative Data Association

Some form of data association is required for the implementation of the proposed cooperative localization approach. In particular, each vehicle needs to associate its own V2F measurements, gathered by the on-board sensing devices, to the corresponding passive features to be able to localize and track all the features in view. Several approaches are available in the multi-target tracking literature [38]–[40], accounting for the arrival of new features and the removal of features no longer in view. In addition to per-vehicle data association, vehicles must agree on a common set of features, to coherently fuse their data about the features' locations. Approaches for this cooperative data association problem, namely multi-sensor multi-target tracking (MTT) methods, exist as well [41]–[43]. In particular, in [42] the MTT algorithm is able to track targets (features) sensed by multiple sensors (vehicles) by performing an ordered estimation of target states and association variables using belief propagation. This method could be incorporated in the proposed ICP algorithm, but taking into account that in our scenario all processing is distributed over the vehicular network, and that both features and vehicles have unknown positions.

Feature selection is an important issue to maximize the localization performance. In particular, the set of features to be included in the ICP processing should be selected by considering that: (a) quasi-static features have to be preferred, as they

can be estimated with higher accuracy and thus provide more benefits to vehicle localization; (b) well separated (in either distance, angle or velocity) features have to be preferred, as they are easier to associate among vehicles; (c) features should be classified (e.g., as poles, pedestrians, bicycles, cars) to ease the data association, both locally and among vehicles. With proper feature selection, a reliable data association should be viable, at a cost of reducing the number of available features.

In this work we do not explicitly treat the data association problem, but we rather assume that the local sensors can provide unique semantic labels for each detected feature (e.g., for a camera sensor, this label could be of the form "person with green jacket and blue trousers"), based on which the proposed positioning algorithm can be performed. In that sense, our algorithm provides a lower bound on the location error for a more practical algorithm with data association. This creates several challenges that must be addressed, but are outside the scope of this paper.

### B. Communication Overhead

Irrespective of the form of data association, the proposed localization method requires significant communication between vehicles, as discussed below.

- 1) *Cooperative data association*: During this phase, vehicles decide on which local feature identifier corresponds to local feature identifiers of other vehicles. Each vehicle can thus maintain a list of vehicles for each feature and a list of features that it shares with other vehicles. Such lists remove the need for all vehicles to keep track of all features. Consensus-based methods can be applied [43].
- 2) *BP iterations*: Once each vehicle has knowledge of features and associated vehicles that agree on the same features, the BP iterations commence. Each BP iteration, as shown in Algorithm 1, mainly consists of consensus iterations. During each consensus iteration, each vehicle broadcasts feature-related information. Each broadcast would comprise transmitter ID, transmitter belief, feature identifier per feature, feature belief per feature.

From the above discussion, it is clear that the total number of broadcasts per vehicle is dominated by the consensus, and thus scales as  $\mathcal{O}(N_{\text{mp}}N_{\text{con}})$ , where  $N_{\text{mp}}$  denotes the number of BP iterations and  $N_{\text{con}}$  the number of consensus iterations per BP iteration.

Considering a data rate of  $R$  bits/s, the time required for communication is lower bounded by:

$$T_s \geq N_{\text{mp}}N_{\text{con}}N_fN_{\text{nei}}N_b/R, \quad (47)$$

where  $N_{\text{nei}}$  is the number of neighboring vehicles and  $N_b$  is the number of bits needed to describe the belief of a feature. As an example, in the case of using the IEEE 802.11p V2V standard, with  $R = 6$  Mbit/s,  $N_{\text{nei}} = 10$  neighbors,  $N_f = 20$  features,  $N_b = 100$  bits, and  $T_s = 1$  s, we find that  $N_{\text{mp}}N_{\text{con}} \leq 300$ , which is a reasonable number, as we will see during the performance evaluation.

*Remark*: To reduce the communication overhead and delay, the value of  $N_{\text{mp}}$  can be made adaptive. In our case, at vehicle  $i$ , we stop the GMP iterations when  $\|\mu_{x_{i,t}}^{(n+1)} - \mu_{x_{i,t}}^{(n)}\| < \gamma_{\text{mp}}$

and  $\|\mathbf{C}_{x_{i,t}}^{(n+1)} - \mathbf{C}_{x_{i,t}}^{(n)}\|^{1/2} < \gamma_{\text{mp}}$ ,  $\forall i \in \mathcal{V}$ , with  $\gamma_{\text{mp}}$  being a threshold and  $\boldsymbol{\mu}_{x_{i,t}}^{(n)}$  and  $\mathbf{C}_{x_{i,t}}^{(n)}$  being respectively the mean and the covariance of the  $i$ th vehicle belief  $b_{i,t}^{(n)}(\mathbf{x}_{i,t}^{(V)})$ . Similarly, the value of  $N_{\text{con}}$  can be made adaptive. In our case  $\|\tilde{\boldsymbol{\Phi}}_{i,x_k}^{(n,r+1)} - \tilde{\boldsymbol{\Phi}}_{i,x_k}^{(n,r)}\| < \gamma_{\text{con}}$  and  $\|\boldsymbol{\Phi}_{i,x_k}^{(n,r+1)} - \boldsymbol{\Phi}_{i,x_k}^{(n,r)}\|^{1/2} < \gamma_{\text{con}}$ ,  $\forall i \in \mathcal{V}$  and  $\forall k \in \mathcal{F}$ , with a threshold  $\gamma_{\text{con}}$  and  $\tilde{\boldsymbol{\Phi}}_{i,x_k}^{(n,r)}$  and  $\boldsymbol{\Phi}_{i,x_k}^{(n,r)}$  being respectively the variables used to determine the first two moments of  $u_{x_k}^{(n)}(\mathbf{x}_k^{(F)})$  (26), needed for the evaluation of the  $k$ th feature belief  $b_{k,t}^{(n)}(\mathbf{x}_k^{(F)})$ .

### C. Computational Complexity

Ignoring the complexity of the data association, the per-vehicle computational complexity of the proposed method is relatively modest in comparison with the centralized approach from Section III. In particular, the consensus iterations require only additions of vectors, which scales linearly in the number of features. In addition each vehicle must invert  $2N_f + 1$  covariance matrices of dimension  $4 \times 4$ , so that the total complexity per time slot scales as  $\mathcal{O}(N_v N_{\text{mp}} (N_{\text{con}} N_f + 4^3 2N_f))$ . In contrast, the complexity of the centralized approach is dominated by the inversion of covariance matrices, with a total complexity per time slot scaling as  $\mathcal{O}((4N_v + 4N_f)^3)$ .

### D. Measurement Synchronization and V2V latency

The proposed algorithm is built on the assumption that all measurements are taken and processed synchronously. However, rarely GNSS and V2F measurements at a given vehicle on the one hand, and across different vehicles on the other hand, will be isochronous. In addition, other factors in the VANET such as V2V communication latency, difference between sensing and transmission times, uncontrolled and/or irregular successive transmission times<sup>1</sup>, and even processing time, will further exacerbate this problem. Temporal re-alignment of measurements at each ICP step is mandatory and could be handled as follows. Assuming the GNSS clocks are well synchronized, each vehicle can adjust its local measurements to its transmission time. This is done by applying a prediction step (typically, in the Kalman filter sense, while adding local process noise), based on the feature/vehicle dynamic models, with a time horizon equal to the time elapsed between sensing and transmission. When the message is then transmitted by the vehicle and received by a neighbor, the recipient can adjust its own local measurements to the receive time (or alternatively, both its local measurements and the received data to any arbitrary forward point in time) so as to perform fusion with all synchronized measurements. A similar procedure can be applied to consensus iterations. Additional techniques are available in the literature to deal with lack of synchronization, such as out-of-order measurements (e.g., [14]).

<sup>1</sup>Depending on the underlying V2X communication standard, transmission times may be event-driven (e.g., in case of sudden emergency break or acceleration) and/or subject to drastic channel congestion control (e.g., their rate may be contextually adjusted).

Considering the latency requirement of safety-critical ITS applications and the repeated data exchange needed for cooperation, ICP – as any other cooperative perception system – is expected to rely on very-low end-to-end communication latency. Current V2X standards (i.e., IEEE 802.11p/WAVE and ETSI ITS-G5) provide latency ranging from 100 ms down to 20 ms in the most demanding use cases, but a latency of 1-10 ms is envisioned for new-generation V2X systems supporting connected automated driving [44], as well as within foreseen 5G modes possibly devoted to vehicular communications. A very-low latency is mandatory also to guarantee a sub-meter location accuracy, as fusion of outdated measurements (and more generally, their lack of spatial coherence over time) can easily reduce the positioning performance (e.g., 1-10 ms corresponds to 2.2-22 cm for vehicles traveling at 80 km/h).

Assuming that ICP can rely on very-low V2V latency and on a temporal re-alignment of measurements as above discussed, beliefs can be re-synchronized to a common time horizon at each message exchange, at the price of an increasing variance. Clearly, as the time difference between the measurements are taken or combined increases, data fusion becomes less reliable, and the accuracy of ICP may start dropping. This imposes to bound the total time required for V2V communication and processing.

## VII. PERFORMANCE EVALUATION

In this section the ICP performance is assessed in two different scenarios. A cross-road area is first simulated in Sec. VII-A with static features and heterogeneous positioning conditions in terms of feature density and V2V connectivity. This scenario is used to investigate the ICP accuracy for varying number of features and vehicles, and also to validate the analytical bound derived in Sec. V. A more complex scenario is then introduced in Sec. VII-B, where vehicular and pedestrian traffics are simulated over a real urban map using the SUMO simulator. This use-case is considered to validate the ICP method in more realistic traffic conditions, with mobile features and vehicles using different types of GNSS device with significant diversity of location accuracy.

### A. Simulated crossroad scenario with mixed rural/urban areas

**Settings.** We first consider the crossroad scenario in Fig. 4 bottom-left map, where the total length of each road is 1.5 km and the center of the intersection is at position  $\mathbf{c} = [750 \text{ m}, 750 \text{ m}]^T$ . Lane and sidewalk widths are respectively set to 3 m and 1.3 m. As illustrated in the three different time frames at the bottom of Fig. 4, the scenario involves  $N_v$  vehicles, grouped in four clusters of  $N_v/4$  vehicles each, that enter at time  $t = 0$  from the four corners of the area, drive straight ahead along their respective lanes and exit on the opposite sides, crossing in the middle. Each vehicle drives through three different areas: a rural area (first road section of 300 m), urban canyon (the central section of 900 m) and again a rural area (last 300 m). Since vehicles need some time to enter/leave different areas, there is a transitory interval in which different vehicles are in different areas, with duration that depends on the specific parameter settings.

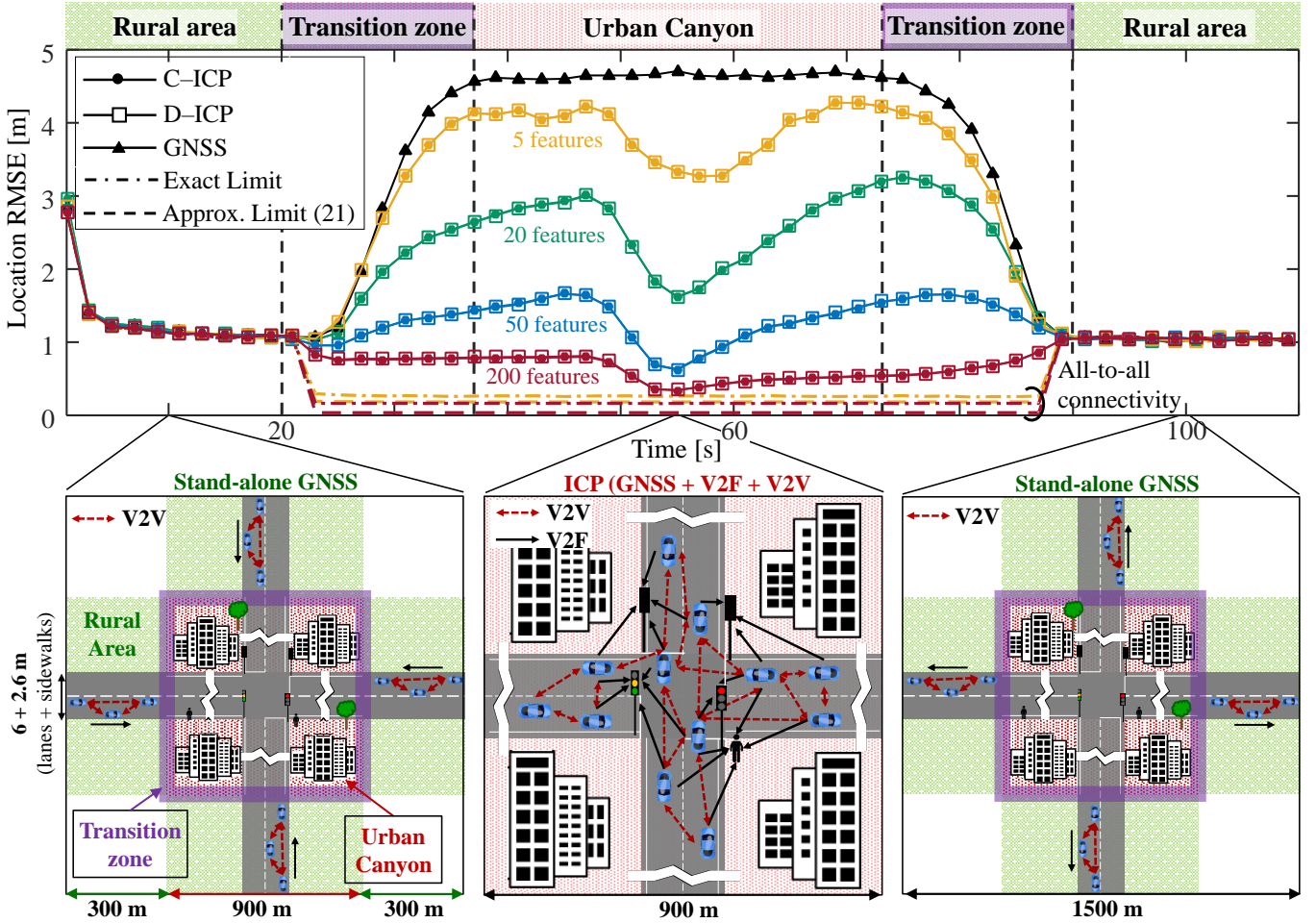


Fig. 4. ICP performance in a crossroad scenario with  $N_v = 12$  vehicles driving through a 1.5 km  $\times$  1.5 km area. The scenario is pictured in the bottom figures at time instant  $t = 10$  s (vehicles are driving through the rural area towards the crossroad),  $t = 55$  s (vehicles are crossing the urban canyon) and  $t = 100$  s (vehicles are back in the rural area). In the rural area only GNSS is available, while in the urban canyon  $N_f \in \{5, 20, 50, 200\}$  features are jointly sensed by the ICP-enabled vehicles to augment the GNSS performance. The violet box highlights the transition from one area to the other. In the top figure, the ICP accuracy versus time is compared with stand-alone GNSS and with the lower bound for all-to-all V2V/V2F connectivity.

In terms of vehicle dynamics, for each vehicle we set the initial velocity to  $\mathbf{v}_0^{(V)} = \mathbf{0}$  km/h. The mean acceleration  $\mathbf{a}_{i,t}$  in (1) is initialized to 1.4 m/s<sup>2</sup> in the driving direction at  $t = 0$  and kept constant until the vehicle reaches a velocity of 50 km/h, then it is set to 0 m/s<sup>2</sup> (i.e., the average driving velocity is 50 km/h). Since vehicles move along roads, the acceleration uncertainty in the direction of road,  $\sigma_{a_{i,\parallel}} = 0.3$  m/s<sup>2</sup>, is assumed to be greater than the one in the orthogonal direction,  $\sigma_{a_{i,\perp}} = 10^{-4}$  m/s<sup>2</sup>. Thus, depending on the driving direction of the vehicle, the acceleration uncertainties along  $x$  and  $y$  axes, respectively  $\sigma_{a_{xi}}$  and  $\sigma_{a_{yi}}$ , are defined. The sampling time is  $T_s = 1$  s. The process noise covariance matrix from (1) is set as  $\mathbf{Q}_{i,t-1}^{(V)} = \mathbf{B}\mathbf{Q}_{i,t-1}^{(V)}\mathbf{B}^T$ , with:

$$\tilde{\mathbf{Q}}_{i,t-1}^{(V)} = \begin{bmatrix} \sigma_{a_{xi}}^2 & 0 \\ 0 & \sigma_{a_{yi}}^2 \end{bmatrix}. \quad (48)$$

For positioning, in the rural area vehicles rely solely on GNSS, while in the urban canyon they can also use features, which are randomly deployed over the area. In this scenario, features are assumed to be static, thus their mobility model in (3) reduces to  $\mathbf{x}_{k,t}^{(F)} = \mathbf{A}\mathbf{x}_{k,t-1}^{(F)} = \mathbf{x}_{k,t-1}^{(F)}$  as  $\mathbf{v}_{k,t}^{(F)} =$

$\mathbf{v}_{k,t-1}^{(F)} = \mathbf{0}$  km/h. Note that in all the simulated methods, vehicle dynamics are incorporated using either a Kalman filter or the GMP.

The GNSS measurement covariance matrix at each vehicle is  $\mathbf{R}_{i,t}^{(\text{GNSS})} = \sigma_{\text{GNSS}}^2 \mathbf{I}_2$ , with  $\sigma_{\text{GNSS}} = 2$  m in the rural area and  $\sigma_{\text{GNSS}} = 15$  m in the urban canyon. The V2F measurement covariance matrix is  $\mathbf{R}_{i,k,t}^{(\text{V2F})} = \sigma_{\text{V2F}}^2 \mathbf{I}_2$ , with  $\sigma_{\text{V2F}} = 0.5$  m. Finally, the communication and sensing ranges at each vehicle are set to  $R_c = 150$  m and  $R_s = 50$  m, respectively. The consensus step-size parameter is set to  $\epsilon = 0.99/\Delta_t$ , while the threshold on the GMP and consensus convergence are set to  $\gamma_{\text{mp}} = \gamma_{\text{con}} = 10^{-2}$ .

In the following, the positioning performance is evaluated through Monte Carlo simulations, in terms of (i) the root mean square error (RMSE) of the position estimate and (ii) the delay of the fix delivery (measured in terms of the number of GMP iterations  $N_{\text{mp}}$  and consensus iterations  $N_{\text{con}}$ ). Three methods are compared, namely the stand-alone GNSS, the centralized and distributed versions of the proposed ICP method.

**Numerical results.** We first investigate the performance for a fixed number of vehicles  $N_v = 12$  and a varying number of

features. Fig. 4 shows the position RMSE of the vehicles as a function of time, for the three positioning methods. Snapshots of the V2V/V2F connectivity are shown at the bottom for time instants  $t = 10$  s (when vehicles are driving through the rural area towards the crossroad),  $t = 55$  s (in the urban canyon) and  $t = 100$  s (back to the rural area). Note that the exponential decay of the RMSE in the first few seconds of simulation results is due to transient effects. When vehicles use stand-alone GNSS, a severe performance degradation is observed as soon as vehicles enter the transition zone. The proposed algorithm can counter this degradation, especially when many features are available. The centralized and distributed ICP methods, namely C-ICP and D-ICP, lead to nearly identical performance, indicating that the proposed solution does not suffer from cycles in the FG. Moreover, assuming all-to-all V2V and V2F connectivity, the exact lower bound (dashed-dot line), obtained from (12) for  $\mathcal{F}_{i,t} = \mathcal{F}$  and  $\mathcal{V}_{k,t} = \mathcal{V}$ ,  $\forall t$ , and the approximated one (dashed line), from (41), are evaluated for  $N_f \in \{5, 200\}$ . For the latter limit, variances are computed by approximating the prior/measurement covariances as diagonal matrices with entries determined as sample averages over the two spatial dimensions (for both vehicles and features). It can be seen that when the connectivity is high, a moderate number of features and vehicles (respectively, 5 and 12) is enough to obtain a centimeter-level accuracy. As predicted from the theoretical analysis in Sec. III, if the number of features is high, e.g.,  $N_f = 200$ , the vehicle location accuracy tends to zero (see (43)).

We now evaluate, for  $N_f \in \{20, 200\}$ , the impact of the number of vehicles. In Fig. 5, the RMSE of the vehicles' position estimate is shown versus time for  $N_v \in \{5, 12, 32\}$  vehicles. It is clear that for a fixed number of features, more vehicles bring clear benefits in terms of positioning accuracy. In both Fig. 4 and Fig. 5, we note an RMSE valley around  $t = 55$  s. This is when most vehicles are in the urban canyon and there is high connectivity with many visible features.

A detailed analysis of the connectivity is provided in Fig. 6. The V2V connectivity (i.e., average number of neighbors at each vehicle) is shown in Fig. 6-(a) and the V2F connectivity (i.e., average number of visible features at each vehicle) is in Fig. 6-(b), for a scenario with  $N_f \in \{20, 200\}$  features and  $N_v \in \{5, 12, 32\}$  vehicles. We observe that since vehicles all start from the rural area and drive towards the urban canyon, there is a high V2V connectivity between 45 s and 65 s, which explains the behavior seen in Fig. 5. As highlighted at the bottom of Fig. 4, at the beginning and at the end of the observation time, there are four subgraphs (one per incoming road) since  $R_c = 150$  m and vehicles are 450 m far from the intersection when they enter in the urban canyon area. On the other hand, in the proximity of the intersection all subgraphs merge into a single graph. The connectivity grows rapidly with the number of vehicles. For this connectivity to be useful, also the number of shared visible features needs to be sufficiently high. In Fig. 6-(b), we observe that this is again the case between 45 s and 65 s, due to the combination of two phenomena: a large number of connected vehicles and a large number of jointly observed features. Both are needed for the proposed algorithm to work well, as confirmed

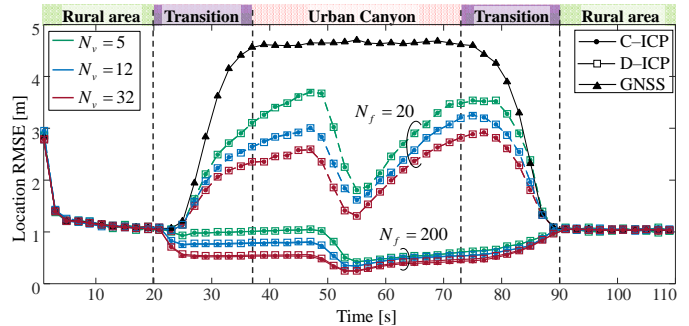


Fig. 5. RMSE of the vehicle position estimate versus time for the crossroad scenario in Fig. 4,  $N_v \in \{5, 12, 32\}$  vehicles and  $N_f \in \{20, 200\}$  features. The performance of the proposed distributed ICP algorithm is compared with both the stand-alone GNSS and centralized ICP approaches.

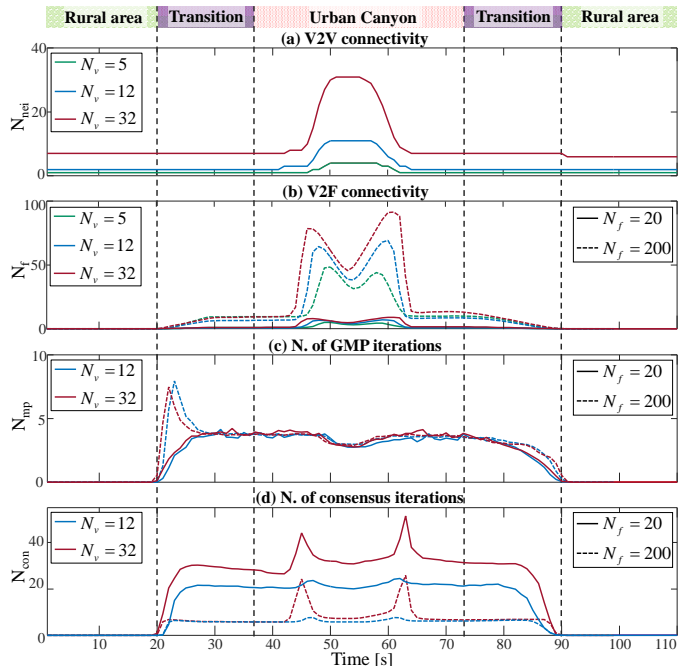


Fig. 6. Graph connectivity and number of iterations versus time for the ICP algorithm in the crossroad scenario of Fig. 4, with  $N_v \in \{5, 12, 32\}$  vehicles and  $N_f \in \{20, 200\}$  features: (a) V2V connectivity, (b) V2F connectivity, (c) number of GMP iteration and (d) number of consensus iterations.

by Figs. 4–5. While high V2V and V2F connectivity are desirable, they come at a cost in delay. For the scenario with  $N_v \in \{12, 32\}$  vehicles and  $N_f \in \{20, 200\}$  features, Figs. 6-(c) and (d) illustrate the number of GMP iterations  $N_{mp}$  and consensus iterations  $N_{con}$  versus time, respectively. We observe that the number of GMP iterations rises rapidly when the vehicles enter the transition area, especially for a larger number of features, and remains roughly constant until they enter the second transition area. It is interesting to note that  $N_{mp}$  is relatively insensitive to the number of vehicles and features. While  $N_{mp}$  remains below 10 for all considered scenarios in Fig. 6-(c),  $N_{con}$  is generally larger (see Fig. 6-(d)). Moreover, it can be noticed that the number of consensus iterations increases around time instants 45 s and 65 s, i.e., respectively when the four subgraphs are fused into a single graph and when the single graph splits in four subgraphs, due



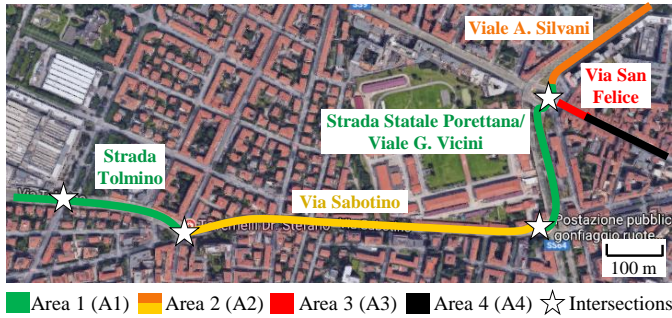


Fig. 7. Map of the Bologna scenario, Italy, with vehicles and pedestrians are simulated over a  $1.2 \text{ km} \times 0.5 \text{ km}$  area. Each street is associated with a different GNSS signal quality (see Table I). Map courtesy of Google Maps.

to the low connectivity between vehicles at those time instants. In contrast to the GMP iterations, the number of consensus iterations increases with the number of vehicles, but decreases with the number of features. In fact, consensus convergence rate depends on the graph connectivity which is related to the number of features that connect single vehicles' subgraphs (see bold connections in Fig. 3). The results from Fig. 6 can be used to evaluate the communication overhead of the proposed distributed algorithm through (47).

### B. Real urban scenario with SUMO-simulated traffic

**Settings.** To assess the ICP performance in a more realistic environment, we use the traffic simulator SUMO [45], which uses real city maps to generate synthetic traces of vehicles and pedestrians. For this experiment, we consider vehicles and pedestrians, constrained to the highlighted streets, in a urban area of size  $1.2 \text{ km} \times 0.5 \text{ km}$  in the city of Bologna, Italy (see Fig. 7). In particular, we generate 10 vehicle (thick black line) and 20 pedestrian (colored line) trajectories (as shown in Fig. 8) with sampling period  $T_s = 1 \text{ s}$ . The traces of the vehicles are synthesized according to a ‘‘Krauss car-following’’ model, with maximum speed of  $14 \text{ m/s}$  (around  $50 \text{ km/h}$ ), while the traces of the pedestrians are generated with an ‘‘inter-trip chain’’ model which includes multi-modal profiles (e.g., purely static, queuing while entering a bus, walking on the sidewalk, suddenly turning to adjacent streets). The maximum pedestrian speed is set to  $1.4 \text{ m/s}$  (about  $5 \text{ km/h}$ ) in our simulation. Each vehicle is assumed to equip a GNSS receiver and, in order to account for the wide diversity in the market, we assume four types of GNSS receivers [46]. Three vehicles are assigned a Standard Positioning Service (SPS) receiver whose position estimates have a standard deviation of  $\bar{\sigma}_{\text{GNSS}} = 3.6 \text{ m}$ , three other vehicles a Satellite-Based Augmentation Systems (SBAS) receiver with  $\bar{\sigma}_{\text{GNSS}} = 1.44 \text{ m}$ , two vehicles a Differential GNSS (DGNSS) receiver with  $\bar{\sigma}_{\text{GNSS}} = 40 \text{ cm}$  and the last two vehicles a RTK receiver with  $\bar{\sigma}_{\text{GNSS}} = 1 \text{ cm}$ . Moreover, since the GNSS accuracy is also sensitive to the surrounding environment, we model four types of environments which affect the quality of the GNSS differently, as shown in Table I. The third column of the table indicates how much the standard deviation of GNSS measurements is incremented with respect to their nominal value

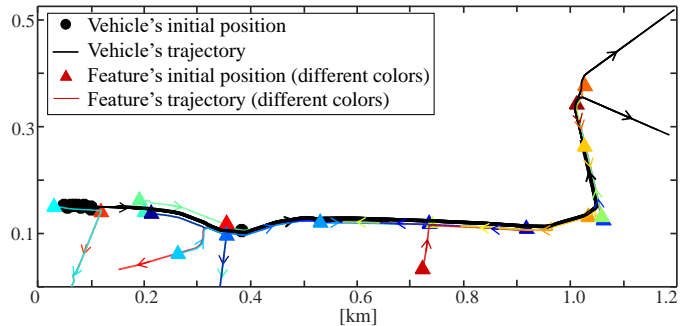


Fig. 8. Superposition of 10 vehicle (thick black line) and 20 pedestrian (colored line) trajectories simulated by SUMO for 200 s in the Bologna scenario of Fig. 7.

TABLE I  
GNSS QUALITY ASSOCIATED TO EACH AREA OF THE BOLOGNA'S SCENARIO IN FIG. 7

Area	Street	Environment	GNSS conditions
A1	Via Tolmino Viale G. Vicini	Open sky, large road with 3 by 3 lanes, scattered med-size buildings	Nominal $\sigma_{\text{GNSS}} = 1\bar{\sigma}_{\text{GNSS}}$
A2	Via Sabotino Viale A. Silvani	Some blockage, narrow road, 3 lanes, scattered medium-size buildings	Slightly degraded $\sigma_{\text{GNSS}} = 2\bar{\sigma}_{\text{GNSS}}$
A3	Via San Felice	Ultra narrow road, 2 lanes, urban canyon	Severely degraded $\sigma_{\text{GNSS}} = 5\bar{\sigma}_{\text{GNSS}}$
A4	Via San Felice	Ultra narrow road, 2 lanes, urban canyon	Lost $\sigma_{\text{GNSS}} = 20\bar{\sigma}_{\text{GNSS}}$

$\bar{\sigma}_{\text{GNSS}}$ . The simulated traces are used to determine the ground-truth reference, to calibrate vehicle/feature mobility models and to produce synthetic erroneous measurements. Tracking is performed by using the mobility models of vehicles and features respectively in (1) and (3). The standard deviation of V2F sensing is set to  $\sigma_{\text{V2F}} = 0.1 \text{ m}$  (as representative of RADAR accuracy [46]). The communication range at each vehicle is set to  $R_c = 200 \text{ m}$ , while the sensing range  $R_s$  is assumed to be lower and varies through simulations. **Numerical results.** Fig. 9 shows the cumulative distribution function (CDF) of the vehicle location error. The proposed distributed ICP method (red dashed line) is compared with the centralized ICP method (blue solid line) and the stand-alone GNSS (black solid line) solutions, for sensing range  $R_s = 100 \text{ m}$  and for different GNSS receivers: Figs. 9-(a) SPS, (b) SBAS, (c) DGNSS and (d) RTK (with cm-level accuracy). Results show that the distributed ICP algorithm achieves sub-meter accuracy for 75% of confidence level in all four cases and outperforms the stand-alone GNSS solution, even when RTK receiver is used though with a small gain. Moreover, the proposed distributed method attains the centralized approach.

In Fig. 10-(a), performances in terms of CDF of the location error are illustrated for different GNSS qualities associated to each street/area (see Table I). As expected, performance improvements are observed when the distributed ICP (dashed line) method is used with respect to the stand-alone GNSS (solid line) one. Fig. 10-(b) shows the CDF of the vehicle location error for the distributed ICP method with different

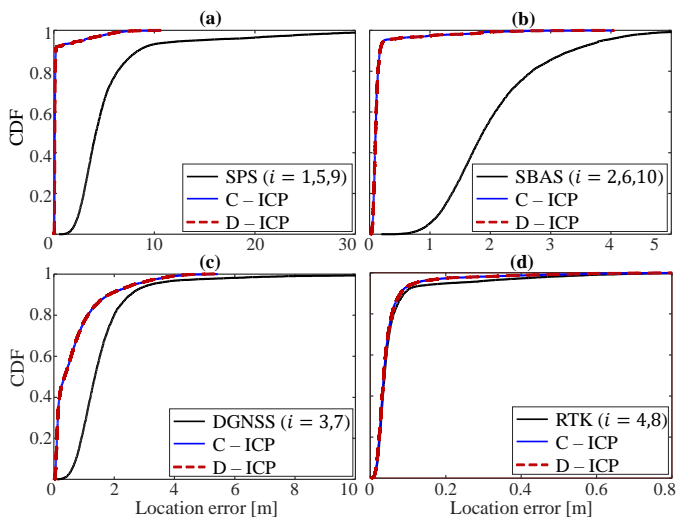


Fig. 9. CDF of the vehicle location error for the Bologna's scenario in Fig. 7, for the distributed (D-ICP) and centralized (C-ICP) cooperative methods and the stand-alone GNSS solution with four types of GNSS receivers: (a) SPS for vehicles  $i = 1, 5, 9$ , (b) SBAS for vehicles  $i = 2, 6, 10$ , (c) DGNSS for vehicles  $i = 3, 7$  and (d) RTK for vehicles  $i = 4, 8$ .

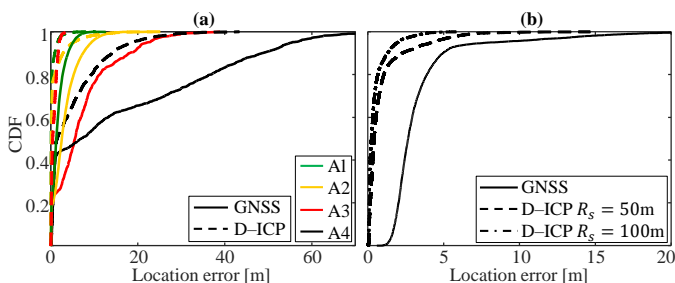


Fig. 10. CDF of the vehicle location error for the Bologna scenario in Fig. 7, for the distributed ICP and stand-alone GNSS. Positioning accuracy (a) over different areas for  $R_s = 50$  m and (b) for  $R_s = 50$  m and  $R_s = 100$  m.

sensing ranges,  $R_s = 50$  m and  $R_s = 100$  m (respectively, dashed and dashed-dot lines), compared with the stand-alone GNSS (solid line). For 50% of confidence level, the ICP approach achieves a location accuracy of 0.46 m for  $R_s = 50$  m and 0.23 m for  $R_s = 100$ , while the stand-alone GNSS accuracy is 2.65 m.

Zoomed view of Bologna city over the intersection between Viale Sabotino (area 2) and Viale G. Vicini (area 1) is shown in Fig. 11. Here, the average performances are evaluated by computing the  $2 \times 2$  mean square error matrix of vehicles' position estimates at convergence,  $\mathbf{MSE} = E[(\hat{\mathbf{p}}_{i,t}^{(V)} - \mathbf{p}_{i,t}^{(V)})(\hat{\mathbf{p}}_{i,t}^{(V)} - \mathbf{p}_{i,t}^{(V)})^T]$ , over 100 independent observations, for both the distributed ICP algorithm (red ellipse) and stand-alone GNSS method for different types of GNSS receivers (coloured contours). For visualization purposes, the error ellipses at 98.9% confidence are plotted around the mean vehicles' position estimates. The V2F and V2V connectivities are also given (respectively black solid and grey dashed-dot lines). Results show that all vehicles improve their position accuracy by using the ICP method compared to the performances obtained by the stand-alone GNSS solution. Note that the location accuracy given by the proposed ICP algorithm is not uniform among

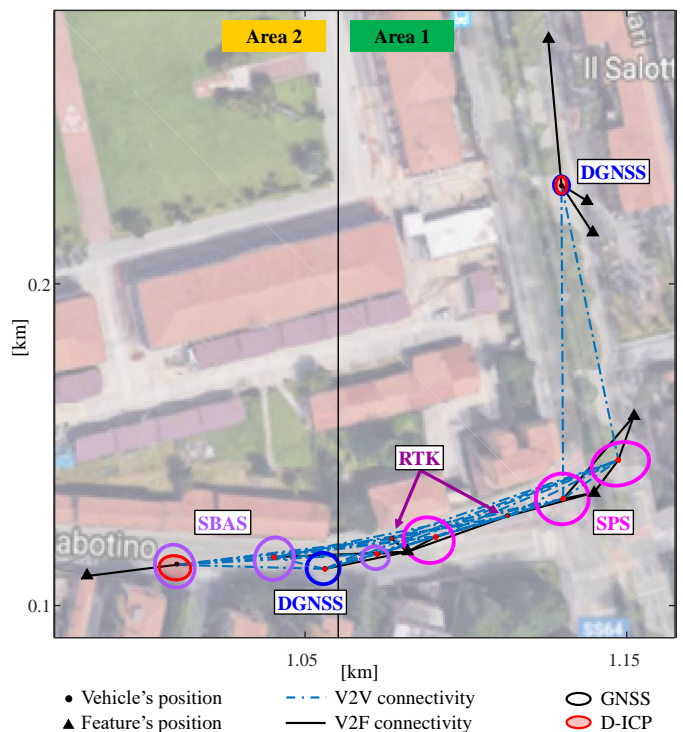


Fig. 11. Localization accuracy for the Bologna scenario of Fig. 7 over areas 1 and 2 at time  $t = 137$  s, for sensing ranges  $R_s = 50$  m: distributed ICP (red ellipse) and stand-alone GNSS (colored contours based on receiver type). Zoomed view over the intersection between Viale Sabotino (A2) and Viale G. Vicini (A1). Map courtesy of Google Maps.

vehicles as it depends on the type of GNSS receiver, on the GNSS signal quality in the area in which vehicles are traveling, but also on the vehicles' and features' positions.

## VIII. CONCLUSION

In this paper, a novel framework of cooperative positioning in vehicular networks was proposed, in which vehicles had to estimate a set of common passive features in a fully distributed way to improve the GNSS-based vehicle positioning. Starting from a FG formulation of the positioning problem, we developed a distributed Gaussian message passing algorithm that employed a consensus-based scheme for the distributed estimation of the features' positions. Simulation results demonstrated that the proposed methodology can accurately estimate the features' positions and (implicitly) improve the vehicle positioning accuracy compared to the stand-alone GNSS solution. Moreover, the ICP method was validated in a real urban scenario using the SUMO traffic simulator.

The framework made several limiting assumptions. First of all, the assumption of a linear measurement model can be removed by considering arbitrary non-linear models with non-parametric (e.g., particles) or parametric (e.g., Gaussian mixtures) message representations. Secondly, the assumption of perfect data association can be removed by including the data association problem in the FG. Investigation of these issues is a topic of further research.



## APPENDIX

Based on the result in (41) and recalling that  $\alpha_p^{(V)} = N_f/\sigma_{V2F}^2 + 1/\sigma_{p,pr}^{(V)2} + 1/\sigma_{GNSS}^2$ , the limiting cases considered in Sect. V are derived as follows:

1) If  $N_v \rightarrow \infty$ , (42) is given by:

$$\sigma_{p,post}^{(V)2} \rightarrow \frac{1}{\alpha_p^{(V)}} \cdot \frac{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2}{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2} = \frac{1}{\alpha_p^{(V)}}. \quad (49)$$

2) If  $N_f \rightarrow \infty$ , (43) is obtained as:

$$\sigma_{p,post}^{(V)2} \rightarrow \frac{1}{N_f/\sigma_{V2F}^2} \cdot \frac{N_f/\sigma_{V2F}^2 + N_f/\sigma_{p,pr}^{(F)2}}{N_f/\sigma_{p,pr}^{(F)2}} \rightarrow \frac{\sigma_{p,pr}^{(F)2} + \sigma_{V2F}^2}{N_f} \rightarrow 0. \quad (50)$$

3) If  $\sigma_{V2F}^2 \rightarrow 0$ , (44) is:

$$\begin{aligned} \sigma_{p,post}^{(V)2} &\rightarrow \frac{1}{N_f/\sigma_{V2F}^2} \cdot \frac{N_f/\sigma_{V2F}^2}{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2 + N_f/\sigma_{p,pr}^{(F)2}} \\ &= \frac{1}{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2 + N_f/\sigma_{p,pr}^{(F)2}}. \end{aligned} \quad (51)$$

4) If  $\sigma_{p,pr}^{(F)2} \rightarrow 0$ , (45) is given as:

$$\begin{aligned} \sigma_{p,post}^{(V)2} &\rightarrow \frac{1}{\alpha_p^{(V)}} \cdot \frac{N_f/\sigma_{p,pr}^{(F)2} + \sigma_{V2F}^2/(\sigma_{p,pr}^{(V)2} \sigma_{p,pr}^{(F)2}) + \sigma_{V2F}^2/(\sigma_{GNSS}^2 \sigma_{p,pr}^{(F)2})}{N_f/\sigma_{p,pr}^{(F)2} + \sigma_{V2F}^2/(\sigma_{p,pr}^{(V)2} \sigma_{p,pr}^{(F)2}) + \sigma_{V2F}^2/(\sigma_{GNSS}^2 \sigma_{p,pr}^{(F)2})} \\ &= \frac{1}{\alpha_p^{(V)}}. \end{aligned} \quad (52)$$

5) If  $\sigma_{p,pr}^{(V)2} \rightarrow 0$  or if  $\sigma_{GNSS}^2 \rightarrow 0$ , then (46) is obtained as:

$$\begin{aligned} \sigma_{p,post}^{(V)2} &\rightarrow \frac{1}{1/\sigma_{GNSS}^2 + 1/\sigma_{p,pr}^{(V)2}} \\ &\frac{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2 + \sigma_{V2F}^2/(\sigma_{p,pr}^{(V)2} \sigma_{p,pr}^{(F)2}) + \sigma_{V2F}^2/(\sigma_{GNSS}^2 \sigma_{p,pr}^{(F)2})}{N_v/\sigma_{p,pr}^{(V)2} + N_v/\sigma_{GNSS}^2 + \sigma_{V2F}^2/(\sigma_{p,pr}^{(V)2} \sigma_{p,pr}^{(F)2}) + \sigma_{V2F}^2/(\sigma_{GNSS}^2 \sigma_{p,pr}^{(F)2})} \\ &= \frac{1}{1/\sigma_{GNSS}^2 + 1/\sigma_{p,pr}^{(V)2}} \rightarrow 0. \end{aligned} \quad (53)$$

## REFERENCES

- [1] A. Pascale and et al., "Wireless sensor networks for traffic management and road safety," *IET Intelligent Transport Systems*, vol. 6, no. 1, pp. 67–77, Mar. 2012.
- [2] E. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*. Norwood: Artech House, 2006.
- [3] M. Skoglund and et al., "Static and dynamic performance evaluation of low-cost RTK GPS receivers," in *IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2016, pp. 16–19.
- [4] S. Thrun and M. Montemerlo, "The graph SLAM algorithm with applications to large-scale mapping of urban structures," *The Int. J. of Robotics Research*, vol. 25, no. 5-6, pp. 403–429, 2006.
- [5] S. E. Shladover and S.-K. Tan, "Analysis of Vehicle Positioning Accuracy Requirements for Communication-based Cooperative Collision Warning," *J. of Intelligent Transportation Systems*, vol. 10, no. 3, pp. 131–140, 2006.
- [6] M. During and K. Lemmer, "Cooperative Maneuver Planning for Cooperative Driving," *IEEE Intelligent Transportation Systems Mag.*, vol. 8, no. 3, pp. 8–22, Fall 2016.
- [7] H. Wymeersch and et al., "Challenges for Cooperative ITS: Improving Road Safety through the Integration of Wireless Communications, Control, and Positioning," in *IEEE Int. Conf. on Computing, Networking and Communications (ICNC)*, 2015, pp. 573–578.
- [8] "Vehicle Safety Communications - Applications (VSC-A) Final Report: Appendix volume 1 System Design and Objective Test," Tech. Rep., 2011.
- [9] "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service," *ETSI Std 302 637-2 V1. 3.2, Oct.*, 2014.
- [10] R. Parker and S. Valaee, "Vehicular Node Localization using Received-Signal-Strength Indicator," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 6, pp. 3371–3380, 2007.
- [11] —, "Cooperative Vehicle Position Estimation," *IEEE Int. Conf. on Communications (ICC)*, pp. 5837–5842, 2007.
- [12] J. Yao and et al., "Improving Cooperative Positioning for Vehicular Networks," *IEEE Trans. on Vehicular Technology*, vol. 60, no. 6, pp. 2810–2823, 2011.
- [13] N. M. Drawil and O. Basir, "Intervehicle-Communication-Assisted Localization," *IEEE Trans. on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 678–691, 2010.
- [14] G. M. Hoang and et al., "Distributed Link Selection and Data Fusion for Cooperative Positioning in GPS-Aided IEEE 802.11p VANETs," in *12th IEEE Workshop on Positioning, Navigation and Communications (WPNC)*, Mar. 2015.
- [15] —, "Cooperative localization in GNSS-aided VANETs with accurate IR-UWB range measurements," in *13th Workshop on Positioning, Navigation and Communications (WPNC)*, Oct. 2016, pp. 1–6.
- [16] R. Raulefs and et al., "Bound-based Spectrum Allocation for Cooperative Positioning," *Trans. on Emerging Telecommunications Technologies*, vol. 24, no. 1, pp. 69–83, 2013. [Online]. Available: <http://dx.doi.org/10.1002/ett.2572>
- [17] E.-K. Lee and et al., "RFID Assisted Vehicle Positioning in VANETs," *Pervasive and Mobile Computing*, vol. 8, no. 2, pp. 167–179, 2012.
- [18] M. Rohani and et al., "Vehicular Cooperative Map Matching," in *IEEE Int. Conf. on Connected Vehicles and Expo (ICCVE)*, 2014, pp. 799–803.
- [19] D. Wu and et al., "Location-based crowdsourcing for vehicular communication in hybrid networks," *IEEE Trans. on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 837–846, 2013.
- [20] —, "Online war-driving by compressive sensing," *IEEE Trans. on Mobile Computing*, vol. 14, no. 11, pp. 2349–2362, 2015.
- [21] C.-H. Ou, "A roadside unit-based localization scheme for vehicular ad hoc networks," *Int. J. of Communication Systems*, vol. 27, no. 1, pp. 135–150, 2014.
- [22] H. Wymeersch and et al., "Cooperative Localization in Wireless Networks," *Proc. of the IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.
- [23] F. Meyer and et al., "Distributed Localization and Tracking of Mobile Networks Including Non-Cooperative Objects," *IEEE Trans. on Signal and Inf. Processing over Networks*, vol. 2, no. 1, pp. 57–61, Mar. 2016.
- [24] G. Soatti and et al., "Consensus-based Algorithms for Distributed Network-State Estimation and Localization," *IEEE Trans. on Signal and Inf. Processing over Networks*, Mar. 2017.
- [25] A. Boukerche and et al., "Vehicular Ad-hoc Networks: A New Challenge for Localization-based Systems," *Computer Communications*, vol. 31, no. 12, pp. 2838–2849, 2008.
- [26] G. Hoang and et al., "On Communication Aspects of Particle-based Cooperative Positioning in GPS-aided VANETs," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 20–25.
- [27] F. de Ponte Müller, "Survey on ranging sensors and cooperative techniques for relative positioning of vehicles," *Sensors*, vol. 17, no. 2, p. 271, 2017.
- [28] G. Soatti and et al., "Enhanced Vehicle Positioning in Cooperative ITS by Joint Sensing of Passive Features," in *IEEE 20th Int. Conf. on Intelligent Transportation Systems (ITSC)*, Oct. 2017.
- [29] F. Gustafsson and F. Gunnarsson, "Mobile Positioning using Wireless Networks: Possibilities and Fundamental Limitations based on Available Wireless Network Measurements," *IEEE Signal Processing Mag.*, vol. 22, no. 4, pp. 41–53, Jul. 2005.
- [30] S. M. Kay, *Fundamentals of Statistical Signal Processing, volume I: Estimation Theory*. Prentice Hall, 1993.
- [31] F. R. Kschischang and et al., "Factor Graphs and the Sum-Product Algorithm," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [32] H.-A. Loeliger, "An Introduction to Factor Graphs," *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 28–41, 2004.
- [33] P. Rusmevichientong and B. Van Roy, "An Analysis of Belief Propagation on the Turbo Decoding Graph with Gaussian Densities," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 745–765, 2001.
- [34] Y. Weiss and W. T. Freeman, "On the Optimality of Solutions of the Max-Product Belief-Propagation Algorithm in Arbitrary Graphs," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 736–744, 2001.
- [35] —, "Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology," *Neural computation*, vol. 13, no. 10, pp. 2173–2200, 2001.

- [36] R. Olfati-Saber and et al., "Consensus and Cooperation in Networked Multi-Agent Systems," *Proc. of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
- [37] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge university press, 1987.
- [38] Y. Bar-Shalom, *Multitarget-multisensor tracking: Advanced applications*, 1990.
- [39] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.
- [40] D. Musicki and R. J. Evans, "Multiscan multitarget tracking in clutter with integrated track splitting filter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1432–1447, Oct. 2009.
- [41] L. Chen and et al., "Data association based on optimization in graphical models with application to sensor networks," *Mathematical and Computer Modelling*, vol. 43, no. 9, pp. 1114–1135, 2006.
- [42] F. Meyer and et al., "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Trans. on Signal Processing*, vol. 65, no. 13, pp. 3478–3493, Jul. 2017.
- [43] B. Wang and et al., "Distributed fusion of labeled multi-object densities via label spaces matching," *arXiv preprint arXiv:1603.08336*, 2016.
- [44] L. Hobert and et al., "Enhancements of V2X communication in support of cooperative autonomous driving," *IEEE Communications Mag.*, vol. 53, no. 12, pp. 64–70, Dec. 2015.
- [45] D. Krajzewicz and et al., "Recent development and applications of SUMO-Simulation of Urban MObility," *Int. J. On Advances in Systems and Measurements*, vol. 5, no. 3&4, 2012.
- [46] B. Denis and et al., "D5.3 - Cooperative Localization Algorithms and Hybrid Data Fusion Schemes [Intermediary]," High Precision Positioning for Cooperative-ITS, Tech. Rep., 2017.



**Nil Garcia** (S14, M16) received the Telecommunications Engineer degree from the Polytechnic University of Catalonia (UPC), Barcelona, Spain, in 2008; and the double Ph.D. degree in electrical engineering from the New Jersey Institute of Technology, Newark, NJ, USA, and from the National Polytechnic Institute of Toulouse, Toulouse, France, in 2015. He is currently a postdoctoral researcher of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. In 2009, he worked as an engineer in the Centre National d'Études Spatiales (CNES). In 2008 and 2009, he had Internships in CNES and NASA. His research interests are in the areas of localization, intelligent transportation systems and 5G.



**Benoît Denis** received the E.E. (2002), M.Sc. (2002), and Ph.D. (2005) degrees from INSA (Rennes, France) in electronics and communication systems. Since December 2005, he has been with CEA-Leti Minatec (Grenoble, France), contributing into French (ANR), European (FP6/FP7/H2020) and extra-European (QNRF) collaborative research projects in the fields of wireless sensor networks and wearable networks, heterogeneous and cooperative networks, vehicular networks, and mobile applications related to the Internet of Things or Connected Intelligent Transportation Systems. His main research interests concern joint wireless localization and communications, ranging/positioning/tracking and hybrid data fusion algorithms, radio channel modeling and cross-layer protocol design. He as (co)authored about 100 scientific papers on the previous topics.



**Gloria Soatti** received the M.Sc. degree in Telecommunication Engineering in 2012 and the Ph.D. degree (cum laude) in Information Technology in Feb. 2017 both from Politecnico di Milano (Italy). She was a visiting researcher at the Department of Signals and Systems, Chalmers University of Technology (Sweden) in 2016. Currently, she is a postdoctoral researcher at the Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB) of Politecnico di Milano. Her research interests are in the field of signal processing, particularly distributed consensus-based approaches for wireless sensor networks, IoT cognitive radios and vehicular networks.



**Ronald Raulefs** received the Dipl.-Ing. degree from the University of Kaiserslautern, Germany, in 1999 and the Dr.-Ing. (PhD) degree from the University of Erlangen-Nuremberg, Germany, in 2008. He is working as senior research member at the Institute of Communications and Navigation of the German Aerospace Center (DLR) in Oberpfaffenhofen, Germany. Ronald Raulefs initiated and lead the EU FP7 project WHERE and its successor project WHERE2 ([www.ict-where2.eu](http://www.ict-where2.eu)) as well as the task on cooperative location and communications in heterogeneous networks. He taught courses on the cooperation between wireless communications and positioning systems, gave tutorials at various conferences and summer schools such as at the IEEE conferences ICC'17 and PIMRC'17. He authored and co-authored 120+ scientific publications in conferences and journals. His current research interests focus on mobile radio communications and localization, including cooperative positioning for future cellular communication systems



**Monica Nicoli** (M99) received the M.Sc. degree (with honors) and the Ph.D. degree in Telecommunication Engineering from Politecnico di Milano, in 1998 and 2002, respectively. During 2001 she was Visiting Researcher with Uppsala University, Sweden. Since 2002 she has been with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, where she is Assistant Professor. Her research interests are in the area of signal processing, with emphasis on wireless communications, distributed and cooperative systems, intelligent transportation systems, wireless positioning and navigation. Dr. Nicoli is an Associate Editor of the EURASIP Journal on Wireless Communications and Networking, and she served as Lead Guest Editor for the Special Issue on Localization in Mobile Wireless and Sensor Networks in 2011. She has been member of the technical program committees of several conferences in the area of signal processing and wireless communications. She received the Marisa Bellisario Award in 1999 and the Premium Award for the Best Paper in IET Intelligent Transport Systems in 2014.



**Henk Wymeersch** (S01, M05) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication Letters (2009–2013), IEEE Transactions on Wireless Communications (since 2013), and IEEE Transactions on Communications (since 2016). His current research interests include cooperative systems and intelligent transportation.