Beyond-Line-of-Sight Identification by Using Vehicle-to-Vehicle Communication

Linjun Zhang and Gábor Orosz

***Abstract*— In this paper, we investigate the identification of the configuration and the dynamics of connected vehicle systems, where wireless vehicle-to-vehicle communication is used to access the motion data of vehicles that are beyond the line of sight. In particular, we first construct a causality detector to determine whether the information received from distant vehicles is relevant to the motion of the receiving vehicle. Then, we design a link-length estimator to identify the number of vehicles between the broadcasting vehicle and the receiving vehicle, which is required for appropriately incorporating the received data into the vehicle control system. Finally, a dynamics identifier is proposed to approximate the nonlinear time-delayed dynamics of vehicle chains, which is needed for the controller design to achieve desired system-level performance. The presented analytical results are validated through numerical simulations using synthetic data and on-road experiments.**

***Index Terms*— Beyond-line-of-sight identification, connected vehicle systems, vehicle-to-vehicle (V2V) communication.**

I. Introduction

GROUND transportation systems play a significant role in modern society. However, road transportation is also responsible for more than 30,000 fatalities per year [1] and 4.2 billion hours wasted in traffic congestions [2] in the United States alone. The demands for safer and more efficient transportation motivate the research on intelligent transportation systems.

In order to decrease the drivers' workload, adaptive cruise control (ACC) was designed to regulate the longitudinal motion of vehicles by exploiting range sensors (e.g., radar, lidar) that monitor the motion of the vehicle immediately ahead [3], [4]. Although ACC can improve passengers' comfort, its enhancements in safety and mobility are limited since the applied sensors are only able to detect objects within the line of sight. To address this problem, wireless vehicle-to-vehicle (V2V) communication can be used to access the motion data of vehicles ahead, even those beyond the line of sight. Incorporating V2V communication into vehicle control systems has great potentials for improving fuel efficiency [5], [6] and enhancing vehicle safety [7]. An application of V2V communication in vehicle control systems is to construct cooperative adaptive cruise control (CACC), where vehicles respond to the motion of the vehicle immediately ahead based on sensing information and they also respond to the motion of the designated platoon leader via V2V communication [8], [9]. The advantages of CACC in improving the capacity of traffic flow has been shown by theoretical studies [10]–[14] and experiments [15]–[17]. However, implementing this concept in real traffic is difficult, since it requires that multiple vehicles equipped with ACC travel next to each other, which rarely occurs in practice due to the low penetration of such vehicles.

Relaxing the aforementioned restrictions, we proposed the concept of connected cruise control (CCC) [18]–[20], which allows the incorporation of human-driven vehicles that may or may not broadcast information. Moreover, CCC requires neither a designated leader nor a prescribed connectivity topology. Such flexibility makes it practical to implement CCC in real traffic, even when the penetration of vehicles equipped with V2V communication is low. Zhang and Orosz [18] proposed a motif-based approach for modularly designing complex vehicle networks while taking into account information delays and connectivity topologies. The impacts of nonlinear dynamics and stochastic delays on the stability of connected vehicle systems were investigated in [21] and [22], respectively, while optimal CCC was studied in [23]. All these studies on CCC assumed that the configuration and the dynamics of the connected vehicle system were known. However, in practice such information may be not available, especially when the traffic flow includes non-broadcasting vehicles. In order to identify the configuration and the dynamics of connected vehicle systems in the neighborhood of the CCC vehicle, we establish the concept of beyond-line-of-sight (BLOS) identification that utilizes information gained via V2V communication.

BLOS identification is challenging due to the following reasons. First, the received data may be not relevant to the motion of the CCC vehicle, even when the broadcasting vehicle is in the same lane. For instance, if the broadcasting vehicle is far from the CCC vehicle while there are no vehicles between them, the information of the broadcasting vehicle is not relevant to the CCC vehicle. Moreover, to appropriately utilize the information received from a distant vehicle, one needs to know the number of non-broadcasting vehicles between the broadcasting vehicle and the CCC vehicle. In addition, the dynamics of connected vehicle systems are typically nonlinear and include delays, which increases the complexity for identifying the dynamics.

In this paper, we propose a systematic approach for BLOS identification by using V2V communication and make the

Manuscript received November 11, 2016; revised June 15, 2017; accepted August 19, 2017. Date of publication October 4, 2017; date of current version May 29, 2018. This work was supported by the National Science Foundation under Award 1351456. The Associate Editor for this paper was L. Li. *(Corresponding author: Linjun Zhang.)*

L. Zhang is with Ford Motor Company, Dearborn, MI 48124 USA (e-mail: linjunzh@umich.edu).

G. Orosz is with the Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48105 USA (e-mail: orosz@umich.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2017.2747582

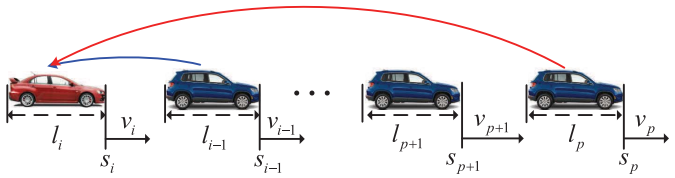


Fig. 1. A connected vehicle system where the CCC vehicle i receives information from multiple vehicles ahead. Symbols s_j, l_j, and v_j denote the position, length, and velocity of vehicle j, respectively. The short-range link (blue) can be realized by range sensors or V2V communication while the long-range links (red) can only be realized by V2V communication since distant vehicles are beyond the line of sight. Note that the vehicle chain allows the incorporation of non-broadcasting vehicles (e.g., vehicle $p+1$).

following contributions. First, we propose a causality detector to determine whether the received information is relevant to the CCC vehicle. Then, we design a link-length estimator for identifying the number of non-broadcasting vehicles between the broadcasting vehicle and the CCC vehicle. Finally, we construct a dynamics identifier to approximate the nonlinear time-delayed dynamics of connected vehicle systems by using linear discrete-time models. The coefficients of the model are obtained by solving an optimization problem where the stability of the linear discrete-time model is explicitly considered as a constraint. We propose an approach that solves the stability-constrained optimization problem in an efficient way. Numerical simulations using synthetic data and on-road experiments are used to validate the analytical results and to demonstrate the robustness of our presented estimator against changing configurations of connected vehicle systems.

The rest of this paper is organized as follows. In Sections II, III, and IV, we present the causality detector, the link-length estimator, and the dynamics identifier, respectively. The results of numerical simulations and on-road experiments are presented in Section V. In Section VI, we conclude our results and discuss future research directions.

II. Causality Detector

In this paper, we focus on the longitudinal motion of vehicles and assume that the CCC vehicle monitors the motion of the vehicle immediately ahead by using range sensors and also receives information from distant vehicles via V2V communication. Moreover, we consider the case that the configuration of a connected vehicle system may vary since vehicles may leave or join the lane. Fig. 1 shows a connected vehicle system where the CCC vehicle i monitors the positions s_j and the velocities v_j of some of the vehicles $j = p, \ldots, i-1$ with p denoting the furthest broadcasting vehicle within the communication range of vehicle i. Note that CCC allows the incorporation of vehicles that do not broadcast information; for example, see vehicle $p+1$ in Fig. 1. The symbol l_j represents the length of vehicle j. Vehicle $i-1$ can be monitored by range sensors (e.g., radar, lidar) or by V2V communication, while distant vehicles can only be monitored via V2V communication since they are beyond the line of sight. The difficulty for beyond-line-of-sight (BLOS) identification increases when the traffic flow includes less broadcasting vehicles. Considering the worst-case scenario, throughout this paper we assume that the CCC vehicle i can only access the motion data of vehicles p and $i-1$.

According to SAE J2739 Message Set Dictionary standard [24], V2V communication provides information that includes position, velocity, and vehicle size. The information is broadcast and received every $\delta t = 0.1$ [s]. For simplicity, we assume that the clocks of all vehicles are synchronized and they collect information at time instants $t_k = k \cdot \delta t$ for $k = 0, 1, 2, \ldots$. We also define the vector of motion data in the form

$$\begin{aligned} \mathbf{s}_j(t_k) &= \big[s_j(t_{k-\Delta}), s_j(t_{k-\Delta+1}), \ldots, s_j(t_k)\big], \\ \mathbf{v}_j(t_k) &= \big[v_j(t_{k-\Delta}), v_j(t_{k-\Delta+1}), \ldots, v_j(t_k)\big], \end{aligned} \tag{1}$$

where the positive constant Δ represents the amount of data stored in the memory. Historical data are used to enhance the confidence about the identification.

To appropriately use the information received via V2V communication, it is important to detect whether the received information is relevant to the motion of the CCC vehicle. To determine such relevance, we propose a causality detector in this section. The motion of two vehicles are said to have causal relationship if the motion of one vehicle influences the motion of the other vehicle. For example, consider the vehicle chain in Fig. 1. If the motion of vehicles p and i have causal relationship, the speed perturbation of vehicle p will lead to a (similar) speed perturbation of vehicle i since perturbations propagate along the chain of vehicles. Otherwise, vehicle i may not respond to the motion of vehicle p. Such non-causality occurs when there are no vehicles between vehicles $i-1$ and p while they maintain a large distance.

We evaluate the causality by using the concept of lag phase, which we define as the time difference between similar motions of two vehicles. In practice, the lag phase may vary in time, but the possible values shall be in a bounded range when the number of vehicles between the broadcasting vehicle and the receiving vehicle is bounded. For implementation, we consider a set of values for lag phases

$$\mathcal{D}_\tau = \{\tau_1, \tau_2, \ldots, \tau_N\}, \tag{2}$$

where $\tau_\ell = \ell \cdot \delta\tau$ for $\ell = 1, \ldots, N$ with $\delta\tau$ representing the resolution of the distribution. To construct the causality detector, we define the weights corresponding to the lag phase $\tau_\ell \in \mathcal{D}_\tau$ between vehicles i and j at time t_k as

$$w_{i,j}^{(t_k)}(\tau_\ell) = \frac{1}{\|\widetilde{\mathbf{v}}_i(t_k) - \widetilde{\mathbf{v}}_j(t_k - \tau_\ell)\|_2}, \tag{3}$$

for $\ell = 1, \ldots, N$, where

$$\widetilde{\mathbf{v}}(t) = \mathbf{v}(t) \big/ \|\mathbf{v}(t)\|_\infty \tag{4}$$

denotes the normalized signal such that all elements in the set $\widetilde{\mathbf{v}}(t)$ are between 0 and 1; cf. (1). The 2-norm and the ∞-norm of vectors in (3) and (4) are defined in [25]. Then we normalize the weights given by (3) as

$$\widetilde{w}_{i,j}^{(t_k)}(\tau_\ell) = \frac{w_{i,j}^{(t_k)}(\tau_\ell)}{\sum_{\ell=1}^{N} w_{i,j}^{(t_k)}(\tau_\ell)}, \tag{5}$$

which yields $\sum_{\ell=1}^{N} \widetilde{w}_{i,j}^{(t_k)}(\tau_\ell) = 1$.

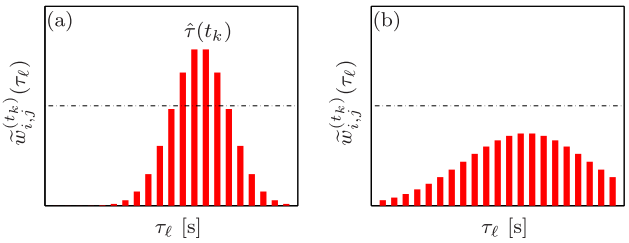


Fig. 2. Normalized weight $\widetilde{w}_{i,j}^{(t_k)}(\tau_\ell)$ for lag phase τ_ℓ at time t_k. The dashed-dotted line denotes the confidence threshold. (a) If the maximum weight is larger then the confidence threshold, the lag phase associated with the maximum weight leads to the estimate $\hat{\tau}(t_k)$. (b) If the maximum weight is smaller than the confidence threshold, it implies lack of confidence about the causal relationship between the motion of two vehicles.

When there is causal relationship between the motion of two vehicles, the distribution (5) becomes more concentrated; see Fig. 2. If the maximum value of the normalized weights (5) is larger than a given threshold (denoted by the dashed-dotted line in Fig. 2), the element in the set (2) associated with the maximum weight is chosen to be the estimate of lag phase $\hat{\tau}(t_k)$; see Fig. 2(a). Otherwise, it implies lack of confidence for causal relationship between the motion of two vehicles; see Fig. 2(b). We remark that when the motion of two vehicles are non-causal, the weights (5) tends to be uniformly distributed. For instance, this happens when one vehicle moves with constant speed while the speed of the other vehicle varies. When the motion of one vehicle is exactly a time-shifted copy of the motion of the other vehicle, the weight of the corresponding lag phase will be one while the other weights are zeros. For periodic motions, there may exist multiple elements in (2) that have the same weight. In this case, we use the smallest element as the estimated lag phase.

Note that the normalized weight (5) alone is not sufficient to evaluate the causality between the motion of two vehicles, because it is only determined by the instantaneous motion that may vary significantly at each time step. To enhance the confidence about the causality detection, we construct a distribution by using both the current and the historical data. In particular, we define

$$P_{i,j}^{(t_k)}(\tau_\ell) = \frac{Q_{i,j}^{(t_k)}(\tau_\ell)}{\sum_{m=1}^{N} Q_{i,j}^{(t_k)}(\tau_m)} \tag{6}$$

through the distribution

$$Q_{i,j}^{(t_k)}(\tau_\ell) = Q_{i,j}^{(t_{k-1})}(\tau_\ell) + \gamma\, \psi_{i,j}^{(t_k)}\, \widetilde{w}_{i,j}^{(t_k)}(\tau_\ell), \tag{7}$$

for $\ell = 1, \ldots, N$; cf. (2). Here, the positive constant γ determines the update rate and $\psi_{i,j}^{(t_k)}$ is given by the Kullback-Leibler divergence (KL-divergence)

$$\psi_{i,j}^{(t_k)} = \sum_{\ell=1}^{N} P_{i,j}^{(t_{k-1})}(\tau_\ell) \log \frac{P_{i,j}^{(t_{k-1})}(\tau_\ell)}{\widetilde{w}_{i,j}^{(t_k)}(\tau_\ell)}, \tag{8}$$

which quantifies the difference between the distributions $P_{i,j}^{(t_{k-1})}$ and $\widetilde{w}_{i,j}^{(t_k)}$; see [26]. Here, we use the KL-divergence because of its following properties:

- It is non-negative, and it is zero if and only if the two distributions are the same.
- It increases when the difference between the two distributions increases.

Since the distribution $P_{i,j}^{(t_k)}$ includes historical information, it is more robust against the variation of instantaneous motion and hence increases the confidence about the causality detection. As initial condition, we use the uniform distribution

$$P_{i,j}^{(t_0)}(\tau_\ell) = Q_{i,j}^{(t_0)}(\tau_\ell) = \frac{1}{N}, \tag{9}$$

for $\ell = 1, \ldots, N$; cf. (2).

To quantify how concentrated the distribution (6) is, we use the concentration function

$$C_{i,j}(t_k) = 1 - \frac{\min_{\tau_\ell} P_{i,j}^{(t_k)}(\tau_\ell)}{\max_{\tau_\ell} P_{i,j}^{(t_k)}(\tau_\ell)}. \tag{10}$$

Large value of $C_{i,j}(t_k)$ indicates high possibility of causal relationship between the motion of the vehicles i and j. Setting a threshold $0 < \overline{C} < 1$, we define the causality indicator

$$R_{i,j}(t_k) = \begin{cases} 1, & \text{if } C_{i,j}(t_k) > \overline{C}, \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

where $R_{i,j} = 1$ indicates the causal relationship between vehicles i and j while $R_{i,j} = 0$ implies lack of confidence about causality.

III. Link-Length Estimator

To appropriately incorporate the motion data received from distant vehicles into CCC, one needs to know the number of vehicles between the broadcasting vehicle and the CCC vehicle. However, in practice such information is not available, especially when some vehicles do not broadcast information. In this section, we present an estimator to identify the number of vehicles between the broadcasting vehicle and the CCC vehicle by using only the position and the velocity data obtained from V2V communication and range sensors.

To make the subsequent expressions more compact, we use the phrase *link length* to represent the number of vehicles between two vehicles. That is, the link length $n_{i,j}$ between vehicles i and j is given by

$$n_{i,j} = i - j, \tag{12}$$

cf. Fig. 1. We also use the *headway* $h_{j,j-1}$ to describe the distance between the rear bumper of vehicle $j-1$ and the front bumper of vehicle j, that is,

$$h_{j,j-1} = s_{j-1} - s_j - l_{j-1}, \tag{13}$$

for $j = p, \ldots, i$; cf. Fig. 1. It follows that the average headway $h_{i,j}$ between vehicle i and vehicle j ($j < i$) is given by

$$h_{i,j} = \frac{s_j - s_i - \sum_{m=j}^{i-1} l_m}{n_{i,j}}. \tag{14}$$

The empirical traffic data in [27] show that, when vehicle j moves at a constant speed v_j^*, it aims to keep a desired

speed-dependent headway $h^*_{j,j-1}$ from vehicle $j-1$. This relationship is called *range policy* that can be described by

$$h^*_{j,j-1} = H_j(v^*_j), \tag{15}$$

where the strictly positive function $H_j(v)$ increases with v over the operating domain $0 < v < v_{\max}$. Here, the maximum speed $v_{\max}$ may be determined by speed limit or by driver's decision. When all vehicles move at the same constant speed and they also use the same range policy, i.e., $v^*_j = v^*$ and $H_j(v) = H(v)$ for all j-s, it leads to the *uniform flow equilibrium*

$$h^*_{i,j} = \frac{s^*_j - s^*_i - \sum_{m=j}^{i-1} l_m}{n_{i,j}} = H(v^*), \tag{16}$$

cf. (12)–(15). Here, $s^*_j = \overline{s}_j + v^*_j t$ represents the equilibrium position of vehicle j for all j-s, where $\overline{s}_j$ is a constant. In real traffic, range policies may vary among different vehicles such that the equilibrium still exists but it is not uniform, i.e., $v^*_j = v^*_k = v^*$ but $h^*_{j,j-1} \neq h^*_{k,k-1}$ for $j \neq k$. In this situation, equation (16) can be written as

$$h^*_{i,j} = \frac{s^*_j - s^*_i - n_{i,j} \cdot l_{\rm av}}{n_{i,j}} = H_{\rm av}(v^*), \tag{17}$$

where the average vehicle length $l_{\rm av}$ and the average range policy $H_{\rm av}(v^*)$ are given by

$$l_{\rm av} = \frac{1}{n_{i,j}} \sum_{m=j}^{i-1} l_m, \quad H_{\rm av}(v^*) = \frac{1}{n_{i,j}} \sum_{m=j+1}^{i} H_m(v^*). \tag{18}$$

Notice that the length of vehicle i and the range policy of vehicle j are not included in (18); cf. (14) and (15).

If all vehicles broadcast information, the average quantities (18) may be directly obtained, and hence the link length can be calculated from (17), yielding

$$n_{i,j} = \frac{s^*_j - s^*_i}{l_{\rm av} + H_{\rm av}(v^*)}. \tag{19}$$

However, when there are vehicles that do not broadcast information, the average quantities (18) are not available. In this case, we use their approximations

$$\overline{l}_{\rm av} \approx l_{\rm av}, \quad \overline{H}_{\rm av}(v^*) \approx H_{\rm av}(v^*), \tag{20}$$

which may be obtained by statistical analysis of traffic data.

In real traffic, vehicles may not move at a constant velocity and the headway varies in time. In practice, we use the averages

$$\begin{aligned}
\xi_s(t_k) &\triangleq \mathbb{E}^{t_k}_{t_0}\big[s_j(t_q) - s_i(t_q)\big],\\
\xi_v(t_k) &\triangleq \mathbb{E}^{t_k}_{t_0}\big[\eta v_j(t_q) + (1-\eta) v_i(t_q)\big],\\
\xi_n(t_k) &\triangleq \mathbb{E}^{t_k}_{t_0}\left[\frac{\xi_s(t_q)}{\overline{l}_{\rm av} + \overline{H}_{\rm av}\big(\xi_v(t_q)\big)}\right],
\end{aligned} \tag{21}$$

for the equilibrium distance, equilibrium velocity, and equilibrium link length, respectively, where $0 \le \eta \le 1$ is a constant weight and the averaging operator is defined by

$$\mathbb{E}^{t_k}_{t_0}\big[x(t_q)\big] = \frac{1}{k+1} \sum_{q=0}^{k} x(t_q). \tag{22}$$

Rewriting (21) in iterative form and defining $\hat{n}_{i,j}(t_k)$ as the estimate of $n_{i,j}$, we propose the link-length estimator

$$\begin{aligned}
\xi_s(t_k) &= \xi_s(t_{k-1}) + \frac{s_j(t_k) - s_i(t_k) - \xi_s(t_{k-1})}{k},\\
\xi_v(t_k) &= \xi_v(t_{k-1}) + \frac{\eta v_j(t_k) + (1-\eta) v_i(t_k) - \xi_v(t_{k-1})}{k},\\
\xi_n(t_k) &= \xi_n(t_{k-1}) + \frac{1}{k}\Big(\frac{\xi_s(t_k)}{\overline{l}_{\rm av} + \overline{H}_{\rm av}\big(\xi_v(t_k)\big)} - \xi_n(t_{k-1})\Big),\\
\hat{n}_{i,j}(t_k) &= \mathrm{int}\big(\xi_n(t_k)\big),
\end{aligned} \tag{23}$$

for $k = 1, 2, \ldots$, where the operator $\mathrm{int}(\xi)$ rounds the estimated link length to the nearest integer. Here, initial condition is set to be $\xi_s(t_0) = \xi_v(t_0) = \xi_n(t_0) = 0$. Note that there is no need to set an initial guess for the estimated link length $\hat{n}_{i,j}$ since k starts from 1.

Theorem 1: The output of the link-length estimator (23) converges to the actual link length, i.e., $\hat{n}_{i,j}(t_k) \to n_{i,j}$ as $k \to \infty$, if and only if the following condition holds:

$$1 - \frac{1}{2n_{i,j}} < \frac{l_{\rm av} + H_{\rm av}(v^*)}{\overline{l}_{\rm av} + \overline{H}_{\rm av}(v^*)} < 1 + \frac{1}{2n_{i,j}}. \tag{24}$$

The proof is given in Appendix. We remark that the link-length estimator (23) does not rely on the motion data of vehicles $j+1, \ldots, i-1$ so that the estimator can be applied when they do not broadcast information. When more vehicles broadcast information, the accuracy of approximations $\overline{l}_{\rm av}$ and $\overline{H}_{\rm av}$ can be improved, leading to more accurate estimate of link length.

The link-length estimator (23) treats all historical data equally, which may reduce its sensitivity when the value of the link length changes. To address this issue, we modify the link-length estimator (23) by using a forgetting factor μ. This leads to the weighted link-length estimator

$$\begin{aligned}
\Omega(t_k) &= 1 + \mu\Omega(t_{k-1}),\\
\xi_s(t_k) &= \xi_s(t_{k-1}) + \frac{s_j(t_k) - s_i(t_k) - \xi_s(t_{k-1})}{\Omega(t_k)},\\
\xi_v(t_k) &= \xi_v(t_{k-1}) + \frac{\eta v_j(t_k) + (1-\eta) v_i(t_k) - \xi_v(t_{k-1})}{\Omega(t_k)},\\
\xi_n(t_k) &= \xi_n(t_{k-1}) + \frac{1}{\Omega(t_k)}\Big(\frac{\xi_s(t_k)}{\overline{l}_{\rm av} + \overline{H}_{\rm av}\big(\xi_v(t_k)\big)} - \xi_n(t_{k-1})\Big),\\
\hat{n}_{i,j}(t_k) &= \mathrm{int}\big(\xi_n(t_k)\big).
\end{aligned} \tag{25}$$

Here, the forgetting factor μ decreases the influence of historical data on the current estimation, which enhances the sensitivity. When $\mu = 1$, the estimator (25) is reduced to (23). Note that small μ may cause oscillations and hence does not lead to convergence. In practice, the value of μ should be chosen by considering the trade-off between the robustness and the sensitivity. The feasible values for choosing μ will be investigated in Section V based on the results of numerical simulations and on-road experiments.

IV. Dynamics Identifier

Consider the connected vehicle system shown in Fig. 1. To design CCC for vehicle i that can achieve certain system-level performance such as string stability, one needs to

know the dynamics of the vehicle chain comprised of vehicles $p, \ldots, i-1$; see [18]–[21]. Such information cannot be directly obtained in real traffic. In this section, we present an identifier to approximate the dynamics of connected vehicle systems.

In particular, we consider the velocities v_p and v_{i-1} as the input and the output of the connected vehicle system, since the speed perturbation of vehicle p propagates along the vehicle chain and eventually affects the speed of vehicle $i-1$. Although the car-following dynamics of human-driven vehicles is typically nonlinear and contains reaction time delays, in order to reduce complexity, we use a linear discrete-time model

$$\begin{aligned}\hat{v}_{i-1}(t_k) &= \sum_{q=1}^{N} \big(- a_q \hat{v}_{i-1}(t_{k-q}) + b_q v_p(t_{k-q}) \big) \\ &= \begin{bmatrix} -\theta_a \\ \theta_b \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \phi_a \\ \phi_b \end{bmatrix}, \end{aligned} \tag{26}$$

where the coefficients and the regressors are given by

$$\begin{aligned} \theta_a &= \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}, \quad \phi_a = \begin{bmatrix} \hat{v}_{i-1}(t_{k-1}) \\ \vdots \\ \hat{v}_{i-1}(t_{k-N}) \end{bmatrix}, \\ \theta_b &= \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix}, \quad \phi_b = \begin{bmatrix} v_p(t_{k-1}) \\ \vdots \\ v_p(t_{k-N}) \end{bmatrix}, \end{aligned} \tag{27}$$

respectively. Note that ϕ_a is comprised of estimated states and ϕ_b is constructed by using measured states, while the coefficient vectors θ_a, θ_b and the regression length N are to be determined.

When there are no information delays, the longitudinal dynamics of each vehicle can be modeled as a second-order system [28]. Hence, for an n-vehicle chain, we set the regression length as $N = 2n$. Since n is typically unknown in real traffic, we use its estimation $\hat{n}$ given by the estimator (25), and hence set $N = 2\hat{n}$. In practice, the order of the longitudinal dynamics of a vehicle may be higher than two due to information delays. Such modeling errors will be minimized by designing the coefficients θ_a and θ_b in (26).

We define the cost function

$$\begin{aligned} J(\theta_a, \theta_b) = \frac{1}{\Delta+1} &\|\mathbf{v}_{i-1}(t_k) - \hat{\mathbf{v}}_{i-1}(t_k)\|_2 \\ &+ c_1 \|\mathbf{v}_{i-1}(t_k) - \hat{\mathbf{v}}_{i-1}(t_k)\|_\infty + c_2 \|\theta_b\|_2, \end{aligned} \tag{28}$$

where $\Delta + 1$ is the number of elements in the vector $\mathbf{v}_{i-1}(t_k)$ and c_1, c_2 are positive constants. Here, the constant Δ and the vectors $\mathbf{v}_{i-1}(t_k)$, $\hat{\mathbf{v}}_{i-1}(t_k)$ are defined according to (1). Note that Δ is chosen to be much larger than N in (26). The first term in (28) is used to evaluate the mean of integral error, the second term represents the peak of error, while the third term is used to control the coefficient values to avoid over-fitting. Note that $J(\theta_a, \theta_b)$ explicitly depends on θ_b through the the third term and also inexplicitly depends on θ_a and θ_b through $\hat{\mathbf{v}}_{i-1}(t_k)$; cf. (26).

Then, the coefficients θ_a and θ_b can be obtained by solving the constrained optimization problem

$$\begin{aligned} &\underset{\theta_a, \theta_b}{\text{minimize}} \;\; J(\theta_a, \theta_b), \\ &\text{subject to} \left\{ |\lambda| < 1 : \lambda^N + \sum_{q=1}^{N} a_q \lambda^{N-q} = 0 \right\}, \end{aligned} \tag{29}$$

where the constraint is used to guarantee the stability of the model (26) by requiring all roots to be inside the unit circle in the complex plane. The optimization problem (29) is challenging, since λ-s are complex numbers and the constraint may lead to disconnected regions for coefficients $a_1, \ldots, a_N$. Now we propose a method that finds a suboptimal solution for the problem (29) in an efficient way. We first generate a pool of candidate coefficients $a_1, \ldots, a_N$ that satisfy the stability constraint, and then minimize the cost function $J(\theta_a, \theta_b)$ based on the candidate pool. The detailed process will be described in the rest of this section.

The characteristic equation in the constraint of (29) can be written in the form

$$\prod_{j=1}^{N} (\lambda + \lambda_j) = 0, \tag{30}$$

that is, the corresponding roots are given by $-\lambda_j$ for $j = 1, \ldots, N$. To ensure the constraint of (29), one needs to guarantee $|\lambda_j| < 1$ for $j = 1, \ldots, N$. To begin with, we randomly sample N numbers inside the unit circle in the complex plane. Since $a_1, \ldots, a_N$ are all real numbers, complex roots must appear in a conjugate way while the real roots shall be mutually independent. To realize this, we randomly sample $N/2$ values $\lambda_1, \ldots, \lambda_{N/2}$. If λ_j ($j = 1, \ldots, N/2$) is a real number, we randomly sample a real value between -1 and 1 for $\lambda_{j+N/2}$; otherwise, set $\lambda_{j+N/2} = \lambda_j'$ where the prime represents the complex conjugate. Indeed, there are many methods for sampling complex numbers in the unit circle. Here, we sample the magnitude and the real part of complex values. If the real part is larger than the magnitude, then the sample is set to be equal to the real part; otherwise it is set to be a complex number, where the imaginary part is calculated from the magnitude and the real part.

Using the sampled characteristic roots $\lambda_1, \ldots, \lambda_N$, one can construct the corresponding coefficients by using the formula

$$a_q = \sum_{\{i_1, \ldots, i_q\} \in \mathcal{C}_N^q} \lambda_{i_1} \cdots \lambda_{i_q}, \tag{31}$$

for $q = 1, \ldots, N$, where the set $\mathcal{C}_N^q$ contains all possible combinations for picking q elements from the set $\{1, \ldots, N\}$. Then, we repeat the aforementioned sampling process for M times, where M is a sufficiently large number. This leads to M sets of candidate coefficients $\{a_q\}_m$ for $m = 1, \ldots, M$. A larger size of candidate pool may improve the accuracy of the solution to (29), but it also requires more computational resources when searching for the optimal solution. Since linear systems with similar coefficients have similar dynamics, the large number of sampled coefficients can be clustered into

groups that are represented by the coefficients closest to the group center. For unsupervised clustering, one can apply different algorithms such as the k-means approach [29] and the self organizing map (SOM) approach [30]. Here, we use the k-means approach, since it is simple for implementation and, based on our observation, addresses this problem well. The k-means approach leads to K sets of representative coefficients $\{a_q\}_k$ for $k = 1, \ldots, K$. We remark that K is typically much smaller than M, but it needs to be selected appropriately. When K is too large, there may be too many groups such that the coefficients are not sufficiently separated; while when K is too small, some representative coefficients may be missed. An appropriate value for K is typically set through trial and error, which will be discussed in Section V based on numerical simulations.

Then, we compute $\min_{\theta_a,\theta_b} J(\theta_a, \theta_b)$ through iterations. First, we randomly initialize $\theta_b^{(0)}$ where the superscript indicates the iteration index. Then, for each iteration $\ell = 1, \ldots, L$, we update $\theta_a^{(\ell)}$ according to

$$\theta_a^{(\ell)} = \underset{\{a_q\}_k, k=1,\ldots,K}{\operatorname{argmin}} J\big(\{a_q\}_k, \theta_b^{(\ell-1)}\big), \tag{32}$$

where $\{a_q\}_k$ is the k-th representative candidates. Since K is typically small, the minimizer of (32) can be obtained by comparing the cost for each $\{a_q\}_k$ where $k = 1, \ldots, K$. Then, we update $\theta_b^{(\ell)}$ by using the gradient descent approach

$$\theta_b^{(\ell)} = \theta_b^{(\ell-1)} - \rho \nabla_{\theta_b} J\big(\theta_a^{(\ell)}, \theta_b^{(\ell-1)}\big), \tag{33}$$

where the step size ρ is used to ensure that the cost decreases at each step, i.e.,

$$J\big(\theta_a^{(\ell)}, \theta_b^{(\ell)}\big) \leq J\big(\theta_a^{(\ell)}, \theta_b^{(\ell-1)}\big). \tag{34}$$

In practice, we first try $\rho = 1/2$ in (33). If the result does not satisfy (34), we iteratively replace ρ by $\rho/2$ until the result satisfies (34). We remark that, once the coefficients θ_a and θ_b are obtained, we freeze these coefficients and use (26) as the approximate dynamic model of the vehicle chain, until the causality detector or the link-length estimator indicates that the configuration of the connected vehicle system has changed.

For BLOS identification, we have presented the causality detector (11), the link-length estimator (25), and the dynamics identifier (26). It is important to note that these three estimators are activated sequentially, and the activation sequence is shown by the flow chart in Fig. 3. Suppose that the CCC vehicle i starts to receive data from vehicle j at time t_0 via V2V communication. Then, the causality detector (11) is activated at time t_Δ when there are sufficient amount of data stored in the memory; cf. (1). If the causality indicator $R_{i,j}$ (cf. (11)) becomes 1 at time t_{ℓ_1}, the link-length estimator (25) starts to estimate the number of vehicles between the broadcasting vehicle j and the CCC vehicle i by using the historical data between $t_{\ell_1} - \Delta \cdot \delta t$ and t_{ℓ_1}. If the causality indicator $R_{i,j}$ becomes 0 sometime later, it implies no causal relationship between the motion of vehicles i and j, leading to the reset of the link-length estimator and the dynamics identifier if they have been activated. Note that the estimated link length $\hat{n}_{i,j}$ may vary after the estimator (25) is activated.

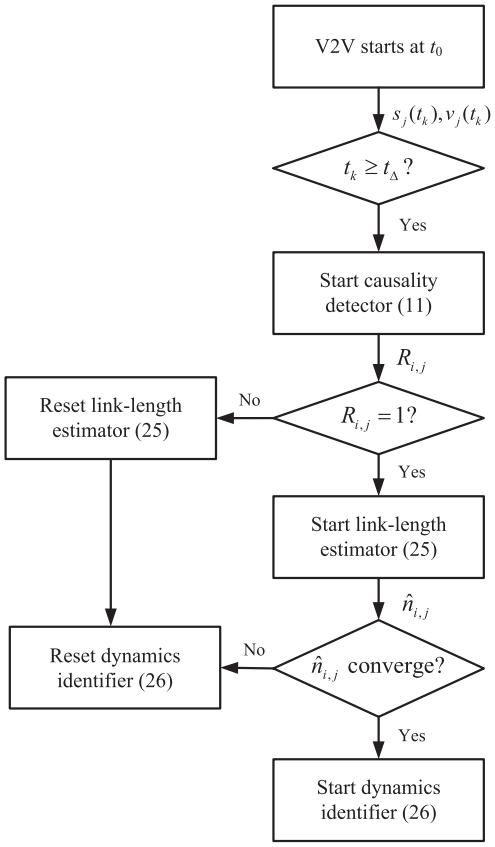


Fig. 3. Flow chart for the activations of causality detector (11), link-length estimator (25), and dynamics identifier (26). The causality detector is activated at t_Δ when the V2V communication has accumulated sufficient amount of data. The determination of the causality triggers the activation of the link-length estimator. After the estimated link length converges, the dynamics identifier is activated.

We say that the estimated link length $\hat{n}_{i,j}$ converges when it remains unchanged for a given time length T. The convergence of the estimated link length at t_{ℓ_2} will trigger the activation of the dynamics identifier (26) that uses the data between $t_{\ell_2} - \Delta \cdot \delta t$ and t_{ℓ_2}. After the coefficients in (26) have been identified, the model is frozen for the subsequent control design. However, if the estimated link length varies, it implies that the configuration of the vehicle chain has changed and hence resets the dynamics identifier.

V. Evaluation of BLOS Identification

In this section, we use numerical simulations and data collected through on-road experiments to validate the analytical results and to demonstrate the robustness of our estimators against varying configurations of connected vehicle systems. In particular, we consider a (4+1)-vehicle chain where the CCC vehicle 4 at the tail monitors the motion of vehicles 0 and 3 while the vehicles 1, 2, 3 only respond to the vehicle immediately ahead, as shown in Fig. 4.

First, we use the experimental data for the speed profile of the leading vehicle 0 and then generate synthetic data for vehicles 1, 2, 3 using the optimal velocity model [27]

$$\begin{aligned} \dot{s}_j(t) &= v_j(t), \\ \dot{v}_j(t) &= \alpha_j\big(H_j^{-1}\big(h_{j,j-1}(t-\sigma_j)\big) - v_j(t-\sigma_j)\big) \\ &\quad + \beta_j(v_{j-1}(t-\sigma_j) - v_j(t-\sigma_j)\big), \end{aligned} \tag{35}$$

Fig. 4. A vehicle network where the CCC vehicle 4 monitors the motion of vehicle 3 via range sensors and also receives information from vehicle 0 by V2V communication.

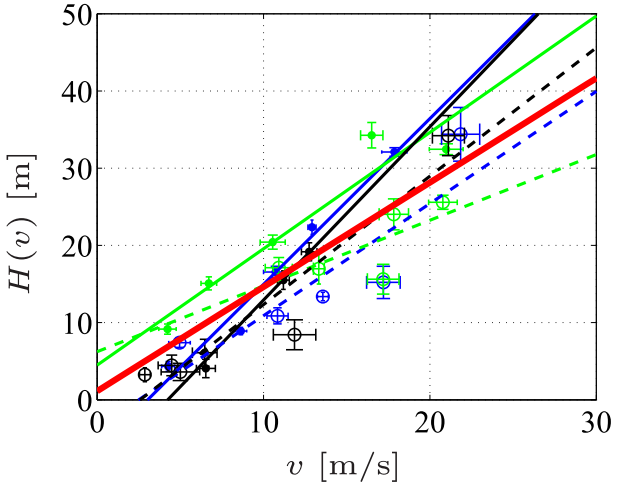


Fig. 5. Data points with error bars represent the equilibrium states with uncertainties, where the solid dots are collected from Experiment I while the circles are collected from Experiment II. The colors blue, green, and black mark the data collected from vehicles 1, 2, 3, respectively. Solid lines and dashed lines are obtained by fitting the linear function (36) to the solid dots and the circles of the corresponding colors. The red solid line represents the average range policy, which is obtained by fitting all the collected data.

for $j = 1, 2, 3$, where the positive constants α_j and β_j are the control gains for the headway and the relative velocity, respectively, and the positive constant σ_j represents the reaction time delay of human-driven vehicle j. The range policy function H_j is given in (15), and its inverse function H_j^{-1} determines the desired velocity according to the distance $h_{j,j-1}$.

In order to obtain the range policies of human drivers, we conducted on-road experiments at a 4.5 [km] long section of Mast road near Dexter, Michigan, USA using four human driven vehicles; see vehicles 0, 1, 2, 3 in Fig. 4. During the experiment, head vehicle 0 set different constant speed values and vehicles 1, 2, 3 kept the same speed while maintaining constant distances according to the drivers' preference. In particular, we conducted two sets of experiments as follows.

- Experiment I: Vehicle 0 gradually increased the values of constant speed.
- Experiment II: Vehicle 0 gradually decreased the values of constant speed.

The collected data are shown in Fig. 5 with the colored dots and the corresponding error bars. In particular, the solid dots mark the data collected during Experiment I while the circles corresponds to Experiment II. The colors blue, green, and black represent vehicles 1, 2, 3, respectively. The linear function

$$H(v) = \kappa v + \rho \tag{36}$$

TABLE I

PARAMETERS, DEFINED IN (36), CORRESPONDING TO THE FITTED LINES IN FIG. 5

Experiment	I (solid)	II (dashed)
Vehicle 1 (blue)	$\kappa = 2.2$ [s] $\rho = -6.5$ [m]	$\kappa = 1.5$ [s] $\rho = -3.6$ [m]
Vehicle 2 (green)	$\kappa = 1.5$ [s] $\rho = 4.5$ [m]	$\kappa = 0.9$ [s] $\rho = 6.2$ [m]
Vehicle 3 (black)	$\kappa = 2.2$ [s] $\rho = -9.5$ [m]	$\kappa = 1.7$ [s] $\rho = -4.3$ [m]

TABLE II

VEHICLE PARAMETERS FOR NUMERICAL SIMULATIONS

Vehicle i	1	2	3
α_i [1/s]	0.2	0.3	0.4
β_i [1/s]	0.4	0.4	0.2
σ_i [s]	0.5	0.4	0.6
l_i [m]	4.8	4.6	4.8
κ_i [s]	1.5	1.5	1.7
ρ_i [m]	-3.6	4.5	-4.3

TABLE III

SCENARIOS TO TEST BLOS IDENTIFICATION ALGORITHMS

Scenario I	Vehicle 1 catches up with vehicle 0 while vehicles 2 and 3 follow vehicle 1.
Scenario II	Vehicles 2 and 3 follow vehicle 0 at the beginning while vehicle 1 cuts in between vehicles 0 and 2.
Experiment	Vehicle 0 varies its speed in time while vehicles 1–3 try to follow it.

is used to fit the data by using the least square approach. The corresponding straight lines (solid for Experiment I and dashed for Experiment II) are shown in Fig. 5 and the results are summarized in Table I. To obtain the average range policy, we also fit the whole data which yields $\kappa_{\text{ave}} = 1.4$ [s] and $\rho_{\text{ave}} = 1.1$ [m] as shown by the red solid line.

The difference in the parameters κ between the two experiments for each driver implies the variability of time headway that one keeps from the vehicle immediately ahead in equilibrium situations. Also, notice that vehicle 2 has considerable smaller κ-s when compared with vehicles 1 and 3. This is attributed to the fact that the driver of vehicle 2 is much more experienced than the other two drivers. We also remark that the experiments were only conducted for middle ranges of the speed so that the saturations for high speed cannot be observed. Nevertheless, Fig. 5 and Table I demonstrates that the range policy can be obtained when collecting data of individual drivers.

Considering the heterogeneity in real traffic, we use different parameters for different vehicles. For the leading vehicle 0, we assume that its length is $l_0 = 4.8$ [m] and it follows a given speed profile. The parameters of vehicles 1, 2, 3 are summarized in Table II. The range policies of vehicles 1, 2, 3 are set by using the fitted lines in Fig. 5; cf. Table I.

To test the performance of the presented BLOS identification algorithms, we consider the two scenarios presented in Table III. We first consider scenario I where the initial distance between vehicles 0 and 1 is 500 [m] the distance between vehicles 2 and 3 and the distance between vehicles 3 and 4

are 40 [m] and 50[m], respectively. The initial velocities of vehicles 1, 2, 3 are all set to be 30 [m/s]. The velocity of vehicle 0 is given by the UMTRI Safety Pilot data [31]. For the window size of the vector (1), we use $\Delta = 500$ that corresponds to $\Delta \cdot \delta t = 50$ [s]. The gain for the distribution update (7) is given by $\gamma = 1$ and the confidence threshold in (11) is set to be $\overline{C} = 0.5$. After the causal relationship is detected, the link-length estimator (25) is activated to identify the number of vehicles between the broadcasting vehicle 0 and the receiving vehicle 4. To find appropriate values for the forgetting factor μ in (25), we conducted a large number of simulations, which show that $0.98 < \mu < 1$ can lead to a balance between the sensitivity and the robustness. For the subsequent simulation, we use $\mu = 0.99$. Moreover, we set $\eta = 0.5$ in (25) for equal weights on v_0 and v_3. To determine the convergence of the estimated link length, we set the time length as $T = 50$ [s]. Once the estimated link length converges, the dynamics identifier (26) is activated. The approximate average vehicle length is set to $\overline{l}_{\text{av}} = 4.7$ [m]. The approximate average range policy is represented by solid red line in Fig. 5. For the cost function (28), we use $c_1 = 0.7$ and $c_2 = 0.2$. For the pool of candidate coefficients, we first randomly generate $M = 10^5$ samples and then apply the k-means clustering approach that leads to $K = 60$ sets of representative coefficients. The maximum iteration steps for solving (29) is set to be 50.

The simulation results corresponding to scenario I are shown in Fig. 6, where vehicles 1, 2, 3 catch up with vehicle 0 ahead. The velocities of vehicles 0, 1, 2, 3 are displayed in Fig. 6(a). Fig. 6(b) shows the output of the causality detector (11), which detects the causality at $t_{\ell_1} = 84.6$ [s]. The time difference between the instant when motion of vehicle 0 begins to influence the motion of vehicle 3 and the instant when the causal relationship is detected is about 50 [s] due to the window size $\Delta = 500$; cf. (1). Then, the link length estimator (25) is activated by using the historical data between $t_{\ell_1} - \Delta \cdot \delta t = 34.6$ [s] and $t_{\ell_1} = 84.6$. Fig. 6(c) shows that the estimated link length (blue dashed) converges to its real value (red solid). The convergence of link-length estimator is also determined at $t_{\ell_2} = t_{\ell_1} = 84.6$ [s], which triggers the activation of the dynamics identifier (26) that identifies the coefficient by using the data between $t_{\ell_2} - \Delta \cdot \delta t = 34.6$ [s] and $t_{\ell_2} = 84.6$ [s]. The coefficients are frozen after t_{ℓ_2}. Fig. 6(d) shows that, after the coefficients are identified, the output of the dynamics identifier tracks the real velocity of vehicle 3, which implies that the identified model given by (26) approximates dynamics and hence can be used for the subsequent CCC design. In terms of our setup, the coefficients can be identified in about 3 [s]. The error between the output of the dynamics identifier and the velocity of vehicle 3, i.e., $|\hat{v}_3 - v_3|$, is characterized in the first row of Table IV.

In practice, the configuration of a connected vehicle system may vary due to lane changes. To test the response of our presented estimator to changing configurations, we consider scenario II where vehicle 1 cuts in between vehicles 0 and 2 at $t = 60$ [s]. For simulation, we assume that vehicle 1 maintains constant speed 10 [m/s] before it cuts in between vehicles 0 and 2, and after the cut-in it responds to the motion of vehicle 0 governed by the dynamics (35). We also assume that vehicle 2 slows down with constant deceleration -0.5 [m/s^2] between $t = 55$ [s] and $t = 60$ [s] to provide enough space for the cut-in of vehicle 1, while vehicle 3 always responds to the motion of vehicle 2 by using the dynamics (35). Fig. 7(a) shows the velocities of

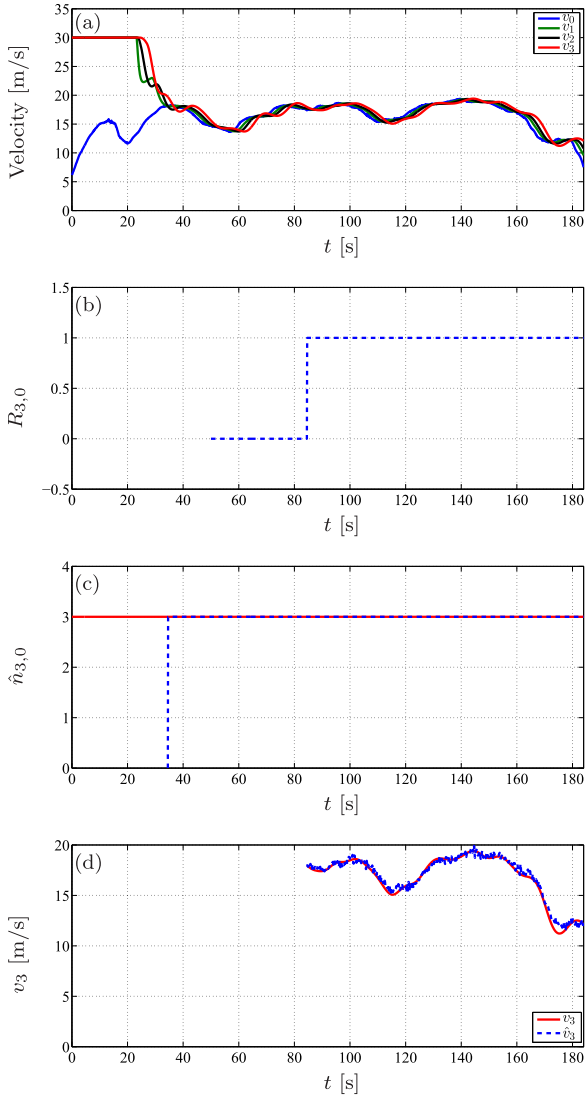


Fig. 6. Simulation results for scenario I where vehicles 1, 2, 3 catch up with vehicle 0. (a) Velocities of vehicles 0–3. (b) The output of the causality indicator (11). (c) The estimated link length (blue dashed) given by (25) converges to its real value (red solid). (d) The output of the dynamics identifier (blue dashed) given by (26) approximates the real velocity of vehicle 3 (red solid), which implies that the identified model can approximate the dynamics of the vehicle chain comprised of vehicles 0–3.

TABLE IV

ERRORS BETWEEN THE OUTPUT OF THE DYNAMICS IDENTIFIER AND THE REAL VELOCITY OF VEHICLE 3, i.e., $|\hat{v}_3 - v_3|$. THE QUANTITIES HAVE THE UNIT [m/s]

Scenario	mean	standard deviation	maximum
Simulation I	0.32	0.28	1.12
Simulation II	0.19	0.21	0.79
Experiment	1.94	1.59	8.95

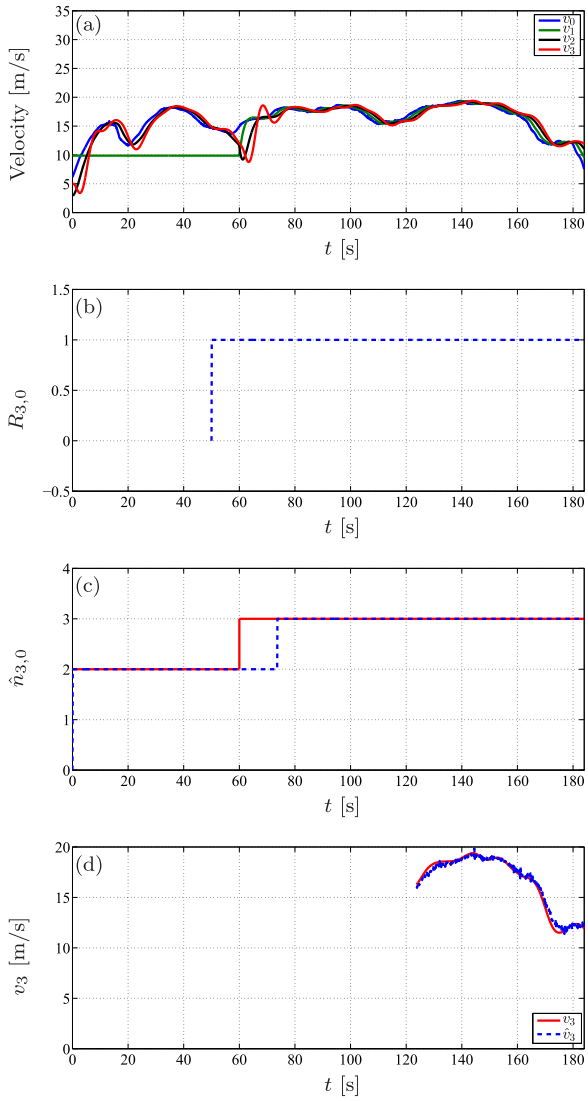


Fig. 7. Simulation results for scenario II where vehicle 1 cuts in between vehicles 0 and 2 at $t = 60$ [s]. (a) Velocities of vehicles 0–3. (b) The output of the causality indicator (11). (c) The estimated link length (blue dashed) given by (25) converges to its real value (red solid), even when the real value changes. (d) The output of the dynamics identifier (blue dashed) given by (26) approximates the real velocity of vehicle 3 (red solid).

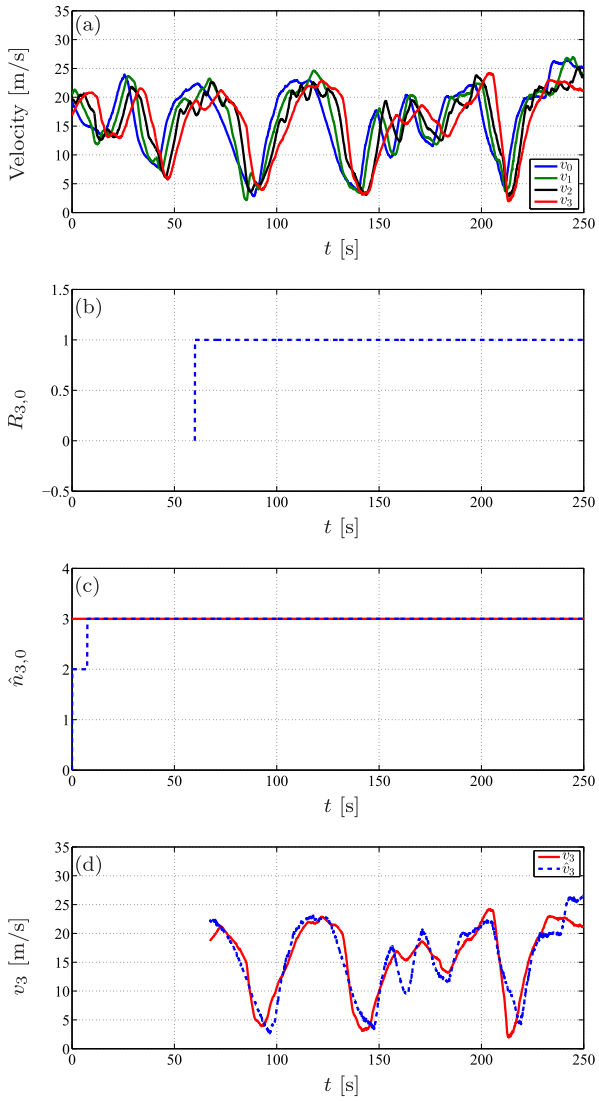


Fig. 8. Application of BLOS identification algorithms to the data collected through on-road experiments. (a) Velocities of vehicles 0–3. (b) The output of the causality indicator (11). (c) The estimated link length (blue dashed) given by (25) converges to its real value (red solid). (d) The output of the dynamics identifier (blue dashed) given by (26) tracks the real velocity of vehicle 3 (red solid) with bounded error.

vehicles 0, 1, 2, 3. Fig. 7(b) indicates the causal relationship between the motion of vehicles 0 and 3. Fig. 7(c) shows that, after the configuration changes, the estimated link length converges to the real value in about 15 [s]. Fig. 7(d) demonstrates that the difference between the identified model and the real nonlinear time-delayed dynamics is small. The corresponding statistics of errors are given in the second row of Table IV.

To evaluate the performance of the presented estimators in real traffic, we use experimental data collected on Mast road using the four vehicles; see vehicles 0–3 in Fig. 4. In this case, vehicle 0 varies its speed in time while vehicles 1–3 try to follow it. Then, we only use the information of vehicles 0 and 3 to estimate the configuration and the dynamics of the corresponding connected vehicle system. To counteract the effects of measurement errors and transient motion of human drivers, we set the window size as $\Delta = 600$ in (1), corresponding to $\Delta \cdot \delta t = 60$ [s]. The results are demonstrated in Fig. 8, where the notations are the same as used in Fig. 6. Fig. 8(b) shows that the causality is detected at $t_{\ell_1} = 60.1$ [s]. This triggers the activation of the link length estimator, which starts to use the historical data between $t_{\ell_1} - \Delta \cdot \delta t = 0.1$ [s] and $t_{\ell_1} = 60.1$ [s]. Fig. 8(c) shows that the estimated link length (blue dashed) converges to the its real value (red solid). The convergence of the estimated link length is determined at $t_{\ell_2} = 67.5$ [s], which triggers the activation of the dynamics identifier that trains the model by using the data between $t_{\ell_2} - \Delta \cdot \delta t = 7.5$ [s] and $t_{\ell_2} = 67.5$ [s]. After the coefficients are identified, the model is frozen. Fig. 8(d) shows that the output of the identified model follows the real velocity of vehicle 3 with bounded errors. This implies that the identified model approximates the dynamics. Comparing Fig. 8(d) with Fig. 6(d) and Fig. 7(d), one may observe that

identification errors for experimental data are larger than the identification errors for simulation data. This is due to the nonlinearities and large reaction time of human driver.

VI. Conclusion

In this paper, we proposed the concept of beyond-line-of-sight (BLOS) identification by exploiting vehicle-to-vehicle (V2V) communication to monitor the motion of distant vehicles. In particular, we focused on addressing three issues that play significant roles in the design of connected cruise control (CCC). First, we presented a causality detector to determine whether the information received from distant vehicles was relevant to the CCC vehicle. Then, we designed a link-length estimator to identify the number of vehicles between the broadcasting vehicle and the CCC vehicle, which was crucial for appropriately utilizing the motion data of broadcasting vehicles. Finally, we constructed a dynamics identifier to approximate the nonlinear time-delayed dynamics of connected vehicle systems, which was required to achieve desired system-level performance. The coefficients of the dynamics identifier were obtained by minimizing the cost and also by ensuring the stability of the model. We proposed a method to solve the stability-constrained optimization problem in an efficient way. The presented analytical results were validated through numerical simulations and on-road experiments. In the future, we will investigate the robustness of the identifier against packet drops in V2V communication. We will also investigate how to improve the efficiency of the dynamics identifier by utilizing advanced clustering algorithms and other optimization approaches.

Appendix
Proof of Theorem 1

For the connected vehicle systems with data given in (1), we have

$$\begin{aligned} s_j - s_i &= s_j^* - s_i^* + \tilde{d}_{i,j}, \\ \eta v_j + (1-\eta) v_i &= v^* + \tilde{v}_{i,j}, \end{aligned} \tag{37}$$

where the uniform flow equilibrium $s_j^* - s_i^* \equiv d_{i,j}^*$ and v^* can be seen as the constant components in the Fourier series of $s_j - s_i$ and $\eta v_j + (1-\eta) v_i$ while the perturbations $\tilde{d}_{i,j}$ and $\tilde{v}_{i,j}$ are composed of all sinusoidal terms of the Fourier series. It follows that

$$\mathbb{E}_{t_0}^{t_k}\big[\tilde{d}_{i,j}(t_p)\big] \to 0, \quad \mathbb{E}_{t_0}^{t_k}\big[\tilde{v}_{i,j}(t_p)\big] \to 0, \tag{38}$$

as $k \to \infty$. Note that measurement errors may also be included in $\tilde{d}_{i,j}$ and $\tilde{v}_{i,j}$.

Substituting (37) and (38) into (21) leads to

$$\begin{aligned} \xi_s(t_k) &\to s_j^* - s_i^*, \\ \xi_v(t_k) &\to v^*, \\ \xi_n(t_k) &\to \frac{s_j^* - s_i^*}{\overline{l}_{\mathrm{av}} + \overline{H}_{\mathrm{av}}(v^*)}, \end{aligned} \tag{39}$$

as $k \to \infty$. Hence, the output of link-length estimator (23) at the limit becomes

$$\lim_{k\to\infty} \hat{n}(t_k) = \mathrm{int}\left(\frac{s_{i-n}^* - s_{i-1}^*}{\overline{l}_{\mathrm{av}} + \overline{H}_{\mathrm{av}}(v^*)}\right). \tag{40}$$

At the uniform flow, we have

$$s_{i-n}^* - s_{i-1}^* = n\big(l_{\mathrm{av}} + H_{\mathrm{av}}(v^*)\big), \tag{41}$$

cf. (16). Substituting (41) into (40) leads to

$$\lim_{k\to\infty} \hat{n}(t_k) = \mathrm{int}\left(\frac{l_{\mathrm{av}} + H_{\mathrm{av}}(v^*)}{\overline{l}_{\mathrm{av}} + \overline{H}_{\mathrm{av}}(v^*)} n\right). \tag{42}$$

When the condition (24) holds, it follows that

$$n - 0.5 < \frac{l_{\mathrm{av}} + H_{\mathrm{av}}(v^*)}{\overline{l}_{\mathrm{av}} + \overline{H}_{\mathrm{av}}(v^*)} n < n + 0.5, \tag{43}$$

cf. (24). Rounding the result to the closest integer yields $\lim_{k\to\infty} \hat{n}(t_k) = n$; cf. (42) and (43).

Acknowledgment

The authors would also like to thank Jin Ge and Commsignia Inc. for their support regarding on-road experiments.

References

[1] *Traffic Safety Facts—Research Note*, Nat. Highway Traffic Safety Admin., Washington, DC, USA, 2014. [Online]. Available: http://www.epa.gov/oms/emisslab/testing/dynamometer.htm

[2] D. Schrank and T. Lomax, "Urban mobility report 2009," Texas A&M Transp. Inst., College Station, TX, USA, Tech. Rep., 2009. [Online]. Available: http://americandreamcoalition.org/highways/mobility_report_2009_wappx.pdf

[3] G. Marsden, M. McDonald, and M. Brackstone, "Towards an understanding of adaptive cruise control," *Transp. Res. C, Emerg. Technol.*, vol. 9, no. 1, pp. 33–51, 2001.

[4] A. Vahidi and A. Eskandarian, "Research advances in intelligent collision avoidance and adaptive cruise control," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 3, pp. 143–153, Sep. 2003.

[5] A. Alam, J. Mårtensson, and K. H. Johansson, "Experimental evaluation of decentralized cooperative cruise control for heavy-duty vehicle platooning," *Control Eng. Pract.*, vol. 38, pp. 11–25, May 2015.

[6] C. R. He and G. Orosz, "Saving fuel using wireless vehicle-to-vehicle communication," in *Proc. Amer. Control Conf.*, 2017, pp. 4946–4951.

[7] D. Tian, J. Zhou, Y. Wang, Z. Sheng, H. Xia, and Z. Yi, "Modeling chain collisions in vehicular networks with variable penetration rates," *Transp. Res. C, Emerg. Technol.*, vol. 69, pp. 36–59, Aug. 2016.

[8] K. C. Dey *et al.*, "A review of communication, driver characteristics, and controls aspects of cooperative adaptive cruise control (CACC)," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 491–509, Feb. 2015.

[9] V. Milanés, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura, "Cooperative adaptive cruise control in real traffic situations," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 296–305, Feb. 2014.

[10] P. Seiler, A. Pant, and K. Hedrick, "Disturbance propagation in vehicle strings," *IEEE Trans. Autom. Control*, vol. 49, no. 10, pp. 1835–1842, Oct. 2004.

[11] Y. Zhao, P. Minero, and V. Gupta, "On disturbance propagation in leader–follower systems with limited leader information," *Automatica*, vol. 50, pp. 591–598, Feb. 2014.

[12] S. Öncü, J. Ploeg, N. van de Wouw, and H. Nijmeijer, "Cooperative adaptive cruise control: Network-aware analysis of string stability," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1527–1537, Aug. 2014.

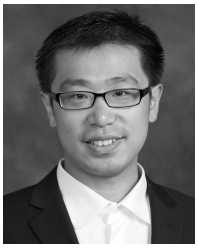
[13] M. di Bernardo, A. Salvi, and S. Santini, "Distributed consensus strategy for platooning of vehicles in the presence of time-varying heterogeneous communication delays," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 102–112, Feb. 2015.

[14] R. Dang, J. Wang, S. E. Li, and K. Li, "Coordinated adaptive cruise control system with lane-change assistance," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2373–2383, Oct. 2015.

[15] B. van Arem, C. J. G. van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 429–436, Dec. 2006.

[16] A. Geiger *et al.*, "Team AnnieWAY's entry to the 2011 grand cooperative driving challenge," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1008–1017, Sep. 2012.

[17] E. Chan, P. Gilhead, P. Jelinek, P. Krejci, and T. Robinson, "Cooperative control of SARTRE automated platoon vehicles," in *Proc. 19th ITS World Congr.*, 2012, pp. 1–9.
[18] L. Zhang and G. Orosz, "Motif-based design for connected vehicle systems in presence of heterogeneous connectivity structures and time delays," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1638–1651, Jun. 2016.
[19] G. Orosz, "Connected cruise control: Modelling, delay effects, and nonlinear behaviour," *Veh. Syst. Dyn.*, vol. 54, no. 8, pp. 1147–1176, 2016.
[20] L. Zhang, J. Sun, and G. Orosz, "Hierarchical design of connected cruise control in the presence of information delays and uncertain vehicle dynamics," *IEEE Trans. Control Syst. Technol.* [Online]. Available: http://ieeexplore.ieee.org/document/7859490/
[21] L. Zhang and G. Orosz, "Consensus and disturbance attenuation in multi-agent chains with nonlinear control and time delays," *Int. J. Robust Nonlinear Control*, vol. 27, no. 5, pp. 781–803, 2017.
[22] W. B. Qin, M. M. Gomez, and G. Orosz, "Stability and frequency response under stochastic communication delays with applications to connected cruise control design," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 2, pp. 388–403, Feb. 2017.
[23] J. I. Ge and G. Orosz, "Optimal control of connected vehicle systems with communication delay and driver reaction time," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 2056–2070, 2017.
[24] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011.
[25] K. Petersen and M. Pedersen, *The Matrix Cookbook*. Lyngby, Denmark: Tech. Univ. Denmark, 2012.
[26] C. M. Bishop, *Pattern Recognition and Machine Learning*. Singapore: Springer, 2006.
[27] G. Orosz, R. E. Wilson, and G. Stépén, "Traffic jams: Dynamics and control," *Philos. Trans. Roy. Soc. A, Math. Phys. Sci.*, vol. 368, no. 1928, pp. 4455–4479, 2010.
[28] L. Zhang and G. Orosz, "Black-box modeling of connected vehicle networks," in *Proc. Amer. Control Conf.*, 2016, pp. 2421–2426.
[29] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.
[30] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, nos. 1–3, pp. 1–6, 1998.
[31] D. Bezzina and J. Sayer, "Safety pilot model deployment: Test conductor team report," Nat. Highway Traffic Safety Admin., Washington, DC, USA, Tech. Rep. DOT HS 812 171, 2015.



Linjun Zhang received the B.Eng. degree in automation from Northeastern University, China, in 2009, the M.Eng. degree in control science and engineering from Beijing University of Aeronautics and Astronautics, China, in 2012, and the Ph.D. degree in mechanical engineering from University of Michigan, Ann Arbor, USA, in 2017. He is currently a Research Engineer for autonomous vehicles with Ford Motor Company. His research interests include intelligent transportation systems, vehicle dynamics and control, nonlinear control, time-delay systems, machine learning, and system identification.

Gábor Orosz received the M.Sc. degree in engineering physics from Budapest University of Technology, Hungary, in 2002, and the Ph.D. degree in engineering mathematics from University of Bristol, U.K., in 2006. He held post-doctoral positions with University of Exeter, U.K., and University of California at Santa Barbara. He joined University of Michigan, Ann Arbor, in 2010, where he is currently an Associate Professor in mechanical engineering. His research interests include nonlinear dynamics and control, time-delay systems, networks and complex systems with applications on connected and automated vehicles, and biological networks.