An Algebraic Evaluation Framework for a Class of Car-Following Models

Zejiang Wang, *Graduate Student Member, IEEE*, Xingyu Zhou, and Junmin Wang, *Senior Member, IEEE*

***Abstract*— Car-following models describe how a driver follows the leading vehicle in the same lane. They serve as the cornerstone of microscopic traffic-flow simulations and play an essential role in analyzing human factors in traffic casualty, congestion, efficiency, and emissions. An extensive and continuously growing number of car-following models in the literature raises the requirement to evaluate and compare different models objectively. Generally, a car-following model is evaluated after model parameter calibration: the optimal residual between the calibrated model output and the measured counterpart is used as a metric to assess a car-following model's performance. However, model parameter calibration, usually formed as a numerical optimization problem, suffers from several issues, such as local optimality and heavy computational burden. More importantly, different formulations of the cost function can lead to distinct calibration outcomes and contradictory conclusions of the model evaluation results. This paper proposes instead a purely algebraic framework for evaluating a class of car-following models whose parameters can be linearly identified. Car-following models with nonlinear relationships among parameters, e.g., the behavioral car-following models, are out of the scope of analysis in this paper. Algebraic manipulations performed on a model finally produce a system error index, which is a uniform metric for evaluating and comparing different car-following models. During the whole process, no cost function needs to be designed a priori, and no computationally expensive numerical optimization is involved. Three car-following models are evaluated and compared under the proposed algebraic framework.**

***Index Terms*— Algebraic estimation, car-following model, model evaluation.**

I. Introduction

CAR-FOLLOWING model describes a driver's longitudinal control maneuver, which is commonly specified as vehicle acceleration [1], with respect to the movement of the preceding vehicle in the same lane. Together with the lane-changing model, they serve as the cornerstone of microscopic traffic-flow simulation [2], [3]. Especially, car-following models play an important role in analyzing human factors in traffic casualty [4], urban congestion [5], as well as fuel consumption and emissions [6].

Because of their importance, car-following models have been intensively studied since the 1950s, and there exist an extensive and continuously increasing number of car-following models in the literature [4], [7]–[21]. Such an abundance evokes the desire to evaluate and compare different models objectively. Generally, a car-following model is evaluated after model parameter calibration, which is the process of finding the best-fitted parameters to minimize the residual between the modeled outputs and the real measurements [22], [23]. Mathematically, it is required to find $\theta^* = \underset{\theta_{lb} \le \theta \le \theta_{ub}}{\arg\min}\ J\left(X - \hat{X}(\theta)\right)$, with the parameter set θ as the optimized variable, θ_{ub} and θ_{lb} as the upper and lower bounds of θ, X being the measured state of interest, $\hat{X}(\theta)$ being the modeled state of interest, and J being the function of goodness of fit. By substituting θ^* back into the cost function J, we finally obtain the optimal model residual as: $J^* = J\left(X - \hat{X}(\theta^*)\right)$. The smaller the J^*, the more performant a car-following model is.

However, car-following model calibration is a more complicated task than it seems [24], [25]. Firstly, the state of interest X needs to be selected. There exist several candidates, such as the longitudinal acceleration [26], speed [15], [23], or position of the ego-car [27], inter-vehicle distance [28], [29], time headway between the ego-car and its predecessor [22], and a mix of several attributes mentioned above [30]. Secondly, the cost function can be formulated in different ways, including the squared error, mean absolute error, root mean squared error, root mean squared logged error, and Theil's inequality coefficient [31]. Thirdly, an efficient *numeric* solver shall be chosen to solve the constrained nonlinear optimization problem. In practice, the mathematical relationship between the optimized variable θ and the cost function J is usually too complex to formulate explicitly. Therefore, gradient-free solvers, such as genetic algorithm [1], [23], [32], [33], downhill simplex method [16], [22] and Box's complex algorithm [34] are employed. Note that none of them can guarantee global optimality. Instead, local optimality, which strongly depends on the initial optimization point, can frequently occur [31]. Authors in [24] pointed out that sometimes, a seemingly promising goodness-of-fit value can be indeed associated with the optimized variables far away from their actual values.

With the three issues mentioned above, it is not strange to find conflicting conclusions of car-following model evaluation results in the literature. For instance, with the state of interest X chosen as the ego car's longitudinal speed and the goodness of fit J selected as the absolute error, authors in [33] concluded that the IDM model [11] could better fit the recorded trajectory data than the Gipps' model [19], [20]. In contrast, by fixing the cost function as the squared error,

Manuscript received April 7, 2021; revised July 14, 2021 and August 22, 2021; accepted September 15, 2021. This work was supported in part by the National Science Foundation under Award 1901632. The Associate Editor for this article was M. Brackstone. *(Corresponding author: Junmin Wang.)*

The authors are with the Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: jwang@austin.utexas.edu).

Digital Object Identifier 10.1109/TITS.2021.3113788

authors in [27] instead asserted that the IDM model should be the least possible car-following model adopted by human beings. Contradictions can also be found in [35] versus [32], [36] versus [26], as examples. As the impact of emerging technologies, e.g., vehicle-to-vehicle (V2V) communication, on transportation systems is frequently analyzed via traffic-flow simulation [5], objectively evaluating the performance of a car-following model serves as a preliminary condition to gauge the effectiveness of novel intelligent transportation systems.

Because the standard framework for *numerically* evaluating different car-following models is lacking, we propose an *algebraic* approach to evaluate and compare a class of car-following models whose parameters can be linearly identified. With algebraic manipulations performed on a given model, model parameters are first expressed in a linearly identifiable form. Then, a least-squares (LS) problem is formulated to solve the underdetermined equation *analytically*. The LS problem's residual error is subsequently converted into parameter error index (PEI) and system error index (SEI). PEI is utilized to gauge the accuracy of the estimated model parameters. More importantly, SEI serves as a uniform metric to evaluate and compare different car-following models. During the whole process, no cost function J needs to be determined *a priori*, and the calculations are purely algebraic without computationally intensive numerical optimization. CHM model [7], GHR model [8], and Edie model [9] are selected as benchmarks to illustrate the proposed algebraic model evaluation framework. Note that because of the nonlinear relationships among parameters, behavioral car-following models, e.g., IDM model [11] and Gipps' model [19], [20], cellular automation based model [12], and psychophysical model [18] cannot be analyzed under the proposed algebraic framework.

This paper principally adopts the longitudinal speed as the state of interest for evaluating and comparing the three models. However, as stated in [28], [29], if one state of interest is selected, the cumulative measurement variable, such as vehicle position or inter-vehicle spacing, is preferable to speed or acceleration. Analysis based on cumulative variables will be covered in our future work.

The rest of this paper is organized as follows: Section II describes the three car-following models under investigation. Section III explains how to reformulate the car-following models to express the model parameters in a linearly identifiable form, which yields an underdetermined equation. To solve the underdetermined equation in Section III, Section IV formulates an LS problem and introduces the concept of residual error, PEI, and SEI. Subsequently, Section V applies synthetic simulation data to illustrate the usage of PEI and SEI. After that, Section VI uses the NGSIM vehicle trajectory dataset to evaluate and compare the three car-following models. Finally, Section VII concludes this paper.

II. Car-Following Models Under Investigation

This section illustrates the three car-following models that we will analyze under the proposed algebraic framework.

CHM model [7]:

$$\dot{v}_x(t) = c\,\Delta v_x(t - T_r), \tag{1}$$

GHR model [8]:

$$\dot{v}_x(t) = c\frac{\Delta v_x(t - T_r)}{\Delta x(t - T_r)}, \tag{2}$$

and Edie model [9]:

$$\dot{v}_x(t) = c v_x(t)\frac{\Delta v_x(t - T_r)}{\Delta x^2(t - T_r)}. \tag{3}$$

In (1), (2), and (3), v_x indicates the longitudinal velocity of the ego-car. Δv_x is the speed difference between the ego-car and the preceding vehicle, as $\Delta v_x = v_{x-1} - v_x$, where v_{x-1} is the longitudinal velocity of the preceding car. Similarly, $\Delta x = x_{-1} - x$ is the inter-vehicle distance gap, with x_{-1} and x being the positions of the preceding car and the ego-car, respectively. Two *constant* parameters include the sensitivity gain c and the driver reaction delay T_r. Inter-driver difference explains that each driver has a personalized sensitivity gain and reaction delay [30], [37].

Remark 1: Even though the car-following models (1), (2), (3) prescribe constant parameters c and T_r, realistic car-following behavior of a specific driver may also vary [38], [39] because of the different traffic situations. This phenomenon is termed as intra-driver difference. In the following, we howbeit obey the assumption of constant parameters, as we want to evaluate the three models *per se*, instead of proposing new car-following models with time-varying parameters.

Models (1), (2), and (3) are selected for three reasons. Firstly, albeit simple, they serve as the basis for many modern car-following models [17]. Secondly, their performances have been evaluated and compared, and contradictory conclusions exist in the literature. For instance, in [15] and [40], by assigning the goodness of fit J as the averaged root mean square and the state of interest X as speed, acceleration, and inter-vehicle distance, the authors stated that the simplest CHM model (1) could outperform the other two models. Conversely, by optimizing the squared errors, authors in [26] concluded that none of the three models could be regarded as more capable of fitting the trajectory data than the others. Thirdly, the three models have the same number of parameters, which implies the same level of structural complexity and flexibility. This desirable feature yields a fair comparison and avoids the situation that models with more parameters are favored thanks to the adaptability brought from extra parameters. Note that at the end of Section III, we explain that several other car-following models can also be analyzed under the proposed algebraic framework.

III. Car-Following Model Reformulation

As a preliminary to derive the system error index (SEI) in Section IV, we first reformulate a car-following model, such that the model parameters are expressed in a linearly identifiable form. The workflow is illustrated in Fig. 1, where the first and last blocks are executed in the time domain, whereas the others are conducted in the frequency domain.

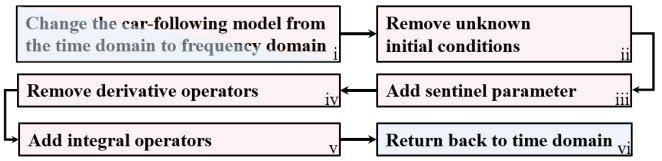

Fig. 1. Car-following model reformulation.

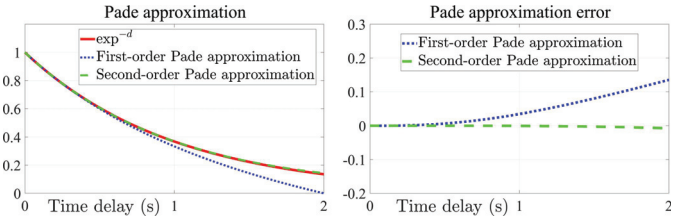

Fig. 2. Padé approximation.

We use the CHM model (1) as an example. We first convert (1) from the time domain to the frequency domain, which corresponds to block i in Fig. 1. To do so, we perform Laplace transform on (1), which yields:

$$sv_x(s) - v_x(0) = c\Delta v_x(s)\, e^{-T_r s}. \tag{4}$$

Then, following block ii in Fig. 1, we remove the unknown initial speed $v_x(0)$ in (4) by differentiating both sides of (4) with respect to the Laplace variable s and obtain:

$$v_x(s) + s\frac{dv_x(s)}{ds} = c\frac{d\Delta v_x(s)}{ds}e^{-T_r s} - cT_r\Delta v_x(s)\, e^{-T_r s}. \tag{5}$$

To continue our analysis, we replace the reaction delay $e^{-T_r s}$ with its approximated polynomial form. As indicated in [10], human reaction delay during a car-following process can reach over 2 seconds. Hence, the second-order Padé approximation is employed to reduce the approximation error. The enhancement of adopting the second-order Padé approximation over the first-order Padé approximation is shown in Fig. 2.

Second-order Padé approximation reads:

$$e^{-T_r s} \approx \frac{12 - 6T_r s + T_r^2 s^2}{12 + 6T_r s + T_r^2 s^2}. \tag{6}$$

Substituting (6) back into (5), we can express (5) as:

$$v_x(s) + s\frac{dv_x(s)}{ds} = \left(c\frac{d\Delta v_x(s)}{ds} - cT_r\Delta v_x(s)\right) \times \frac{12 - 6T_r s + T_r^2 s^2}{12 + 6T_r s + T_r^2 s^2}, \tag{7}$$

which can be equivalently formulated as:

$$\begin{aligned} &12\left(v_x(s) + s\frac{dv_x(s)}{ds}\right) \\ &= -6T_r\left(sv_x(s) + s^2\frac{dv_x(s)}{ds}\right) + c\left(12\frac{d\Delta v_x(s)}{ds}\right) \\ &\quad - T_r^2\left(s^2 v_x(s) + s^3\frac{dv_x(s)}{ds}\right) \\ &\quad - cT_r\left(6s\frac{d\Delta v_x(s)}{ds} + 12\Delta v_x(s)\right) \\ &\quad + cT_r^2\left(s^2\frac{d\Delta v_x(s)}{ds} + 6s\Delta v_x(s)\right) \\ &\quad - cT_r^3\left(s^2\Delta v_x(s)\right). \end{aligned} \tag{8}$$

Afterward, per block iii in Fig. 1, we extend (8) with a sentinel parameter b_x [41] by multiplying the left side of (8) with $s + b_x$ and the right side of (8) with $s + 1$, as:

$$\begin{aligned} &12\left(sv_x(s) + s^2\frac{dv_x(s)}{ds}\right) + b_x\left(12v_x(s) + 12s\frac{dv_x(s)}{ds}\right) \\ &= -T_r\left(6sv_x(s) + 6s^2\left(\frac{dv_x(s)}{ds} + v_x(s)\right) + 6s^3\left(\frac{dv_x(s)}{ds}\right)\right) \\ &\quad - T_r^2\left(s^2 v_x(s) + s^3\left(\frac{dv_x(s)}{ds} + v_x(s)\right) + s^4\left(\frac{dv_x(s)}{ds}\right)\right) \\ &\quad + c\left(12\frac{d\Delta v_x(s)}{ds} + 12s\frac{d\Delta v_x(s)}{ds}\right) \\ &\quad - cT_r^3\left(s^2\Delta v_x(s) + s^3\Delta v_x(s)\right) \\ &\quad - cT_r\left(12\Delta v_x(s) + s\left(6\frac{d\Delta v_x(s)}{ds} + 12\Delta v_x(s)\right)\right. \\ &\quad \left. + s^2\left(6\frac{d\Delta v_x(s)}{ds}\right)\right) \\ &\quad + cT_r^2\left(6s\Delta v_x(s) + s^2\left(\frac{d\Delta v_x(s)}{ds} + 6\Delta v_x(s)\right)\right. \\ &\quad \left. + s^3\left(\frac{d\Delta v_x(s)}{ds}\right)\right). \end{aligned} \tag{9}$$

Note that b_x can be regarded as an extended state of the car-following model, and it is estimated along with the rest of the original parameters inside the car-following model. However, unlike the other unknown parameters, its true value is known *a priori*, as $b_x = 1$. Therefore, by observing the closeness between the estimated sentinel parameter $\hat{b}_x$ and its true value, we can infer the status of the other estimated parameters [41]. The usage of the sentinel parameter will be demonstrated in Section IV.

Next, following block iv in Fig. 1, we multiply s^{-5} on both sides of (9) and move all the parameters to the right side, as:

$$\begin{aligned} &12\left(s^{-4}v_x(s) + s^{-3}\frac{dv_x(s)}{ds}\right) \\ &= -b_x\left(12s^{-5}v_x(s) + 12s^{-4}\frac{dv_x(s)}{ds}\right) \\ &\quad - T_r\left(6s^{-4}v_x(s) + 6s^{-3}\left(\frac{dv_x(s)}{ds} + v_x(s)\right)\right. \\ &\quad \left. + 6s^{-2}\left(\frac{dv_x(s)}{ds}\right)\right) \\ &\quad - T_r^2\left(s^{-3}v_x(s) + s^{-2}\left(\frac{dv_x(s)}{ds} + v_x(s)\right)\right. \\ &\quad \left. + s^{-1}\left(\frac{dv_x(s)}{ds}\right)\right) \\ &\quad + c\left(12s^{-5}\frac{d\Delta v_x(s)}{ds} + 12s^{-4}\frac{d\Delta v_x(s)}{ds}\right) \end{aligned}$$

$$
\begin{aligned}
&- cT_r\left(12s^{-5}\Delta v_x(s) + s^{-4}\left(6\frac{d\Delta v_x(s)}{ds} + 12\Delta v_x(s)\right)\right.\\
&\left.+ s^{-3}\left(6\frac{d\Delta v_x(s)}{ds}\right)\right)\\
&+ cT_r^2\left(6s^{-4}\Delta v_x(s) + s^{-3}\left(\frac{d\Delta v_x(s)}{ds} + 6\Delta v_x(s)\right)\right.\\
&\left.+ s^{-2}\left(\frac{d\Delta v_x(s)}{ds}\right)\right)\\
&- cT_r^3\left(s^{-3}\Delta v_x(s) + s^{-2}\Delta v_x(s)\right). \qquad (10)
\end{aligned}
$$

Comparing (10) and (9), we notice that the derivate operators $s^{\alpha}, \alpha \in \mathrm{N}^{+}$, in (9) are all eliminated. Eliminating derivative operators is crucial because derivative operators can amplify the measurement noises in the time domain. Moreover, we multiply s^{-5}, instead of s^{-4}, on both sides of (9) such that all the measured variables, i.e., v_x and Δv_x, are integrated at least once (see block v in Fig. 1) after we return (10) from the frequency domain back to the time domain. The (iterated) integrations serve as low-pass filters to wipe out high-frequency measurement noises [42], [43].

Recall that the inverse Laplace transform reads:

$$
\begin{cases}
L^{-1}\left(\dfrac{dX(s)}{ds}\right) = -tx(t), \\
L^{-1}\left(\dfrac{X(s)}{s}\right) = \displaystyle\int_0^t x(\tau)\, d\tau.
\end{cases} \qquad (11)
$$

Following block vi in Fig. 1 and applying (11) on (10), we can obtain a linear equation in the time domain, in terms of the model parameters, as:

$$
P_{CHM}(t)^T \Theta = q_{CHM}(t), \qquad (12)
$$

where $P_{CHM}(t) = \left[p^1_{CHM}(t), p^2_{CHM}(t), \ldots, p^7_{CHM}(t)\right]^T$, and the parameter vector is $\Theta = \left[-b_x, -T_r, -T_r^2, c, -cT_r, cT_r^2, -cT_r^3\right]^T$.

The elements in the vector $P_{CHM}(t)$ are:

$$
p^1_{CHM}(t) = 12\left(\int^5 v_x(t) + \int^4 -tv_x(t)\right), \qquad (13)
$$

$$
\begin{aligned}
p^2_{CHM}(t) = 6\Bigg(&\int^4 v_x(t) + \int^3 -tv_x(t) + v_x(t)\\
&+ \int^2 -tv_x(t)\Bigg), \qquad (14)
\end{aligned}
$$

$$
p^3_{CHM}(t) = \int^3 v_x(t) + \int^2 -tv_x(t) + v_x(t) - \int tv_x(t), \qquad (15)
$$

$$
\begin{aligned}
p^4_{CHM}(t) = 12\Bigg(&\int^5 -t\Delta v_x(t)\\
&+ \int^4 -t\Delta v_x(t)\Bigg), \qquad (16)
\end{aligned}
$$

$$
\begin{aligned}
p^5_{CHM}(t) = 12&\int^5 \Delta v_x(t) + \int^4 -6t\Delta v_x(t) + 12\Delta v_x(t)\\
&+ \int^3 -6t\Delta v_x(t), \qquad (17)
\end{aligned}
$$

$$
\begin{aligned}
p^6_{CHM}(t) = 6&\int^4 \Delta v_x(t) + \int^3 6\Delta v_x(t) - t\Delta v_x(t)\\
&+ \int^2 -t\Delta v_x(t), \qquad (18)
\end{aligned}
$$

$$
p^7_{CHM}(t) = \int^3 \Delta v_x(t) + \int^2 \Delta v_x(t), \qquad (19)
$$

along with

$$
q_{CHM}(t) = 12\left(\int^4 v_x(t) + \int^3 -tv_x(t)\right). \qquad (20)
$$

Note that we use $\int^{(n)}\phi(t)$ to represent the iterated integrals $\int_{t_0}^{t}\int_{t_0}^{\sigma_1}\cdots\int_{t_0}^{\sigma_{n-1}}\phi(\sigma_n)\, d\sigma_n\cdots d\sigma_1$, with $t_0 = 0$. If $n = 1$, we simplify $\int^{(1)}\phi(t)$ as $\int\phi(t)$.

Remark 2: To adopt cumulative measurement variables, such as vehicle positions or inter-vehicle gap for comparing and evaluating different models [28], [29], we can reformulate (1) as $\ddot{x}(t) = c\Delta\dot{x}(t - T_r)$. Note that in this case, two unknown initial conditions: $x(0), \dot{x}(0)$ would need to be removed when the model is converted to the frequency domain.

Equation (12) expresses the CHM model parameters in a linearly identifiable form. Indeed, both the GHR model and the Edie model, albeit nonlinear, can also yield a similar equation.

For the GHR model (2), we assign $\eta(t) = \Delta v_x(t)/\Delta x(t)$, obtaining:

$$
\dot{v}_x(t) = c\eta(t - T_r). \qquad (21)
$$

Following (4)-(11), we shall be able to reach (12) again, as:

$$
P_{GHR}(t)^T \Theta = q_{GHR}(t), \qquad (22)
$$

where $q_{GHR}(t) = q_{CHM}(t)$ in (20), and $P_{GHR}(t)$ can be generated from $P_{CHM}(t)$ by replacing the term $\Delta v_x(t)$ in (13)-(19) by $\eta(t)$, if applicable.

Similarly, for the Edie model in (3), we first shift the velocity $v_x(t)$ to the left side of the equation, yielding:

$$
\dot{v}_x(t)/v_x(t) = c\gamma(t - T_r), \qquad (23)
$$

with $\gamma(t) = \Delta v_x(t)/\Delta x^2(t)$. Defining $\omega(t) = \ln(v_x(t))$, we have naturally:

$$
\dot{\omega}(t) = \dot{v}_x(t)/v_x(t). \qquad (24)
$$

Therefore, (23) can be equivalently written as:

$$
\dot{\omega}(t) = c\gamma(t - T_r). \qquad (25)
$$

Contrasting (25) with (1), it becomes clear that the parameters inside the Edie model are also linearly identifiable, as:

$$
P_{Edie}(t)^T \Theta = q_{Edie}(t), \qquad (26)
$$

where $P_{Edie}(t)$ and $q_{Edie}(t)$ can be generated from $P_{CHM}(t)$ and $q_{CHM}(t)$ by replacing $\Delta v_x(t)$ and $v_x(t)$ in (13)-(20) with $\gamma(t)$ and $\omega(t)$, respectively.

It should be noted that a class of car-following models, whose parameters are linearly identifiable, can be analyzed under the following algebraic framework. Noteworthy examples include the optimal velocity with relative velocity model [13], Newell model [14], adaptive time-gap model [16],

spring-mass-damper model [21], to name a few. However, it is recommended to first check if the generated vectors Θ contain practically an equal number of elements inside to evaluate different car-following models fairly. This requirement eliminates the situation that one (more complicated) model is favored not because of its capability to reflect the traffic flow's authentic dynamics, but due to the adaptability brought from the extra parameters.

IV. Least-Squares Problem, Residual Error, and Error Index

Once a car-following model has been reformulated such that the parameters are expressed in a linearly identifiable form like (12), we can then construct a least-squares (LS) problem to solve it *analytically*. The residual error of the LS problem serves as the basis for deriving the system error index, which is the uniform metric to evaluate and compare different car-following models.

A. Least-Squares Problem Formulation

In what follows, we unify the linearly identified forms (12), (22), and (26) as:

$$P(t)^T \Theta = q(t), \tag{27}$$

where $P(t)$ and Θ are two column vectors, and $q(t)$ is a scalar.

To solve the underdetermined Eq. (27), we can formulate an LS problem, as:

$$\Theta^* = \arg\min_{\Theta} J(\Theta, t), \tag{28}$$

where

$$J(\Theta, t) = \int_0^t \left(P^T(\tau)\Theta - q(\tau)\right)^2 d\tau. \tag{29}$$

Defining

$$M_{Pq} = \int_0^t P(\tau) q(\tau) d\tau, \tag{30}$$

$$M_{qP} = M_{Pq}^T = \int_0^t q(\tau) P^T(\tau) d\tau, \tag{31}$$

$$M_{PP} = \int_0^t P(\tau) P^T(\tau) d\tau, \tag{32}$$

and

$$M_{qq} = \int_0^t q(\tau) q(\tau) d\tau, \tag{33}$$

we can express (29) as:

$$\begin{aligned} J(\Theta, t) &= M_{qq} - M_{qP}\Theta - \Theta^T M_{Pq} + \Theta^T M_{PP}\Theta \\ &= \left(\Theta - (M_{PP})^{-1} M_{Pq}\right)^T M_{PP} \\ &\quad \times \left(\Theta - (M_{PP})^{-1} M_{Pq}\right) \\ &\quad + M_{qq} - M_{qP}(M_{PP})^{-1} M_{Pq}. \end{aligned} \tag{34}$$

Therefore, the optimal solution to minimize (29) can be analytically calculated as:

$$\Theta^* = (M_{PP})^{-1} M_{Pq}, \tag{35}$$

with M_{PP} in (36), as shown at the bottom of the next page, and

$$\begin{aligned} M_{Pq} &= \int_0^t P(\tau) q(\tau) d\tau \\ &= \left[\int_0^t p_1(\tau) q(\tau) d\tau, \int_0^t p_2(\tau) q(\tau) d\tau, \ldots, \right. \\ &\quad \left. \int_0^t p_n(\tau) q(\tau) d\tau\right]^T. \end{aligned} \tag{37}$$

Specifically for the models (1), (2), and (3), we have $\Theta^T = \left[-b_x, -T_r, -T_r^2, c, -cT_r, cT_r^2, -cT_r^3\right]$(see (12)). Therefore, with Θ^*, we can finally obtain $-\hat{b}_x = \Theta^*(1)$, $-\hat{T}_r = \Theta^*(2)$, and $\hat{c} = \Theta^*(4)$. Note that the estimated parameter set Θ^* in (35) is derived from algebraic manipulations performed on the model equation itself. In contrast to the Luenberger observer [44] or Kalman Filter [45], algebraic parameter identification does not rely on the Lyapunov stability theory [46]. As a result, there does not exist an asymptotical convergence phase for identification. Instead, Θ^* can be directly determined when the matrix M_{PP} in (35) becomes invertible [47]. In other words, we can obtain the estimated parameter vector Θ^* in (35) if and only if the matrix M_{PP} is non-singular and well-conditioned. However, the matrix M_{PP} is singular at $t = 0$, as all its elements are zero at that moment. Therefore, immediately after $t = 0$, the identification results will demonstrate strong oscillations. However, as M_{PP} is positive semi-definite, we can reasonably conjecture that its numerical conditioning can substantially improve as time goes by. Once the minimum absolute eigenvalue of M_{PP} becomes far away from zero, Θ^* can be obtained. The required period for M_{PP} becoming non-singular is hard to estimate, as it can be influenced by various factors, such as the excitation level of the system inputs, system sampling rate, and signal-to-noise ratio (SNR) [48].

Remark 3: If the matrix M_{PP} in (35) remains singular during a long period, we will not be able to accurately estimate Θ^*. Based on (13)-(19), we can straightforwardly identify two trivial cases yielding consistently singular M_{PP} for the CHM model: *i).* $v_x(t) = 0$, and *ii).* $\Delta v_x(t) = 0$. The first case implies that the ego-car stands still, which is irrelevant in reality. The second case suggests the relative speed between the ego-car and the leading car always remains zero, e.g., both cars run at an identical speed, which implies no interaction between the ego-car and its predecessor. It is recommended to check if the recorded data would yield a singular M_{PP} during a long period before applying the proposed algebraic method.

The sentinel parameter b_x introduced in Section III serves as an indicator to inform when we can assume that the rest of the identified parameters are close enough to their true values. If the estimated sentinel parameter $\hat{b}_x$ remains close to its true value $b_x = 1$, we shall conjecture that $\hat{T}_r = -\Theta^*(2)$ and $\hat{c} = \Theta^*(4)$ are also accurately estimated [41].

Mathematically speaking, we output $-\Theta^*(2)$ and $\Theta^*(4)$ as the estimated reaction time and sensitivity gain when the following criteria are satisfied:

$$\begin{cases} \left|\sigma\left(\hat{b}_x\right) \Big/ E\left(\hat{b}_x\right)\right| \le \Delta_{\text{var}}, \\ \left|\hat{b}_x - 1\right| \le \Delta_{err}. \end{cases} \tag{38}$$

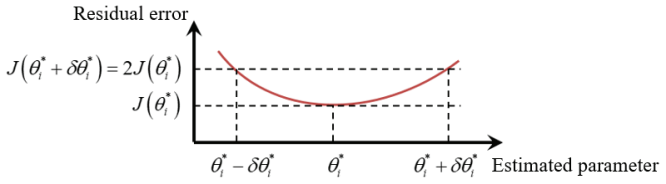

Fig. 3. Parameter error index.

In (38), $E\left(\hat{b}_x\right) = \left(\int_{t-T_{sw}}^{t} \hat{b}_x(\tau)d\tau\right)\Big/T_{sw}$, with T_{sw} as the sliding window length, is the moving-averaged identified sentinel parameter $\hat{b}_x$. $\sigma\left(\hat{b}_x\right) = \sqrt{E\left(\hat{b}_x^2\right) - E\left(\hat{b}_x\right)^2}$ is the standard deviation of $\hat{b}_x$, and $\left|\sigma\left(\hat{b}_x\right)\Big/E\left(\hat{b}_x\right)\right|$ is the sentinel parameter's *fluctuation metric*. Finally, Δ_{var} and Δ_{err} are thresholds. The first requirement in (38) indicates that the estimated sentinel parameter $\hat{b}_x$ has entered into its steady state. The second requirement guarantees the estimation accuracy of $\hat{b}_x$. Selecting the thresholds Δ_{var} and Δ_{err} is a trade-off between estimation precision and speed. Smaller thresholds could yield more accurate estimation, but at the cost of a longer estimation period, and vice versa [48]. In Section V, we fixed $\Delta_{\text{var}} = 1e-5$, and $\Delta_{err} = 1e-2$.

B. Residual Error and Error Indices

Substituting Θ^* in (35) back to (34), we obtain:

$$J\left(\Theta^*, t\right) = M_{qq} - M_{qP}\left(M_{PP}\right)^{-1} M_{Pq}. \tag{39}$$

We term (39) hereafter as the *residual error*. As illustrated below, the residual error serves as a basis for deriving the parameter error index (PEI) and the system error index (SEI).

The theoretical estimation error of a parameter θ_i^* in Θ^* cannot be directly measured, because the true parameter value is unknown. Instead, the parameter error index (PEI) provides an *indirect* assessment of the estimation error. The idea of PEI is exhibited in Fig. 3.

PEI of a *specific parameter* is defined as [49]:

$$\delta\theta_i = \max \delta\theta_i^*, \tag{40}$$

such that

$$J\left(\Theta^* + \delta\theta_i^*, t\right) = 2\left(M_{qq} - M_{qP}\left(M_{PP}\right)^{-1} M_{Pq}\right). \tag{41}$$

Alternatively speaking, PEI is the maximum variation of θ_i^*, such that the residual error $J(\Theta^*, t)$ is doubled. Note that in (41), the term $\Theta^* + \delta\theta_i^*$ is expressed in shorthand and should be understood as the sum of two vectors, where the second vector is $\left[0, \ldots \delta\theta_i^*, \ldots 0\right]^T$.

Substituting (34) and (35) into (41), we can simplify (41) as:

$$\delta\theta_i^{*T} M_{PP} \delta\theta_i^* = M_{qq} - M_{qP}\left(M_{PP}\right)^{-1} M_{Pq}. \tag{42}$$

The optimization problem (40) under the constraint (42) can be *analytically* solved with a Lagrange multiplier, and the solution, termed as PEI of the estimated parameter θ_i^*, is:

$$\delta\theta_i = \sqrt{\left(M_{qq} - M_{qP}\left(M_{PP}\right)^{-1} M_{Pq}\right)\Big/\left(M_{PP}^i\right)}, \tag{43}$$

where M_{PP}^i is the *ith* diagonal element of the matrix M_{PP}.

If $\delta\theta_i$ is large, which implies that the estimated θ_i^* can vary significantly without influencing the residual error (39) too much (less than doubled), then we shall suspect the estimated θ_i^*. Because any estimation between θ_i^* and $\theta_i^* + \delta\theta_i$ would practically yield similar residual errors in this case. Instead, if $\delta\theta_i$ is small, which means that the residual error (39) is quite sensitive to the change of θ_i^*, then we could reasonably deduce that θ_i^* is estimated with more accuracy. Remember that PEI is not the real estimation error. On the contrary, it can reflect the minimum expected orders of magnitude of the estimation error [50]. For instance, if $\delta\theta_1$ of the estimated parameter θ_1^* is much larger than $\delta\theta_2$ of the estimated parameter θ_2^*, then we may conclude with confidence that θ_2^* is more accurately estimated than θ_1^*. Indeed, if $\delta\theta_1$ is only slightly larger than $\delta\theta_2$, then no conclusion with respect to the relative accuracy of the estimated parameters θ_1^* and θ_2^* can be drawn.

In addition to the PEI, which reflects the estimation accuracy of a specific *parameter*, the residual error in (39) can also indicate the structural correctness, and therefore, the performance of a car-following *model* in fitting the measurement data. Note that the residual error is attributable to two parts: model inaccuracy and measurement noises. As mentioned in Section III (see Eq. (10)), iterated integrations serve as low-pass filters to wipe out high-frequency measurement noises. Therefore, the residual error is principally contributed by the model inaccuracy.

The algebraic estimation provides a static and exact formulation of the model parameters via algebraic manipulations performed on the model equation [47]. Therefore, if a car-following model can sufficiently reflect a driver's actual maneuver, the filtered measurement data would accurately fit the linearly identifiable form in (27). In this case, the estimated parameter set Θ^* in (35) shall be very close to the true parameters of this model, and the residual error in (39) would

$$\begin{aligned} M_{PP} &= \int_0^t P(\tau) P^T(\tau) d\tau \\ &= \begin{bmatrix} \int_0^t p_1(\tau) p_1(\tau) d\tau & \int_0^t p_1(\tau) p_2(\tau) d\tau & \cdots & \int_0^t p_1(\tau) p_n(\tau) d\tau \\ \int_0^t p_2(\tau) p_1(\tau) d\tau & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \int_0^t p_n(\tau) p_1(\tau) d\tau & \cdots & \cdots & \int_0^t p_n(\tau) p_n(\tau) d\tau \end{bmatrix}, \end{aligned} \tag{36}$$

be close to zero. However, a car-following model is only an approximation of the complicated human behavior in reality. Assuming the modeling error is condensed as an additional term $q'(t)$, we can modify (27) as $P(t)^T \Theta = q(t) + q'(t)$. With $q'(t)$, Eq. (35) still yields the best-fitted estimated parameters, in the sense of the least squares. However, this estimation will diverge from the true value, which then substantially augments the residual error (39) as a consequence.

The system error index (SEI) is defined as [49], [51]:

$$SEI(t) \sqrt{\frac{M_{qq} - M_{qP}(M_{PP})^{-1} M_{Pq}}{M_{qq}}} = \sqrt{\frac{J(\Theta,t)|_{\Theta=\Theta^*}}{J(\Theta,t)|_{\Theta=0}}} \in [0,1]. \quad (44)$$

In (44), the optimal residual error $J(\Theta,t)|_{\Theta=\Theta^*}$ in (39) is generated by substituting the best-fitted parameters (35) into the cost function (29). Meanwhile, $M_{qq} = J(\Theta,t)\big|_{\Theta=0}$ corresponds to the cost function (29) as well, but with all the parameters assigned as zero. If (44) is close to zero, we can infer that the model is quite accurate. Instead, if (44) is close to one, which means the residual error has a magnitude comparable to M_{qq}, then we can conjecture the model itself is pretty inaccurate.

V. Analysis With Synthetic Simulation Data

This section utilizes synthetic simulation data to illustrate the effectiveness of the parameter error index and the system error index. Synthetic data is used because it is the only way to know the estimated parameters' true values [24].

Real measurements of a preceding car, i.e., its longitudinal speed and position, were directly injected into a car-following model of the ego-car (following car), and the simulated states from the ego-car model, i.e., $\hat{v}_x$ and $\hat{x}$, were named as the synthetic data. The initial conditions of the ego-car model were set equal to the actual measurements, as $\hat{v}_x(0) = v_x(0), \hat{x}(0) = x(0)$. In the following analysis, three pairs of leader/follower were randomly selected from the reconstructed NGSIM dataset [52], which corresponds to a 15-minute video record (4:00 PM - 4:15 PM) on the I-80 highway near San Francesco. Note that the raw NGSIM dataset contains various errors, such as sudden stop of a vehicle, discontinuity in trajectories, and unrealistically frequent rear-end collisions [37]. Those errors have been cleaned up in the reconstructed dataset. The reconstructed NGSIM datasets in [52] only cover the highway driving scenario. Urban driving data processing and analysis will be reserved for future studies.

To acquire trajectory information, e.g., position and speed, of each vehicle from the original NGSIM video, the authors in [52] conducted the following procedures: Firstly, they manually re-extracted vehicle trajectories from the video. This step is crucial as video processing software cannot adequately handle the projection errors, shadows, and occlusions. During this process, a Gaussian kernel filter was employed to smooth the position data of each vehicle. Then, the corresponding velocity profile was derived by differentiating the smoothed position data. A median filter was applied to yield the final velocity profile to mitigate sporadic positioning measurement errors. Similar speed smoothing approaches can also be found in [53].

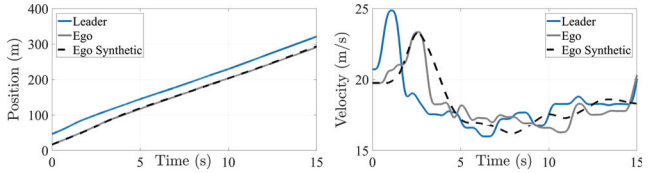

Fig. 4. Synthetic data from the CHM model.

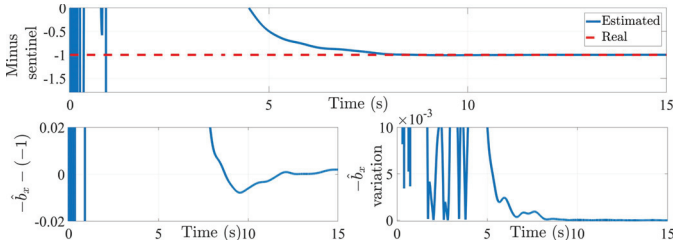

Fig. 5. Estimated sentinel parameter of the CHM model.

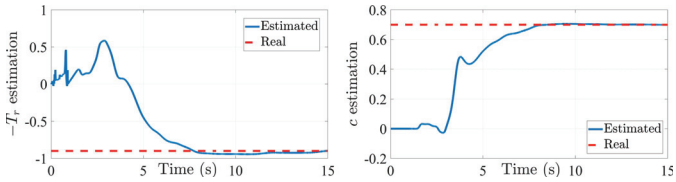

Fig. 6. Estimated reaction delay and sensitivity gain of the CHM model.

TABLE I

Parameter Estimation Results of the CHM Model

	Estimated value	True value	Absolute Error	Parameter Error Index
$-T_r$	-0.941	-0.900	0.041	4.926e-5
c	0.706	0.700	0.006	6.317e-6

We first show the synthetic data of the CHM model (1) in Fig. 4.

In Fig. 4, the blue and grey solid lines correspond to the recorded longitudinal velocity and position of the leading car and the following car (ego-car), respectively. The black dashed lines show the synthetic data of the ego-car.

Using the real data of the leading car and the synthetic data of the ego-car, we draw the estimated minus sentinel parameter $-\hat{b}_x$ of the CHM model in the top subplot of Fig. 5. Besides, the bottom-left subplot shows the discrepancy between $-\hat{b}_x$ and its truth value -1. The bottom-right subplot demonstrates the fluctuation metric of $-\hat{b}_x$. (See Eq. (38)).

In the beginning, because of the singularity of the matrix M_{PP} in (35), $-\hat{b}_x$ demonstrates a strong oscillation. As the numerical conditioning of M_{PP} gradually improved, $-\hat{b}_x$ veered towards and remained close to its true value as $-\hat{b}_x \to -1$. At around $t \approx 9.51s$, the criteria in (38) become valid, and we output $\hat{T}_r = -\Theta^*(2)$ and $\hat{c} = \Theta^*(4) . - \hat{T}_r$ and $\hat{c}$ from the CHM model are depicted in Fig. 6.

Similar to $-\hat{b}_x$, $\hat{T}_r$ and $\hat{c}$ gradually, but not asymptotically, converged towards their true values. At $t \approx 9.51s$, the parameter estimation results and the parameter error indices (43) are compiled in Table I.

Therefore, both the reaction delay and the sensitivity gain of the CHM model are correctly estimated. Moreover, the

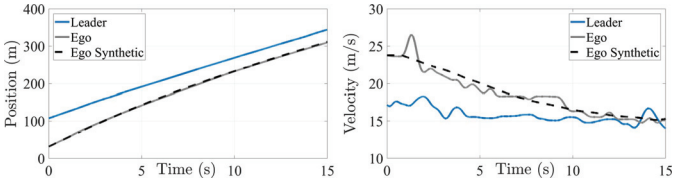

Fig. 7. Synthetic data from the GHR model.

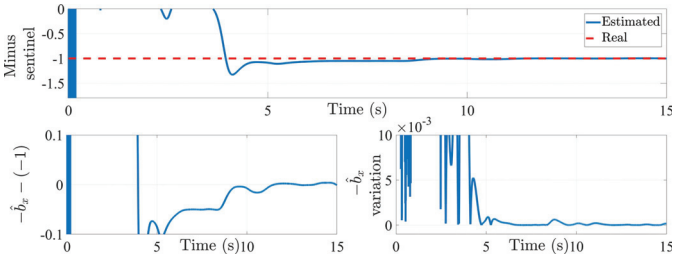

Fig. 8. Estimated sentinel parameter of the GHR model.

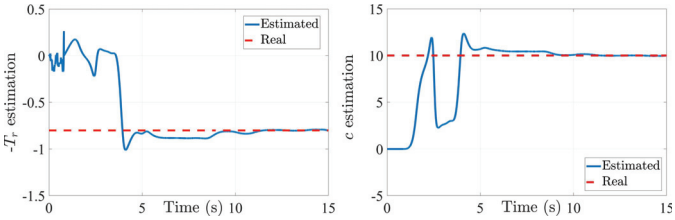

Fig. 9. Estimated reaction delay and sensitivity gain of the GHR model.

TABLE II

PARAMETER ESTIMATION RESULTS OF THE GHR MODEL

	Estimated value	True value	Absolute Error	Parameter Error Index
$-T_r$	-0.815	-0.800	0.015	4.336e-6
c	10.061	10.000	0.061	1.339e-5

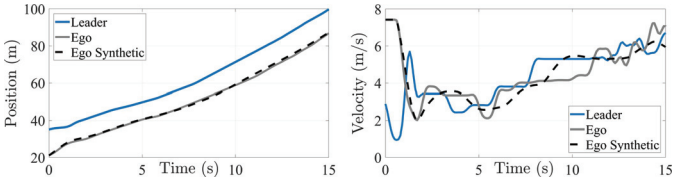

Fig. 10. Synthetic data from the Edie model.

parameter error indices reflect the fact that c is estimated with relatively higher accuracy than $-T_r$, as reflected in their absolute estimation errors.

Similar to the CHM model, the synthetic data, sentinel parameter, and parameter identification results of the GHR model are demonstrated in Fig. 7, Fig. 8, and Fig. 9, respectively.

The final estimation results are summarized in Table II.

Similar to Table I, the reaction delay and the sensitivity gain of the GHR model are correctly estimated as well. Moreover, the magnitudes of PEI also reflect the relative estimation accuracy of $-\hat{T}_r$ and $\hat{c}$.

Finally, the synthetic data, sentinel parameter, and identified parameters of the Edie model are demonstrated in Fig. 10, Fig. 11, and Fig. 12, respectively.

The estimation results are summarized in Table III.

Hence, parameters of the Edie model are correctly estimated as well, and the PEI values of $-\hat{T}_r$ and $\hat{c}$ are compatible with their absolute estimation errors.

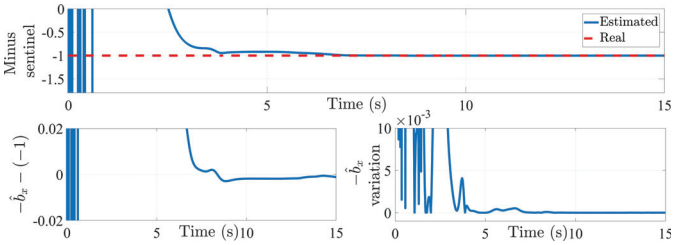

Fig. 11. Estimated sentinel parameter of the Edie model.

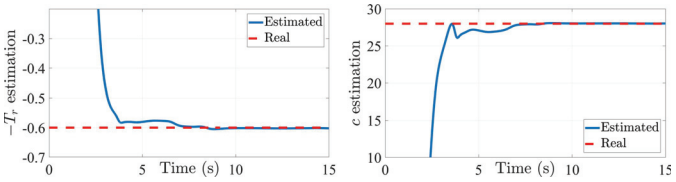

Fig. 12. Estimated reaction delay and sensitivity gain of the Edie model.

TABLE III

PARAMETER ESTIMATION RESULTS OF THE EDIE MODEL

	Estimated value	True value	Absolute Error	Parameter Error Index
$-T_r$	-0.597	-0.600	0.003	2.046e-5
c	27.925	28.000	0.075	3.599e-4

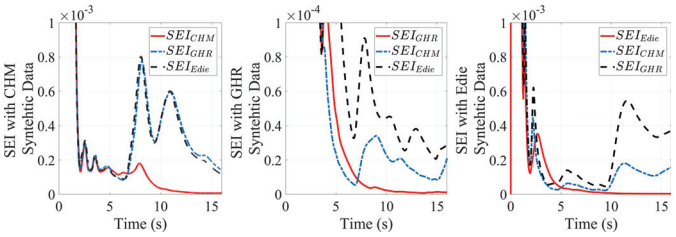

Fig. 13. System error indices comparison based on synthetic data.

The last plot based on the synthetic data is to show the effectiveness of the system error index (44). From (12), (22), (26), and (44), we can formulate three SEIs, as $SEI_{CHM}(t)$, $SEI_{GHR}(t)$, and $SEI_{Edie}(t)$. Synthetic data from the CHM model, the GHR model, and the Edie model in Fig. 4, Fig. 7, and Fig. 10 are individually injected into the three SEI generators, one by one. Fig. 13 depicts the SEI comparison results.

We can witness that one specific SEI generator will yield the lowest value if it receives the synthetic data from the corresponding car-following model. For instance, when the synthetic data from the CHM model is utilized, it is $SEI_{CHM}(t)$, which is constructed according to (1), that yields the globally lowest SEI value. Note that at the beginning of the simulation, all the three SEI values are pretty high because of the singular matrix M_{PP}. Therefore, it is recommended to bypass the initial fluctuation phase when SEI is utilized to evaluate and compare different car-following models with real measurement data.

VI. MODEL EVALUATION WITH SYSTEM ERROR INDEX

In contrast to Section V where synthetic data were applied, we utilize the reconstructed NGSIM datasets [52] in this section for evaluating and comparing the three car-following

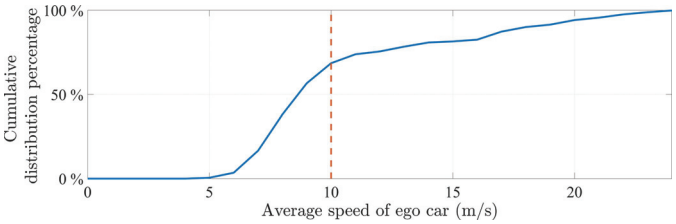

Fig. 14. Cumulative percentage of ego-car's average speed.

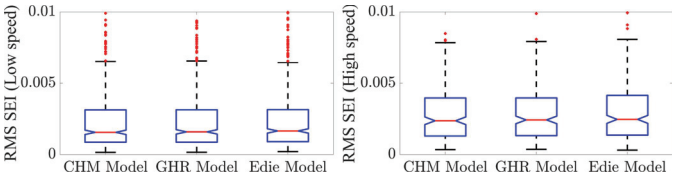

Fig. 15. RMS of system error indices of the three car-following models.

models (1), (2), and (3). Candidate leader-follower pairs were selected according to the following four criteria [26], [54]:

a) Neither the leader nor the follower changed its lane.
b) The maximum inter-vehicle gap was less than 70 meters.
c) The time headway of the follower (ego-car) to the precedent car was less than 2.5s.
d) The overall recorded period was at least 15s.

Requirement a) falls to the basic definition of car-following. Requirements b) and c) guarantee the interaction between the two cars such that the ego-car itself was not in a free-driving status. Requirement d) eliminates short-term, occasionally formed leader-follower pairs. Finally, 772 leader-follower pairs were identified. None of them yields $v_x(t) = 0$ or $v_x(t) = 0$ during a long period. The maximum average speed of the follower reached 24m/s, while its minimum average speed was less than 4m/s. The cumulative percentage of the ego-car's average speed is depicted in Fig. 14.

Based on Fig. 14, we roughly divide the 772 leader-follower pairs into two categories. The high-speed group contains all the trajectories with the follower's average speed larger than 10m/s, and the rest is classified as the low-speed group. There are in total 529 and 243 low-speed and high-speed pairs, respectively.

The root mean squares of the three SEIs for each trajectory: $RMS\left(SEI_j^i\left(t|\,t \geq t^*\right)\right), j \in \{CHM, GHR, Edie\}, i \in \{1, 2\ldots, 722\}$, are calculated. In Fig. 15, we separate the results of the low-speed cases from the high-speed cases. Note that we removed the initial fluctuation phases of the three SEIs (see Fig. 13) by calculating the RMS after $t \geq t^*$, where t^* is the moment after which the three SEIs do not demonstrate drastic variations. To determine t^*, we exploit (38) again, as:

$$t^* = \arg\max_t \left(\begin{array}{l} \left|\sigma\left(SEI_j(t)\right)/E\left(SEI_j(t)\right)\right| \leq \Delta_{SEI}, \\ j \in \{CHM, GHR, Edie\} \end{array} \right). \tag{45}$$

In other words, after t^*, the fluctuation metrics of the three SEIs are all below the threshold Δ_{SEI}. We fixed Δ_{SEI} as 1e-3.

As indicated in the box plots in Fig. 15, the three car-following models yield quite similar distributions of the root mean squares of the system error indices. To statistically

TABLE IV
LOW-SPEED CASE ANOVA TEST TABLE

Source	SS	df	MS	F	Prob>F
Between	0.00001	2	5.68e-6	0.7	**0.494**
Within	0.01277	1584	8.06e-6		
Total	0.01278	1586			

TABLE V
HIGH-SPEED CASE ANOVA TEST TABLE

Source	SS	df	MS	F	Prob>F
Between	0.00001	2	3.34e-6	0.45	**0.639**
Within	0.00541	726	7.45e-6		
Total	0.00542	728	7.45e-6		

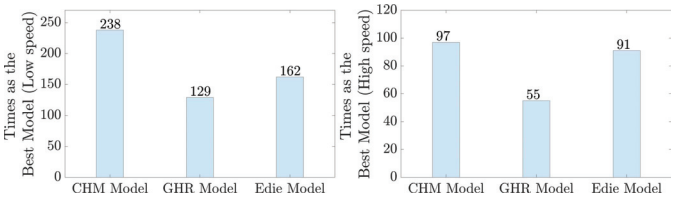

Fig. 16. Times regarded as the best model.

compare the means of $RMS\left(SEI\left(t|\,t \geq t^*\right)\right)$ of the three models and determine whether any one of those means is statistically significantly different from the others, we execute the one-way analysis of variance (ANOVA). The one-way ANOVA tests the null hypothesis:

$$\begin{aligned} H_0 : \mu &\left(RMS\left(SEI_{CHM}^i\left(t|\,t \geq t^*\right)\right)\right) \\ &= \mu\left(RMS\left(SEI_{GHR}^i\left(t|\,t \geq t^*\right)\right)\right) \\ &= \mu\left(RMS\left(SEI_{Edie}^i\left(t|\,t \geq t^*\right)\right)\right), \end{aligned} \tag{46}$$

by comparing the variation of $RMS\left(SEI\left(t|\,t \geq t^*\right)\right)$ between groups to the variation within groups. If the ratio of between-group variation to within-group variation is significantly high, then we can conclude that the group means are significantly different from each other. The test results of both the low-speed ($i = 1, \ldots, 529$) and the high-speed ($i = 1, \ldots, 243$) trajectories are summarized in Table IV and Table V, respectively.

Table IV yields $F(2, 1584) = 0.7$, $p = 0.494 \gg 0.05$ and Table V produces $F(2, 726) = 0.45$, $p = 0.639 \gg 0.05$. As the p-values of both tables are larger than the significant level of 0.05, we cannot reject the null hypothesis (46). In other words, there does not exist a statistically significant difference among the three models. This finding echoes well with the conclusion in [26] that the three car-following models (1), (2), (3) have practically the same performance, and no one can be regarded better than the others.

In addition, we show the counts of each car-following model evaluated as the *best model*. The best model corresponds to the one producing the lowest $RMS\left(SEI\left(t|\,t \geq t^*\right)\right)$. The results are shown in Fig. 16.

Therefore, the CHM model is recognized as the best model with the most counts, followed by the Edie model and finally, the GHR model. However, this result should not be interpreted as the GHR model is the worst among the three, because the performance of a model is reflected in both the counts that it is considered the best and also the counts that it is *not*

considered the best. Interestingly, the relative percentage that the Edie model was evaluated as the best one augmented as the ego-car's average speed increased. This result is expected because the Edie model was particularly designed for describing relatively high-speed, non-congested traffic flows [9].

VII. Conclusion

This paper proposes an algebraic framework for evaluating and comparing a class of car-following models whose parameters are linearly identifiable. In opposite to the model evaluation strategies based on numerical optimization, no cost function needs to be designed *a priori*. Moreover, no numerical solver is involved. These two beneficial features guarantee that the evaluation result is robust to local optimality and independent of the formulation of the goodness of fit. The key idea of the proposed strategy is to carry algebraic manipulations on the car-following model and derive the so-called system error index (SEI), which serves as a uniform metric for evaluating and comparing different car-following models. This method sheds some light on quickly prescreening a newly proposed car-following model and investigating potential modeling performance enhancement via model structure modification. A performant car-following model facilitates assessing the impact of emerging technologies, e.g., V2V communication, on transportation systems. Applying the proposed algebraic evaluation framework on the CHM model, the GHR model, and the Edie model reveals that none of them is statistically significantly better than the other two models in terms of the averaged RMS of the SEIs. Because of the nonlinear relationships among parameters, behavioral car-following models, e.g., IDM model or Gipps' model, cellular automation based model, and psychophysical model, cannot be analyzed under the proposed algebraic framework. Instead, authors in [55] recently proposed another uniform framework based on the Pareto-efficiency to evaluate linear/linearizable and nonlinear car-following models. Evaluating the models (1), (2), (3) with the methods in [55] will be studied in the next step.

References

[1] A. Kesting and M. Treiber, "Calibrating car-following models by using trajectory data: Methodological study," *Transp. Res. Record, J. Transp. Res. Board*, vol. 2088, no. 1, pp. 148–156, Jan. 2008, doi: 10.3141/2088-16.

[2] B. Ciuffo, M. Makridis, T. Toledo, and G. Fontaras, "Capability of current car-following models to reproduce vehicle free-flow acceleration dynamics," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 11, pp. 3594–3603, Nov. 2018, doi: 10.1109/TITS.2018.2866271.

[3] D. Yang, L. Zhu, Y. Liu, D. Wu, and B. Ran, "A novel car-following control model combining machine learning and kinematics models for automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 1991–2000, Jun. 2019, doi: 10.1109/TITS.2018.2854827.

[4] M. Lindorfer, C. F. Mecklenbräuker, and G. Ostermayer, "Modeling the imperfect driver: Incorporating human factors in a microscopic traffic model," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2856–2870, Sep. 2018, doi: 10.1109/TITS.2017.2765694.

[5] T. Kim and K. Jerath, "Congestion-aware cooperative adaptive cruise control for mitigation of self-organized traffic jams," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 24, 2021, doi: 10.1109/TITS.2021.3059237.

[6] Y. Ma and J. Wang, "Energetic impacts evaluation of eco-driving on mixed traffic with driver behavioral diversity," *IEEE Trans. Intell. Transp. Syst.*, early access, Nov. 24, 2020, doi: 10.1109/TITS.2020.3036326.

[7] R. E. Chandler, R. Herman, and E. W. Montroll, "Traffic dynamics: Studies in car following," *Oper. Res.*, vol. 6, no. 2, pp. 165–184, 1958, doi: 10.1287/opre.6.2.165.

[8] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Oper. Res.*, vol. 7, no. 4, pp. 499–505, 1959, doi: 10.1287/opre.7.4.499.

[9] L. C. Edie, "Car-following and steady-state theory for noncongested traffic," *Oper. Res.*, vol. 9, no. 1, pp. 66–76, Feb. 1961, doi: 10.1287/opre.9.1.66.

[10] M. Brackstone and M. McDonald, "Car-following: A historical review," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 2, no. 4, pp. 181–196, Dec. 1, 1999, doi: 10.1016/S1369-8478(00)00005-X.

[11] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, pp. 1805–1824, Aug. 2000, doi: 10.1103/PhysRevE.62.1805.

[12] X. Li, Q. Wu, and R. Jiang, "Cellular automaton model considering the velocity effect of a car on the successive car," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, Nov. 2001, Art. no. 066128, doi: 10.1103/PhysRevE.64.066128.

[13] R. Jiang, Q. Wu, and Z. Zhu, "Full velocity difference model for a car-following theory," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 1, Jun. 2001, Art. no. 017101, doi: 10.1103/PhysRevE.64.017101.

[14] G. F. Newell, "A simplified car-following theory: A lower order model," *Transp. Res. B, Methodol.*, vol. 36, no. 3, pp. 195–205, 2002, doi: 10.1103/PhysRevE.62.1805.

[15] J. Wu, M. Brackstone, and M. McDonald, "The validation of a microscopic simulation model: A methodological case study," *Transp. Res. C, Emerg. Technol.*, vol. 11, no. 6, pp. 463–479, Dec. 2003, doi: 10.1016/j.trc.2003.05.001.

[16] A. Tordeux, S. Lassarre, and M. Roussignol, "An adaptive time gap car-following model," *Transp. Res. B, Methodol.*, vol. 44, nos. 8–9, pp. 1115–1131, 2010, doi: 10.1016/j.trb.2009.12.018.

[17] H. N. Koutsopoulos and H. Farah, "Latent class model for car following behavior," *Transp. Res. B, Methodol.*, vol. 46, no. 5, pp. 563–578, 2012, doi: 10.1016/j.trb.2012.01.001.

[18] B. Higgs, M. Abbas, and A. Medina, "Analysis of the Wiedemann car following model over different speeds using naturalistic data," in *Proc. Road Saf. Simulation*, Indianapolis, IN, USA, Sep. 2011, pp. 1–22.

[19] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transp. Res. B, Methodol.*, vol. 15, no. 2, pp. 105–111, 1981, doi: 10.1016/0191-2615(81)90037-0.

[20] B. Ciuffo, V. Punzo, and M. Montanino, "Thirty years of Gipps' car-following model," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2315, no. 1, pp. 89–99, Jan. 2012, doi: 10.3141/2315-10.

[21] Z. Li, F. Khasawneh, X. Yin, A. Li, and Z. Song, "A new microscopic traffic model using a spring-mass-damper-clutch system," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 8, pp. 3322–3331, Aug. 2020, doi: 10.1109/TITS.2019.2926146.

[22] E. Brockfeld, R. D. Kühne, and P. Wagner, "Calibration and validation of microscopic traffic flow models," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1876, no. 1, pp. 62–70, Jan. 2004, doi: 10.3141/1876-07.

[23] J. Monteil, R. Billot, J. Sau, C. Buisson, and N.-E. El Faouzi, "Calibration, estimation, and sampling issues of car-following parameters," *Transp. Res. Rec., J. Tranp. Res. Board*, vol. 2422, pp. 131–140, Oct. 2014, doi: 10.3141/2422-15.

[24] V. Punzo, B. Ciuffo, and M. Montanino, "Can results of car-following model calibration based on trajectory data be trusted?" *Transp. Res. Rec., J. Tranp. Res. Board*, vol. 2315, pp. 11–24, Dec. 2012, doi: 10.3141/2315-02.

[25] V. Punzo, M. Montanino, and B. Ciuffo, "Do we really need to calibrate all the parameters? variance-based sensitivity analysis to simplify microscopic traffic flow models," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 184–193, Feb. 2015, doi: 10.1109/TITS.2014.2331453.

[26] S. Ossen and S. Hoogendoorn, "Car-following behavior analysis from microscopic trajectory data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1934, pp. 13–21, Jul. 2005, doi: 10.1177/0361198105193400102.

[27] C. P. I. J. van Hinsbergen, W. J. Schakel, V. L. Knoop, J. W. C. van Lint, and S. P. Hoogendoorn, "A general framework for calibrating and comparing car-following models," *Transportmetrica A, Transp. Sci.*, vol. 11, no. 5, pp. 420–440, 2015, doi: 10.1080/23249935.2015.1006157.

[28] M. Treiber and A. Kesting, "Microscopic calibration and validation of car-following models—A systematic approach," *Proc.-Social Behav. Sci.*, vol. 80, pp. 922–939, Jun. 2013, doi: 10.1016/j.sbspro.2013.05.050.

[29] V. Punzo and M. Montanino, "Speed or spacing? Cumulative variables, and convolution of model errors and time in traffic flow models validation and calibration," *Transp. Res. B, Methodol.*, vol. 91, pp. 21–33, Sep. 2016, doi: 10.1016/j.trb.2016.04.012.
[30] S. Ossen, S. P. Hoogendoorn, and B. G. H. Gorte, "Interdriver differences in car-following: A vehicle trajectory–based study," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1965, no. 1, pp. 121–129, Jan. 2006, doi: 10.1177/0361198106196500113.
[31] Y. Hollander and R. Liu, "The principles of calibrating traffic microsimulation models," *Transportation*, vol. 35, no. 3, pp. 347–362, May 2008, doi: 10.1007/s11116-007-9156-2.
[32] J. Zheng, K. Suzuki, and M. Fujita, "Evaluation of car-following models using trajectory data from real traffic," *Proc.-Social Behav. Sci.*, vol. 43, pp. 356–366, Jan. 2012, doi: 10.1016/j.sbspro.2012.04.109.
[33] N. Raju, S. Arkatkar, and G. Joshi, "Evaluating performance of selected vehicle following models using trajectory data under mixed traffic conditions," *J. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 617–634, Nov. 2020, doi: 10.1080/15472450.2019.1675522.
[34] R. Balakrishna, C. Antoniou, M. Ben-Akiva, H. N. Koutsopoulos, and Y. Wen, "Calibration of microscopic traffic simulation models," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1999, no. 1, pp. 198–207, Jan. 2007, doi: 10.3141/1999-21.
[35] V. Punzo and F. Simonelli, "Analysis and comparison of microscopic traffic flow models with real traffic microscopic data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1934, pp. 53–63, Jan. 2005, doi: 10.1177/0361198105193400106.
[36] P. Ranjitkar, T. Nakatsuji, and A. Kawamua, "Car-following models: An experiment based benchmarking," *J. Eastern Asia Soc. Transp. Stud.*, vol. 6, pp. 1582–1596, Jan. 2005, doi: 10.11175/easts.6.1582.
[37] N. Chiabaut, L. Leclercq, and C. Buisson, "From heterogeneous drivers to macroscopic patterns in congestion," *Transp. Res. B, Methodol.*, vol. 44, no. 2, pp. 299–308, Feb. 2010, doi: 10.1016/j.trb.2009.07.009.
[38] J. Taylor, X. Zhou, N. M. Rouphail, and R. J. Porter, "Method for investigating intradriver heterogeneity using vehicle trajectory data: A dynamic time warping approach," *Transp. Res. B, Methodol.*, vol. 73, pp. 59–80, Mar. 2015, doi: 10.1016/j.trb.2014.12.009.
[39] V. Papathanasopoulou, I. Markou, and C. Antoniou, "Online calibration for microscopic traffic simulation and dynamic multi-step prediction of traffic speed," *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 144–159, Jul. 2016, doi: 10.1016/j.trc.2016.04.006.
[40] P. Ranjitkar, T. Nakatsuji, and A. Kawamura, "Experimental analysis of car-following dynamics and traffic stability," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1934, no. 1, pp. 22–32, Jan. 2005, doi: 10.1177/0361198105193400103.
[41] C. Garcia-Rodriguez, J. A. Cortes-Romero, and H. Sira-Ramirez, "Algebraic identification and discontinuous control for trajectory tracking in a perturbed 1-DOF suspension system," *IEEE Trans. Ind. Electron.*, vol. 56, no. 9, pp. 3665–3674, Sep. 2009, doi: 10.1109/TIE.2009.2026383.
[42] M. Mboup, C. Join, and M. Fliess, "Numerical differentiation with annihilators in noisy environment," *Numer. Algorithms*, vol. 50, no. 4, pp. 439–467, Apr. 2009, doi: 10.1007/s11075-008-9236-1.
[43] Z. Wang and J. Wang, "Ultra-local model predictive control: A model-free approach and its application on automated vehicle trajectory tracking," *Control Eng. Pract.*, vol. 101, Aug. 2020, Art. no. 104482, doi: 10.1016/j.conengprac.2020.104482.
[44] D. G. Luenberger, "Observers for multivariable systems," *IEEE Trans. Autom. Control*, vol. 11, no. 2, pp. 190–197, Apr. 1966, doi: 10.1109/TAC.1966.1098323.
[45] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. Amer. Soc. Mech. Eng.*, vol. 82, pp. 35–45, Mar. 1960, doi: 10.1115/1.3662552.
[46] S. Lyashevskiy and Y. Chen, "The Lyapunov stability theory in system identification," in *Proc. Amer. Control Conf.*, Albuquerque, NM, USA, vol. 1, Jun. 1997, pp. 617–621, doi: 10.1109/ACC.1997.611873.
[47] M. Fliess and H. Sira-Ramírez, "An algebraic framework for linear identification," *ESAIM Control. Optim. Calc. Var.*, vol. 9, pp. 151–168, Jan. 2003, doi: 10.1051/cocv:2003008.
[48] J. Becedas, G. Mamani, and V. Feliu, "Algebraic parameters identification of DC motors: Methodology and analysis," *Int. J. Syst. Sci.*, vol. 41, no. 10, pp. 1241–1255, Oct. 2010, doi: 10.1080/00207720903244097.
[49] K. Wang, J. Chiasson, M. Bodson, and L. M. Tolbert, "A nonlinear least-squares approach for identification of the induction motor parameters," *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1622–1628, Oct. 2005, doi: 10.1109/TAC.2005.856661.
[50] J. Chiasson, *Modeling and High-Performance Control of Electric Machines*. New York, NY, USA: Wiley, 2005.
[51] A. M. A. Oteafy, J. N. Chiasson, and S. Ahmed-Zaid, "Development and application of a standstill parameter identification technique for the synchronous generator," *Int. J. Electr. Power Energy Syst.*, vol. 81, pp. 222–231, Oct. 2016, doi: 10.1016/j.ijepes.2016.02.030.
[52] B. Coifman and L. Li, "A critical evaluation of the next generation simulation (NGSIM) vehicle trajectory dataset," *Transp. Res. B, Methodol.*, vol. 105, pp. 362–377, Nov. 2017, doi: 10.1016/j.trb.2017.09.018.
[53] M. Montanino and V. Punzo, "Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns," *Transp. Res. B, Methodol.*, vol. 80, pp. 82–106, Oct. 2015, doi: 10.1016/J.TRB.2015.06.010.
[54] T. Toledo, H. N. Koutsopoulos, and M. Ben-Akiva, "Integrated driving behavior modeling," *Transp. Res. C, Emerg. Technol.*, vol. 15, no. 2, pp. 96–112, Apr. 2007, doi: 10.1016/j.trc.2007.02.002.
[55] V. Punzo, Z. Zheng, and M. Montanino, "About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes," *Transp. Res. C, Emerg. Technol.*, vol. 128, Jul. 2021, Art. no. 103165, doi: 10.1016/J.TRC.2021.103165.

Zejiang Wang (Graduate Student Member, IEEE) received the B.E. degree (Hons.) in mechanical engineering and automation from Southeast University, Nanjing, China, in 2014, the Dipl.Ing. degree from ENSTA ParisTech, Palaiseau, France, and the double M.S. degree in design, modeling, and architecture of complex industrial systems from the École Polytechnique, University of Paris-Saclay, Palaiseau, in 2017. He is currently pursuing the Ph.D. degree with the Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, USA.

His research interests include vehicle dynamics and control, intelligent transportation systems, and cyber-physical systems.

Xingyu Zhou received the B.S. degree (Hons.) in mechanical engineering from Purdue University, West Lafayette, IN, USA, in 2016, and the M.S. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2018. He is currently pursuing the Ph.D. degree with the Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, USA. His research interests include vehicle dynamics and control, robust control, and ADAS systems.

Junmin Wang (Senior Member, IEEE) received the B.E. degree in automotive engineering and the first M.S. degree in power machinery and engineering from Tsinghua University, Beijing, China, in 1997 and 2000, respectively, the second and third M.S. degrees in electrical engineering and mechanical engineering from the University of Minnesota, Twin Cities, Minneapolis, MN, USA, in 2003, and the Ph.D. degree in mechanical engineering from The University of Texas at Austin, Austin, TX, USA, in 2007. He is the Lee Norris & Linda Steen Norris Endowed Professor with the Walker Department of Mechanical Engineering, The University of Texas at Austin. He is the author or coauthor of more than 350 peer-reviewed publications, including 178 journal articles, and holds 13 U.S. patents. He is an IEEE Vehicular Technology Society Distinguished Lecturer and a fellow of SAE and ASME.