Identifiability of car-following dynamics

Yanbing Wang [*]   Maria Laura Delle Monache [†]

Daniel B. Work[*]

Abstract

The advancement of in-vehicle sensors provides abundant datasets to estimate parameters of car-following models that describe driver behaviors. The question of parameter identifiability of such models (i.e., whether it is possible to infer its unknown parameters from the experimental data) is a central system analysis question, and yet still remains open. This article presents both structural and practical parameter identifiability analysis on four common car-following models: i) the constant-time headway relative-velocity (CTH-RV) model, ii) the optimal velocity model (OV), iii) the follow-the-leader model (FTL) and iv) the intelligent driver model (IDM). The structural identifiability analysis is carried out using a differential geometry approach, which confirms that, in theory, all of the tested car-following systems are structurally locally identifiable, i.e., the parameters can be uniquely inferred under almost all initial condition and admissible inputs by observing the space gap alone. In a practical setting, we propose an optimization-based numerical direct test to determine parameter identifiability given a specific experimental setup (the specific initial conditions and input are known). The direct test conclusively finds distinct parameters under which the CTH-RV and FTL are not identifiable under the given initial condition and input trajectory.

1 Introduction

Motivation. With the advancement of sensor technologies, abundant traffic data is readily available to study traffic patterns and individual driving behaviors. Examples of such datasets include overhead camera data [31, 51] and floating-car data [2, 10, 57, 63]. The collection of such datasets support tasks such as the parameterization of microscopic models which describe individual vehicle car-following behaviors.

---

[*]Department of Civil and Environmental Engineering and the Institute for Software Integrated Systems, Vanderbilt University, Nashville, TN 37240

[†]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France and Department of Civil and Environmental Engineering, University of California, Berkeley. Present address: Department of Civil and Environmental Engineering, University of California, Berkeley, 760 Davis, MC 1710, Berkeley, CA, 94720-1710, USA

Studies of car-following models and calibration of such models have mostly been focused on data-fitting quality. Model calibration is usually posed as an optimization problem such that the best fit parameters are found by minimizing the error between the model prediction and corresponding measurement data [29, 57], or through probabilistic approaches to find the most likely parameter candidate [1, 78]. Although the approaches report good accuracy of the estimated parameters, they lack a theoretical guarantee that a unique parameter set can be recovered, which is provided by an identifiability analysis of the evolution-observation system [7,39]. In other words, using data-fitting quality as the metric alone for model calibration does not guarantee the uniqueness of the best-fit parameters, nor can it tell how robust the current estimation is under a different experimental dataset or numerical setup [39]. Such ambiguity might result in a different parametrization when repeating the experiment, leading to an inaccurate characterization of car-following behaviors, such as analytically determining string stability [50].

Besides that different driver behaviors on a microscopic scale can greatly change traffic features on a macroscopic scale [34, 35, 57], knowledge of microscopic model parameters is key to understand individual driving behavior of the adaptive cruise control and the interplay between automated vehicles and human drivers [5, 16, 20, 21, 29, 47, 48, 78]. Therefore, the identifiability of car-following models is an important question to advance calibration of these models.

Problem statement. The parameter identifiability problem refers to the ability to uniquely recover the unknown model parameters from the observation data. Mathematically, consider a general form of a dynamical system:

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), \theta) \\ y(t) = g(x(t), u(t), \theta) \\ x(0) = x_0, \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state variable vector, $u(t)$ is the scalar input, and $\theta \in \mathbb{R}^{n_\theta}$ is the parameter vector; f, g are analytic vector functions describing the evolution of the state and the measurement model, and x_0 is the initial condition. The generic question for parameter identifiability is: *Given the system* (1) *and a known input* $u(t)$, *can we uniquely determine the model parameters* θ *from the output* $y(t)$ *for a given initial condition?* In the context of car-following models, the system of equations can be written as:

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} \dot{s}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} u(t) - v(t) \\ f_{\mathrm{CF}}(x(t), u(t), \theta) \end{bmatrix} \\ y(t) = x(t) \\ x(0) = \begin{bmatrix} s_0 \\ v_0 \end{bmatrix}. \end{cases} \tag{2}$$

The state vector $x(t)$ in (2) is given by the space gap $s(t)$ and the velocity $v(t)$ of the ego vehicle. The space gap is taken as the distance between the lead vehicle

2

and the ego (follower) vehicle while the speed is that of the ego vehicle. The input $u(t)$ is the velocity profile of a lead vehicle and f_{CF} is usually a 2nd order ordinary differential equation (ODE) that describes the car-following dynamics.

With in-vehicle sensor data such as radar data being available, it is now possible to directly measure $s(t)$ and $v(t)$, as well as the velocity profile of the lead vehicle $u(t)$. For (2), the parameter identifiability question becomes: *Given the output time-series $y(t)$ (the velocity and space gap measurements) and the input $u(t)$, can we uniquely determine the model parameters θ in f_{CF}?*

One type of identifiability is *structural identifiability* (or theoretical identifiability), which investigates the input-output configuration and relies on algebraic calculations to determine if the parameter set is unique from the observation [40, 76]. It is completely defined by the model structure and does not consider data quality or model errors. A *structural identifiability* analysis can be performed prior to the estimation of the unknowns in order to detect potential model structural issues and separate them from numerical problems, which should be dealt with differently [72]. However, structurally identifiable systems do not preclude the existence of some specific initial conditions under which the system becomes unidentifiable [60, 68]. For this limitation, *practical identifiability* can be complimentary to provide insights on the dynamical models.

Practical identifiability is usually based on numerical methods to determine, for a given experimental setup, whether the parameters are unique. This is sometimes modified to also consider the presence of moderate measurement errors. In practice, a realistic way to formulate the identifiability problem when performing an experiment becomes: *for a given initial condition and an input trajectory that are known, whether there exists distinct valid parameters that produce the same outputs?* Notice that ensuring uniqueness of the parameters (e.g. there are no distinct parameters that produce the same output) does not imply well-posedness of the inverse problem. When solving the inverse problem, one needs to ensure additional properties, e.g. continuity w.r.t. the initial datum, to guarantee the recovery of the parameters and reconstruction the input.

This paper investigates, from both theoretical and practical angles, parameter identifiability of a microscopic car-following system instantiated as one of four different car-following models, i.e., the constant-time headway relative-velocity (CTH-RV) model [5, 20, 47], the optimal velocity (OV) model [5], the follow-the-leader (FTL) model [15] and the intelligent driver model (IDM) [62]. Only CTH-RV is a linear model while other three are nonlinear. We address the following questions:

1. Are the car-following systems of the form (2) structurally identifiable?

2. Are they identifiable in practice given a specific experimental setup (i.e., given a known initial condition and a known input)?

3. Are they identifiable in practice given a specific experimental setup when moderate measurement errors are allowed?

3

We address the question of structural identifiability using a *differential geometry approach*, and practical identifiability using a *numerical direct test*. The main findings to the above three questions are:

1. The investigated car-following models are all structurally locally identifiable, i.e., almost all points in the parameter space of each of the car following models can be uniquely identified for all admissible inputs given almost all initial conditions. Special initial conditions and parameter sets that lose structural identifiability are also discovered.

2. Given a specific experimental setup (the specific initial condition and input that are known), the CTH-RV and FTL are not identifiable, i.e., there exist distinct parameters that produce the same output.

3. Given a specific experimental setup when measurement errors are present, all models are not practically identifiable, i.e., there exist distinct parameters that produce the same output within a small error bound.

Contributions. Throughout the literature of microscopic traffic modeling we found that the focus has been on model calibration techniques instead of parameter identifiability. In this light, the main contributions of this article are the following. (*i*) We provide structural identifiability analysis of four common car-following models using a differential geometry framework. This allows us to detect if there are structural issues of the dynamic model and to distinguish them from other possible causes of calibration failures, such as the choice of optimization algorithms. (*ii*) We subsequently provide a numerical *direct test*, which is a constrained optimization problem, to check identifiability for a specific experimental design (for a given and known initial condition and input; (*iii*) finally we investigate the practical identifiability for a specific experimental design when moderate measurement errors are present, also with a *direct test*. This study therefore provides both analytical and practical insights on parameter identification.

The organization of this paper is as follows. In Section 2 we summarize the existing microscopic traffic model calibration work, as well as methods derived in other research fields to tackle parameter identifiability problems. In Section 3 we present the four microscopic car-following models that we analyze in this work. We introduce the methodological tools used, including the *differential geometry* approach and the *numerical direct test* approach, in Section 4. Section 5 provides the structural identifiability results of the selected car-following models, and Section 6 presents the practical identifiability results.

2   Related work

The inverse problem of learning microscopic car following model parameters through experimental data (e.g., microscopic-level trajectory data and macroscopic level aggregated counts from fixed sensors) is part of the process to calibrate complex microscopic traffic simulation software. A recent review of these

calibration techniques appears in [37], and is commonly categorized into four types [35, 52, 55, 56]. Briefly put, *type I* calibration views the calibration problem as a likelihood estimation problem where the distribution of the parameter likelihood is calculated for the future time step based on historical driving data (e.g., [25, 52, 56, 57, 63, 64, 78]). *Type II* calibration directly uses a global search to find the best-fit parameters for which the simulated complete trajectory most closely represent the observed trajectory (e.g., [13, 20, 41, 53, 77]). *Type III* calibration considers the long-range interactions amongst vehicles within a platoon (e.g., [21, 33, 34]), and *type IV* calibration relies on mesoscopic or macroscopic traffic flow patterns such as the headway distributions (e.g., [26, 35]). This paper considers parameter identifiability of microscopic models given microscopic data (e.g., velocity and spacing), which are closely related to *Type I* and *II*. More background on these *Type I* and *II* calibration methods can be found in book chapters such as [12,63], review articles [23,25,28,36,37] and the references therein.

In addition to the question of determining the best fit parameters given experimental data, several studies have also considered the confidence level and the sensitivity of parameter estimation. For example, Punzo et al. [56] considered a variance-based global sensitivity analysis to produce the importance ranking of the IDM parameters, and to consequently reduce the parameters to be further calibrated. Monteil & Bouroche [49] considered a systematic statistical approach to first use a global sensitivity analysis to reduce the parameter space, then log likelihood estimation for the insensitive parameters and finally likelihood-ratio for interval estimation; Treiber & Kesting [64] investigated on the data sampling interval, completeness and parameter orthogonality and their effects on parameter calibration of the IDM, instead of solely on data-fitting quality. The sensitivity approaches considered in the mentioned works are closely related to identifiability [46, 49]. However, a formal analysis of identifiability on microscopic traffic models (from a theoretical perspective) has not yet been explicitly studied in the previous work.

Although a systematic study on the parameter identifiability for microscopic traffic models still remains unaddressed in transportation research, it has been developed in and extensively applied to other research fields such as waste water treatment process [9], robot dynamics [30], and biological processes [65], where the system evolution equations are also expressed as ODEs. We summarize the related approaches for tackling structural and practical identifiability respectively.

Structural identifiability of dynamical systems is closely related to the algebraic formulation of the dynamic equations, and provides theoretical possibility for uniquely inferring the system unknowns *a priori* before collecting experimental data [18,40]. It can be analyzed using similarity transformation [11,22,66,74] (applicable to autonomous systems), Laplace transform [7], power series expansion [19,54], implicit functions [79], differential algebra [14,40,59] and differential geometry [70] (applicable to systems with external inputs). Multiple software packages have also been developed for direct implementation of some of the techniques, for example, DAISY [8], COMBOS [44], SIAN [24] and STRIKE-

GOLDD [71, 72]. All the mentioned tools are suitable for analyzing systems of which dynamics are written as rational functions. STRIKE-GOLDD, in particular, is capable for analyzing nonlinear, non-rational system dynamics as well, making it suitable for analyzing a wider class of car-following models.

In practice when an experiment is fully defined (the specific values of the initial condition and the input trajectory are known), numerical methods are often useful to assess identifiability. The numerical methods can be used to explore the profile likelihood [32, 58], which allows to derive likelihood-based confidence intervals for each parameter and recovers the functional relations between parameters due to non-identifiability. The sensitivity matrix [61] combines numerical calculations with a tractable symbolic computation to investigate local structural identifiability.

Numerical methods can also help to assess practical identifiability, which relaxes the noise-free assumption of f and g in the structural analysis. Methods such as Monte Carlo simulations [45] help to check the relative error of estimated parameters under noisy output measurements. Numerically evaluating the profile likelihood can help to distinguish practically unidentifiable parameters due to measurement error from structurally unidentifiable ones by the shape of the likelihood profile [42]..

3 Car-following models

In this section, we briefly present four commonly used car-following models that describe the acceleration dynamics in the form of an ODE.

Constant-time headway relative-velocity (CTH-RV) model The CTH-RV model has been predominately used to describe the car-following behaviors of ACC vehicles [6, 20, 38, 47, 48, 78]. It is a simple model with linear dynamics with respect to the space gap and relative velocity $\Delta v(t) := u(t) - v(t)$:

$$\dot{v}(t) = k_1(s(t) - \tau v(t)) + k_2(\Delta v(t)), \tag{3}$$

where k_1 and k_2 are non-negative gains and τ is the constant time-headway. The three (time-invariant) parameters constitute the parameter vector of this model, i.e., $\theta = [k_1, k_2, \tau]$, whose identifiability is to be determined.

The equilibrium initial condition x_0^* for the car-following system (2) under the acceleration dynamic (3) is:

$$x_0^* = [\tau u_0, u_0]^T, \tag{4}$$

where $u(0) = u_0$ is the initial velocity of the leader. It is easy to see that this initial condition results in $\dot{s}(0) = u_0 - v_0 = 0$ and $\dot{v}(0) = f_{CF}(x_0^*, u_0, \theta) = 0$.

Optimal velocity (OV) model We also consider the car-following model proposed by Bando et al. [4, 5], where an optimal velocity term $V(s(t))$ is intro-

duced to describe the desired spacing-speed relationship at equilibrium:

$$V(s(t)) = a \left(\tanh \left(\frac{s(t) - h_m}{b} \right) + \tanh \left(\frac{h_m}{b} \right) \right), \tag{5}$$

where the parameters a, h_m, and b determine the shape of the optimal velocity function, which increases monotonically as a function of $s(t)$, and asymptotically plateaus at a maximum speed as $s \to \infty$. Consequently the vehicle accelerates and decelerates to achieve the optimal velocity:

$$\dot{v}(t) = \alpha \left(V(s(t)) - v(t) \right), \tag{6}$$

where the parameter α determines the sensitivity of the stimulus, which is the difference between the desired velocity $V(s(t))$ and the actual velocity $v(t)$. The parameter vector of this model is $\theta = [\alpha, a, h_m, b]$. The corresponding equilibrium initial condition for (2) under (6) is:

$$x_0^* = \left[h_m - b * \tanh^{-1} \left(\tanh \left(\frac{h_m}{b} \right) - \frac{u_0}{a} \right), u_0 \right]^T. \tag{7}$$

Follow-the-Leader (FTL) model   The third model is one of the simplest follow-the-leader variations of the Gazis-Herman-Rothery (GHR) car-following model, which originated from research conducted by General Motors in the 1950s [15–17]. We consider the following form:

$$\dot{v}(t) = \frac{C \Delta v(t)}{s(t)^\gamma}, \tag{8}$$

where the parameter C and γ are constants describing the sensitivity of $\Delta v(t)$ and $s(t)$, respectively, or the acceleration. The parameter vector of interest is $\theta = [C, \gamma]$. The corresponding equilibrium initial condition x_0^* for (2) under (8) is:

$$x_0^* = [s_0, u_0]^T. \tag{9}$$

Intelligent driver model (IDM)   The intelligent driver model was proposed in [62] to model a realistic driver behavior, such as asymmetric accelerations and decelerations. It is of the form:

$$\dot{v}(t) = a \left[1 - \left(\frac{v(t)}{v_f} \right)^4 - \left(\frac{s^*(v(t), \Delta v(t))}{s(t)} \right)^2 \right] \tag{10}$$

where the desired space gap s^* is defined as:

$$s^*(v(t), \Delta v(t)) = s_j + v(t)T + \frac{v(t)\Delta v}{2\sqrt{ab}}. \tag{11}$$

The parameters of the model $\theta = [s_j, v_f, T, a, b]$ are the freeflow speed v_f, the desired time gap T, the jam distance s_j, the maximum acceleration a and the desired deceleration b. The corresponding equilibrium initial condition x_0^* is:

$$x_0^* = \left[\frac{s_j + u_0 * T}{\sqrt{1 - (u_0/v_f)^4}}, u_0 \right]^T. \tag{12}$$

7

4 Methodology

In this section we highlight the methods for identifiability analysis. The structural local identifiability is carried out through a *differential geometry* method, and the practical identifiability considering a full experimental setup and measurement error is examined using a *numerical direct test* method.

4.1 Differential geometry framework for structural identifiability

We first introduce the concepts of structurally global and structurally local identifiability via precise definitions. Note that the following definitions are modified from references such as [19, 40, 72, 75], in order to suit the specific format of our analysis.

Definition 1. *Let $\theta \in \mathbb{R}^{n_p}$ denote the generic parameter vector, $\mathcal{X}_0$ a set of generic initial conditions and $\mathcal{U}$ a set of admissible inputs. Let $y(t, \theta, x_0, u)$ be the output function from the state-space model (1). If for all $t > 0$,*

$$y(t, \theta, x_0, u) \equiv y(t, \theta^*, x_0, u) \Rightarrow \theta = \theta^* \tag{13}$$

for almost all $\theta \in \mathbb{R}^{n_p}$, almost all $x_0 \in \mathcal{X}_0$ and every input $u \in \mathcal{U}$ then the model is said to be structurally globally identifiable.

Definition 2. *A dynamical system given by (1) is structurally locally identifiable (s.l.i.) if for almost all $\theta \in \mathbb{R}^{n_p}$ there exists a neighborhood $\mathcal{N}(\theta)$ such that, for all $\theta_1, \theta_2 \in \mathcal{N}(\theta)$, the implication (13) holds for all $t > 0$.*

Remark 1. *Notice that Definitions 1. and 2. consider the generic parameter vector and initial conditions, and thus may not be applicable to parameters and initial conditions that fall into a measure zero set.*

Remark 2. *We consider admissible inputs are polynomial inputs of degree n (with n non-zero time derivative) that enables structurally identifiable systems. In particular, no admissible input exists for structurally unidentifiable systems.*

The differential geometry framework [67] considers identifiability as an augmented observability property for a general nonlinear system of ODEs such as (1), and it can be evaluated in the same manner. The main idea of identifiability consists in considering the parameters θ as additional states with zero dynamics. Hence, the parameter-augmented state and the dynamics become $\tilde{x} = [x, \theta] \in \mathbb{R}^{n_{\tilde{x}}}$, with $n_{\tilde{x}} = n_x + n_\theta$ and

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ \theta \end{bmatrix} \quad \text{and} \quad \begin{cases} \dot{\tilde{x}}(t) = f(\tilde{x}(t), u(t)) \\ y(t) = g(\tilde{x}(t), u(t)) \\ \tilde{x}(0) = \tilde{x}_0 = [x_0, \theta]^T. \end{cases} \tag{14}$$

In the general case of a nonlinear system with time-varying input, to identify the parameters we need to take into account the changes to the output due

to the changing input in the augmented state. Thus we use the extended Lie derivatives. The extended Lie derivatives can, then, be used to build the rows of the observability-identifiability matrix $\mathcal{O}_I$ [27,71], whose rank is to be evaluated for the identifiability analysis. The Lie derivative of $g(\tilde{x}(t), u(t))$ in the direction of $f(\tilde{x}(t), u(t))$ is given by:

$$L_f g(\tilde{x}, u) = \frac{\partial g(\tilde{x}, u)}{\partial \tilde{x}} f(\tilde{x}, u) + \sum_{j=0}^{j=\infty} \frac{\partial g(\tilde{x}, u)}{\partial u^{(j)}} u^{(j+1)} \tag{15}$$

where $u^{(j)}$ is the jth time derivative of the input u. The Lie derivatives of higher order are:

$$L_f^i g(\tilde{x}, u) = \frac{\partial L_f^{i-1} g(\tilde{x}, u)}{\partial \tilde{x}} f(\tilde{x}, u) + \sum_{j=0}^{j=\infty} \frac{\partial L_f^{i-1} g(\tilde{x}, u)}{\partial u^{(j)}} u^{(j+1)}. \tag{16}$$

The observability-identifiability matrix $\mathcal{O}_I(\tilde{x}, u)$ for a general nonlinear system (14) becomes

$$\mathcal{O}_I(\tilde{x}, u) = \begin{pmatrix} \frac{\partial}{\partial \tilde{x}} g(\tilde{x}, u) \\ \frac{\partial}{\partial \tilde{x}} (L_f g(\tilde{x}, u)) \\ \frac{\partial}{\partial \tilde{x}} (L_f^2 g(\tilde{x}, u)) \\ \vdots \\ \frac{\partial}{\partial \tilde{x}} (L_f^{n_{\tilde{x}}-1} g(\tilde{x}, u)) \end{pmatrix} \tag{17}$$

Finally, a nonlinear observability-identifiability condition can be used to infer the structural local identifiability of (14):

Theorem 1 (Nonlinear Observability – Identifiability Condition (OIC)). *If a model (14) satisfies* $\mathrm{rank}(\mathcal{O}_I(\tilde{x}_0, u)) = n_x + n_\theta$, *with* $\mathcal{O}_I(\tilde{x}_0, u)$ *given by (17) and* $\tilde{x}_0$ *being a (possibly generic) point in the augmented state space, the model is locally observable and locally structurally identifiable in a neighborhood* $\mathcal{N}(\tilde{x}_0)$ *of* $\tilde{x}_0$.

Remark 3. *The differential geometry approach yields results that are valid almost everywhere, i.e., for all values of the system variables (initial conditions and inputs) except for a set of measure zero.*

Throughout the rest of the paper, we focus on analyzing structural local identifiability for almost all initial conditions in the augmented initial state $\tilde{x}_0$, i.e., we compute $\mathcal{O}_I(\tilde{x}_0, u)$ symbolically with $\tilde{x}_0 = [x_0, \theta]^T$. We also explore the special cases for $\tilde{x}_0$ (the initial conditions and parameters that belong to a set of measure zero) such that structural identifiability is lost. More details of the setup and analysis are presented in Section 5.

4.2 Numerical direct test

The *direct test* is a conceptually straightforward way to test identifiability considering a fully-defined system (i.e., the given initial condition and input trajectory are known). The direct test offers analysis that differs from structural

identifiability in that it looks for the worst-case unidentifiable parameters in the following sense. It poses the identifiability problem as for finding the maximally distinct (as specified by an objective function) parameters θ_1 and θ_2 such that the output of two systems (which are identical except for the parameters) differ by no more than a threshold ϵ (possibly equal to zero). It is an idea first proposed in [73] as a numerical alternative to algebraic computation.

The generic form of the direct test reads:

$$
\begin{aligned}
\underset{\theta_1,\theta_2\in\Theta}{\text{maximize}} \quad & d(\theta_1,\theta_2) \\
\text{subject to} \quad & e(\theta_1,\theta_2) \le \epsilon,
\end{aligned}
\tag{18}
$$

where the objective function $d(\theta_1,\theta_2)$ is the distance between two parameters in the parameter space Θ. The constraint $e(\theta_1,\theta_2) \le \epsilon$ caps the difference between the output of the system under θ_1 and θ_2 at ϵ, which is a small, user-defined threshold of measurement error. In order to evaluate $e(\theta_1,\theta_2)$, one must solve the ODE (1) for the same known initial condition and the same given input trajectory under the two parameters θ_1 and θ_2. This problem formulation finds the most distinct (as quantified by d) parameters such that the fully defined system produces similar outputs (or the same outputs when $\epsilon = 0$). The difference $e(\cdot,\cdot)$ is only due to θ for a fully-defined system. Therefore $e(\cdot,\cdot)$ is a function of only two parameters θ_1 and θ_2.

Denote the decision variables of (18) at optimality as θ_1^*, θ_2^*, and the objective function value at optimality as $\delta^* := d(\theta_1^*, \theta_2^*)$. Naturally, under the tightest constraint $\epsilon = 0$, all parameters in Θ for system (1) with a specific and known initial condition and input trajectory produce a unique output if and only if $\delta^* = 0$. When this occurs, we say that the system (1) with the prescribed initial and input condition is identifiable; the same system is unidentifiable under the same initial and input conditions if $\delta^* > \delta$, where δ is a small positive number that is problem-specific.

When measurement error is allowed, we consider a relaxed constraint, $\epsilon > 0$. If δ^* is small, then we say the system is practically identifiable, i.e., similar outputs can be produced only by parameters that are similar. On the other hand, if δ^* is large, we found two distinct parameters that generate similar outputs, indicating that the same system is practically unidentifiable.

Results on the change of δ^* as ϵ changes provides insights on the sensitivity of identifiability with respect to the measurement error, and will be presented in the numerical experiment (Section 6).

5 Structural identifiability analysis

In this section, we analyze structural local identifiability of the car-following system (2) under the various car following models in Section 3. The analysis reveals that all the systems are structurally locally identifiable, i.e., it is theoretically possible to uniquely infer the unknown parameters under any admissible

input and almost all initial conditions. We also explore the specific initial conditions (those belonging to a set of measure zero) such that each model becomes unidentifiable.

The organization of this section is the following. First we give an overview of the analysis, including the toolbox for implementing the differential geometry analysis for a given car following model, and interpret the results of the analysis. Next, we provide the structural local identifiability analysis for the CTH-RV (3) model in detail, by showing the explicit observability-identifiability matrices $\mathcal{O}_I(\tilde{x}_0, u)$ under both a generic and specific initial conditions, and a set of admissible inputs for which the system is structurally identifiable. The structural local identifiability results of the other models ((6), (8) and (10)) are then summarized.

5.1 Implementation overview

We deploy a direct implementation of the framework for structural identifiability through *STRuctural Identifiability taKen as Extended-Generalized Observability with Lie Derivatives and Decomposition* (STRIKE-GOLDD) [69, 72], an open-source MATLAB toolbox that computes the observability-identifiability matrix $\mathcal{O}_I(\tilde{x}_0, u)$ and analyzes the local structural parameter identifiability, state observability, and input reconstructability of nonlinear dynamic models of ODEs. Because only parameter identifiability is concerned in this paper as the states and the inputs can be directly measured, we use the toolbox solely for the purpose of identifiability.

In addition, we use STRIKE-GOLDD to also determine *a priori* the minimum degree n of a polynomial input $u(t)$ to enable a structurally identifiable system [71]. For example, if $\mathcal{O}_I$ has full rank under $u^{(n)}(t) = 0$ but not $u^{(n-1)}(t) = 0$, then an input $u(t)$ of degree n or above is sufficient for structural identifiability. Otherwise, $u(t)$ with degree higher than n should be explored.

For structurally unidentifiable systems, no admissible input exists. For structurally identifiable systems, the degree n (i.e., number of non-zero time derivatives) for an input to be admissible depends on the initial condition. We explore the admissible input condition for structurally identifiable car-following systems under both generic initial conditions and a special set of initial conditions where higher-order input may be required to be admissible.

5.2 Structural local identifiability analysis for CTH-RV

Recall CTH-RV ODE (3). The parameter-augmented system dynamics can be written as

$$\begin{cases} \dot{\tilde{x}}_1(t) = \dot{s}(t) = u(t) - v(t) \\ \dot{\tilde{x}}_2(t) = \dot{v}(t) = k_1(s(t) - \tau v(t)) + k_2(u(t) - v(t)) \\ \dot{\tilde{x}}_{3-5}(t) = \dot{\theta}(t) = 0 \\ y(t) = s(t) \\ \tilde{x}(0) = [x_0, \theta]^T. \end{cases} \tag{19}$$

The following propositions and proofs show the structural identifiability of CTH-RV from the calculated observability-identifiability matrices.

Proposition 1. *According to Definition 2, the CTH-RV system* (19) *is structurally locally identifiable under almost all initial condition x_0 and an admissible input $u(t)$ with degree $n \geq 0$ up to a set of measure zero.*

Proof. Consider almost all initial state conditions $x_0 = [s_0, v_0]^T$. We prove that input $u(t)$ with degree $n \geq 0$ is admissible for structural local identifiability, i.e., $\mathcal{O}_I(\tilde{x}_0, u)$ is full rank with constant input $u(t) = u_0$. This implies that $\mathcal{O}_I(\tilde{x}_0, u)$ for $n > 0$ is also full rank [71].

Constructing the observability-identifiability matrix $\mathcal{O}_I(\tilde{x}_0, u)$ requires at least $n_{\tilde{x}} - 1 = 4$ extended Lie Derivatives, where $n_{\tilde{x}} = 5$ is the dimension of the augmented state. The four extended Lie Derivatives are:

$$L_f^1 g(\tilde{x}, u) = u(t) - v(t)$$
$$L_f^2 g(\tilde{x}, u) = k_2(v(t) - u(t)) - k_1(s(t) - \tau v(t))$$
$$L_f^3 g(\tilde{x}, u) = k_1(v(t) - u(t)) - (k_2 + k_1\tau)(k_2(v(t_-u(t)) - k_1(s(t) - \tau v(t))) \quad (20)$$
$$L_f^4 g(\tilde{x}, u) = -(k_1 - (k_2 + k_1\tau)^2)(k_2 * (v(t) - u(t)) - k_1(s(t) - \tau v(t)))$$
$$\qquad - k_1(k_2 + k_1\tau)(v(t) - u(t)).$$

Consequently, $\mathcal{O}_I(\tilde{x}_0, u)$ for system (19) is a 5×5 matrix:

$$
\mathcal{O}_I(\tilde{x}_0, u) =
$$
$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 \\
-k_1 & k_2 + k_1\tau & \tau v_0 - s_0 & v_0 - u_0 & k_1 v_0 \\
-k_1(k_2 + k_1\tau) & (k_2 + k_1\tau)^2 - k_1 & o_{43} & o_{44} & o_{45} \\
o_{51} & o_{52} & o_{53} & o_{54} & o_{55}
\end{bmatrix} \quad (21)
$$

with specific entries defined as follows:

$$o_{43} = v_0 - u_0 + (k_2 + k_1\tau)(s_0 - \tau v_0) - \tau(k_2(v_0 - u_0)) - k_1(s_0 - \tau v_0)$$
$$o_{44} = k_1(s_0 - \tau v_0) - k_2(v_0 - u_0) - (k_2 + k_1\tau)(v_0 - u_0)$$
$$o_{45} = -k_1(k_2(v_0 - u_0) - k_1(s_0 - \tau v_0)) - k_1 v_0(k_2 + k_1\tau)$$
$$o_{51} = k_1(k_1 - (k_2 + k_1\tau)^2)$$
$$o_{52} = -k_1(k_2 + k_1\tau) - (k_2 + k_1\tau)(k_1 - (k_2 + k_1\tau)^2)$$
$$o_{53} = (s - \tau v_0)(k_1 - (k_2 + k_1\tau)^2) - (k_2 + k_1\tau)(v_0 - u_0) + (2\tau(k_2 + k_1\tau) - 1)$$
$$\qquad (k_2(v_0 - u_0) - k_1(s_0 - \tau v_0)) - k_1\tau(v_0 - u_0)$$
$$o_{54} = (2k_2 + 2k_1\tau)(k_2(v_0 - u_0) - k_1(s_0 - \tau v_0)) -$$
$$\qquad k_1(v_0 - u_0) - (k_1 - (k_2 + k_1\tau)^2)(v_0 - u_0)$$
$$o_{55} = k_1^2(v_0 - u_0) - k_1 v_0(k_1 - (k_2 + k_1\tau)^2).$$

Analytically, it is difficult to prove that the matrix $\mathcal{O}_I(\tilde{x}_0, u)$ is full rank. A symbolic calculator such as MATLAB can provide a general case: the matrix in general is not rank deficient (see an example in Appendix A.1). In this case, the system is structurally locally identifiable according to Theorem 1. Since constant input $u(t) = u_0$ is admissible, $u(t)$ with degree $n \geq 0$ is also admissible according to [71]. $\qquad \square$

However, Theorem 1 provides a general result that is valid for all values of the state and parameters except for a set of measure zero. It is possible that for some special initial conditions the matrix is not full rank. The symbolic calculator leads to generic conclusions on the rank of a matrix and fails to capture a specific set (of measure zero). Therefore, it is possible to categorize the system as locally structurally identifiable whereas for special initial conditions the result is uninformative. The following proposition gives an example of special initial condition that leads to unidentifiable CTH-RV system:

Proposition 2. *The CTH-RV system* (19) *is unidentifiable given an equilibrium initial condition* (22) *and constant input* $u(t) = u_0$.

Proof. We prove that the system (19) is unidentifiable under constant input $u(t) = u_0$ with an equilibrium initial conditions specified as

$$x(0) = x_0^* = \begin{bmatrix} \tau u_0 \\ u_0 \end{bmatrix}. \tag{22}$$

where u_0 is the initial value of the lead vehicle velocity. The equilibrium initial condition results in $\dot{s}(0) = 0$ and $\dot{v}(0) = f_{\text{CTHRV}}(u_0, x_0^*, \theta) = 0$. The same four Lie Derivatives (20) will be used to calculate $\mathcal{O}_I(\tilde{x}_0, u)$. However, due to the specified initial conditions and the constant input condition, $\mathcal{O}_I(\tilde{x}_0, u)$ becomes:

$$\mathcal{O}_I(\tilde{x}_0, u) =$$
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ -k_1 & k_2 + k_1\tau & 0 & 0 & k_1 u_0 \\ k_1(k_2 + k_1\tau) & k_1 - (k_2 + k_1\tau)^2 & 0 & 0 & -k_1 u_0(k_2 + k_1\tau) \\ o_{51} & o_{52} & 0 & 0 & o_{55} \end{bmatrix} \tag{23}$$

$$o_{51} = k_1(k_1 - (k_2 + k_1\tau)^2)$$
$$o_{52} = -k_1(k_2 + k_1\tau) - (k_2 + k_1\tau)(k_1 - (k_2 + k_1\tau)^2)$$
$$o_{55} = -k_1 u_0(k_1 - (k_2 + k_1\tau)^2).$$

It is obvious that $\text{rank}(\mathcal{O}_I(\tilde{x}_0, u)) = 3$, and therefore this system is structurally unidentifiable under equilibrium initial conditions. Furthermore, the specific unidentifiable parameter(s) can be detected by removing each of the columns of $\mathcal{O}_I(\tilde{x}_0, u)$ and recalculating the rank. If the rank does not change after removing the ith column, than the corresponding ith variable in the augmented state $\tilde{x}$

13

is unidentifiable (or observable) [72]. It is straightforward to see that removing the 3rd or the 4th column of $\mathcal{O}_I(\tilde{x}_0, u)$ does not change the rank, and therefore, the corresponding parameters k_1 and k_2 are unidentifiable. $\qquad\square$

Corollary 1. *Given an equilibrium initial condition x_0^* for the CTH-RV system* (19), *the admissible input required to enable structural identifiability is with degree $n \geq 1$.*

In the case of equilibrium initial condition, a constant input is not admissible to enable structural identifiability. Therefore, time-varying input may be needed to excite the system. We check the order of the input required by replacing the higher order derivatives of $u(t)$ in $\mathcal{O}_I$ with zero and recalculating rank($\mathcal{O}_I$). This procedure gives a full rank ($\mathcal{O}_I$) when $\ddot{u} = 0$, as long as $\dot{u} \neq 0$. Therefore an input with degree $n \geq 1$ is admissible for structural identifiability of CTH-RV, even if the system starts from an equilibrium initial condition.

Another unidentifiable case that *differential geometry* fails to detect constitutes a special relationship between the initial condition and the parameters. As long as such relationship is established, there exists no admissible input, and the initial conditions need not to be at equilibrium. This case is demonstrated below.

Proposition 3. *Given a generic initial condition $x_0 = [s_0, v_0]^T$ for the CTH-RV system* (19), *if $\tau = \frac{s_0}{v_0}$ and $k_2 = \frac{v_0}{s_0}$, no admissible input exists.*

Proof. Please see the Appendix A.2 for details. $\qquad\square$

Remark 4. *Measuring the space gap alone $y(t) = s(t)$ leads to the same identifiability results as measuring both states $y(t) = [s(t), v(t)]^T$. One can arrive to this result by checking the rank of the observability-identifiability matrix $\mathcal{O}_I(\tilde{x}_0, u)$. This finding is in agreement with [57, 64] that using space gap instead of velocity or acceleration profile to calibrate a car-following model leads to sensitivity in the objective function.*

Visualization   As an illustration of the above analysis for CTH-RV, we visualize the output differences $e(\cdot, \cdot)$ corresponding the settings of Propositions 1-3. We first fix one parameter as θ_{true}, and compare the output error under another parameter θ. Let $e(\theta, \theta_{\text{true}})$ be defined as the *mean-squared-error* (MSE):

$$e(\theta, \theta_{\text{true}}) = \int_0^{\mathcal{T}} \|y(t, \theta_{\text{true}}) - y(t, \theta)\|_2^2 dt, \tag{24}$$

which describes the error between the true output simulated by θ_{true} and the output simulated by θ, under the same initial conditions and input. Let $\mathcal{T}$ denote the total time over which the ODE is solved.

We visualize $e(\theta, \theta_{\text{true}})$ as one sweeps over the parameter space for CTH-RV (see Figure 1a-1c). Figure 1a shows that the error is zero only at $\theta = \theta_{\text{true}}$, indicating an identifiable system. Figure 1b and 1c, on the contrary, show that

Table 1: Structural local identifiability summary.

Model f_{CF}	Parameters θ	Admissible polynomial input $u(t)$ degree n	
		Generic (almost all) x_0	Equilibrium $x_0^*\dagger$
CTH-RV	$k_1,\, k_2,\, \tau$	$n \geq 0$	$n \geq 1$
OV	$\alpha,\, a,\, b,\, h_m$	$n \geq 0$	$n \geq 1$
FTL	$C,\, \gamma$	$n \geq 0$	N/A
IDM	$a,\, b,\, s_j,\, v_f,\, T$	$n \geq 0$	$n \geq 1$

$\dagger$ The equilibrium initial condition for each model is listed in Section 3.

even for a structurally locally identifiable system, there exists an initial condition, parameters, and input combinations such that distinct parameters exist and produce the same output, as suggested by Proposition 2 and 3. Specifically, we see in Figure 1b that k_1 and k_2 are unidentifiable under an equilibrium initial condition and a constant input. This is consistent with the finding from [78]; in Figure 1c, we see that that k_1 is unidentifiable when the initial speed and space gap satisfies $s_0 = \tau v_0$, even though the initial condition is not at a equilibrium state (i.e., $v_0 \neq u_0$).

## 5.3 Structural local identifiability analysis for other models

Following the same procedure as was used to analyze the CTH-RV, we check the observability-identifiability matrix $\mathcal{O}_I(\tilde{x}_0, u)$ for the other three models, i.e., the OV model (6), the FTL model (8) and the IDM (10). The structural identifiability results for each of the models under almost all initial conditions and the corresponding *a priori* set of admissible inputs are summarized in Table 1. For specific initial conditions (e.g., an equilibrium initial condition), the system may require a higher-order input to enable structural identifiability (also see Table 1).

Table 1 suggests that all the models are structurally locally identifiable, i.e., parameters for each model can be uniquely identified under any admissible input given almost all initial conditions, whether it is known or not, even if only space-gap is being measured. Moreover, a constant input degree $n = 0$ is an admissible input to enable structural identifiability of all the models. Additionally, all inputs $u(t)$ with degree $n \geq 0$ are also admissible to enable structural identifiability.

However, the analysis does not preclude the existence of an initial condition (e.g., an equilibrium initial condition) for which the models become unidentifiable for a constant input. For some models, higher-order inputs are required to enable structural identifiability. For CTH-RV, OV and IDM, any input with degree $n \geq 1$ is an admissible input to enable identifiablity for the equilibrium initial condition. On the other hand, the FTL completely loses identifiability given an equilibrium initial condition $v_0 = u_0$, since no admissible inputs can
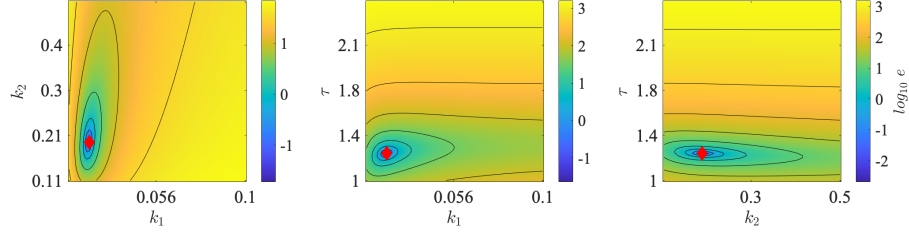
(a) An identifiable scenario: initial condition $x_0 = [60, 20]^T$, input $u(t) = 31$. $\theta_{\text{true}} = [k_1, k_2, \tau]^T = [0.0216, 0.1943, 1.2293]^T$.
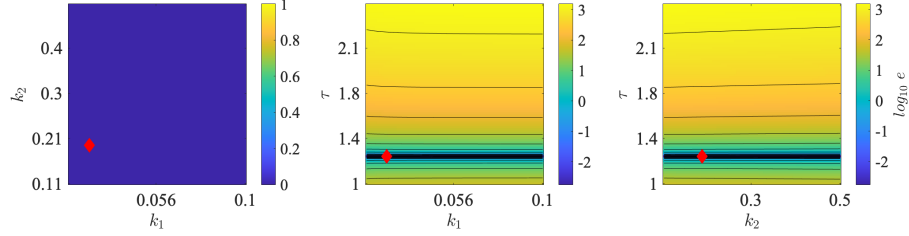
(b) An unidentifiable scenario: equilibrium initial condition $x_0^* = [38.1, 31]^T$ and $u(t) = 31$. $\theta_{\text{true}} = [k_1, k_2, \tau]^T = [0.0216, 0.1943, 1.2293]^T$.
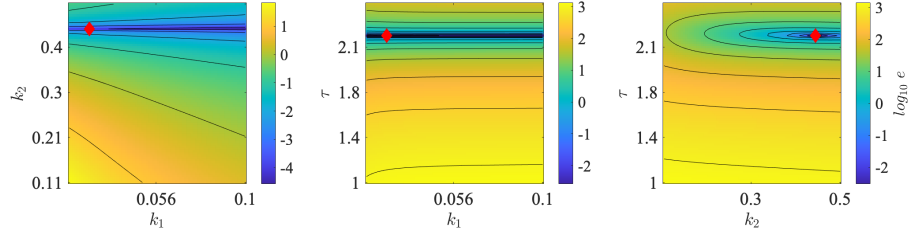
(c) An unidentifiable scenario: initial condition $x_0 = [72.7, 32.5]^T$ and $u(t)$ is shown in Figure 2. $\theta_{\text{true}} = [k_1, k_2, \tau]^T = [0.0216, 0.4472, 2.2361]^T$.

Figure 1: Visualization of $e(\theta, \theta_{\text{true}})$ for CTH-RV. Red diamond indicates θ_{true}. Even for structurally locally identifiable systems, there may exist an initial condition and input combination such that certain parameter become unidentifiable.

Table 2: Parameter bounds.

Models	Parameter θ	Lower bound $\theta_{\mathbf{min}}$	Upper bound $\theta_{\mathbf{max}}$
CTHRV	$[k_1, k_2, \tau_h]$	[0.001, 0.01, 0.1]	[1, 1, 3]
OV	$[\alpha, a, h_m, b]$	[0.5, 10, 2, 18]	[3.3, 32, 30, 45]
FTL	$[C, \gamma]$	[100, 1]	[600, 3]
IDM	$[s_j, v_f, T, a, b]$	[3, 21, 0.1, 0.1, 0.5]	[25, 41, 3, 3, 5]

be found to enable identifiability.

Overall, the main outcome of the structural local identifiability analysis is that we find no fundamental intrinsic structural problems in any of the analyzed models. This means it is possible to design an experiment (i.e., chose an initial condition and an input) in order to uniquely infer the unknown model parameters.

The structural local identifiability analysis provides theoretical results that are generally valid for almost all numerical values of the initial conditions and the parameters, but the test could fail to detect unidentifiable scenarios (even for a simple CTH-RV model, it fails to detect a special initial condition (Proposition 2) or parameter dependency (Proposition 3) that causes non-identifiability). In practice, identifiability would be impossible when close to these special conditions (because of the rank deficiency of the observability-identifiability matrix). The probability of being in a small open set around these measure-zero sets, although unlikely, is not zero. Therefore, it is important to ask a different question in practice: *given the experimental design (i.e., given a specific initial condition and input trajectory that are known), whether there exists multiple distinct parameters that would produce the same output.* If multiple distinct parameters are indistinguishable in the output space, then the corresponding model(s) are unidentifiable. We proceed to numerical methods in the next section to address this question.

6 Practical identifiability analysis

Although the initial condition and parameter subspace that loses structural identifiability for CTH-RV can be analytically solved for, it is much challenging to exhaustively find the special cases for other models composed of irrational functions. To this end, we use a numerical method to analyze a more specific notion of identifiability: *given a specific initial and input condition that are known and non-trivial, i.e., non-equilibrium initial condition and time-varying input, are there distinct parameters that produce the same or similar output?* If such distinct parameter sets are found, then the corresponding model is unidentifiable to this specific experimental setup, irrespective of the cause (initial condition or parameters fall into a measure zero set or the input is not informative enough). In this section, we present identifiability analysis using *direct test* for models (3), (6), (8) and (10). First, we provide the details of the numerical experiment

setup, including a specific optimization formulation. Next, we solve (18) with 1) the tightest constraint $\epsilon = 10^{-6}$ and 2) the relaxed constraint $\epsilon > 10^{-6}$. Because the choice of the cut-off threshold ϵ is arbitrary, we incrementally increase ϵ to observe the sensitivity of model calibration to the measurement error.

For δ-identifiability at both $\epsilon = 10^{-6}$ and $\epsilon > 10^{-6}$, we first provide the detailed analysis on the CTH-RV model (3), and summarize the results for the others.

6.1 Implementation details

Consider the following optimization problem formulation for (18):

$$
\begin{aligned}
&\underset{\theta_{1,i},\theta_{2,i}\in[\theta_{\min,i},\theta_{\max,i}],\forall i}{\text{maximize}} && d(\theta_1,\theta_2) = \frac{1}{\sqrt{n_\theta}}\sqrt{\sum_{i=1}^{n_\theta}\left(\frac{\theta_{1,i}-\theta_{2,i}}{\theta_{\max,i}-\theta_{\min,i}}\right)^2} \\
&\text{subject to} && e(\theta_1,\theta_2) = \frac{1}{K}\sum_{k=0}^{k=K}\|y_k(\theta_1)-y_k(\theta_2)\|_2^2 \leq \epsilon.
\end{aligned}
\tag{25}
$$

The subscript i on θ denotes the ith element in the parameter vector. The resulting distance (objective function) is a normalized Euclidean distance scalar between 0 and 1 bounded by $\theta_{\min}$ and $\theta_{\max}$ (Table 2), i.e., $d(\theta_{\min},\theta_{\max}) = 1$. The trajectory difference $e(\cdot,\cdot)$ in the constraint is a mean squared error to measure the difference between $y_k(\theta_1)$ and $y_k(\theta_2)$, i.e., the discrete-time space-gap simulated with an Euler method using θ_1, θ_2, respectively. The number of decision variables is $2 \times n_\theta$ where n_θ is the dimension of the parameters for each model. Recall that we denote the solution to (18) as δ^* and the corresponding parameter pair as (θ_1^*, θ_2^*).

The optimization solver used to solve (25) is `patternsearch` in the MATLAB global optimization toolbox [43]. The `patternsearch` is a gradient-free optimization solver that works with potentially non-smooth objective functions. More details regarding the solver can be found in [3].

6.2 Practical identifiability with error-free measurement

In the numerical setting, we specify the tightest constraint at approximately zero (specifically, $\epsilon = 10^{-6}$) and solve for the optimization problem (25) with a time-varying input profile (Figure 2) and a fixed, non-equilibrium initial condition $x_0 = [72.7, 32.5]^T$. We use this initial condition and input for all of the models in the analysis for the remaining of this paper. Given this specific experiment setup, direct test aims to find indistinguishable (thus unidentifiable) parameters that separate farthest in the parameter space. This specific initial condition is chosen to illustrate the utility of the direct test, i.e. given an arbitrary but non-equilibrium initial condition, can we find two distinct parameters that generate the same output.
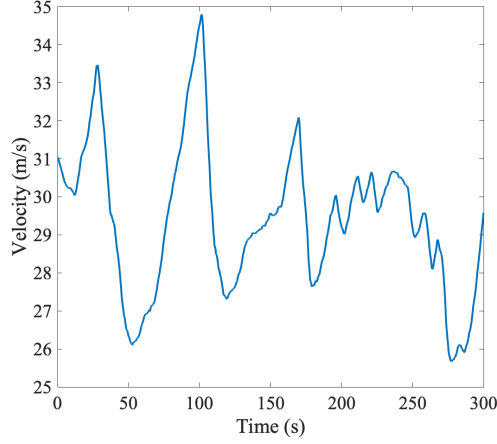
Figure 2: A time-varying lead vehicle velocity profile.

Table 3: Direct test finds two indistinguishable parameter sets for CTH-RV with $\epsilon = 1\text{e-}6$

Model	Parameter	θ_1^*	θ_2^*	δ^*
	k_1	1	0.001	
CTHRV	k_2	0.4472	0.4472	0.5774
	τ	2.236	2.236	
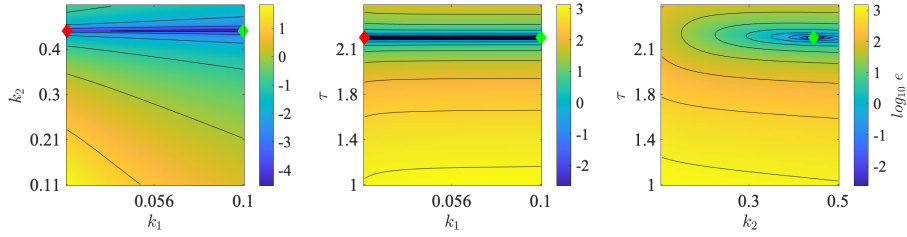
Figure 3: Visualize indistinguishable parameter sets θ_1^* (red diamond) and θ_2^* (green diamond) from Table 3.
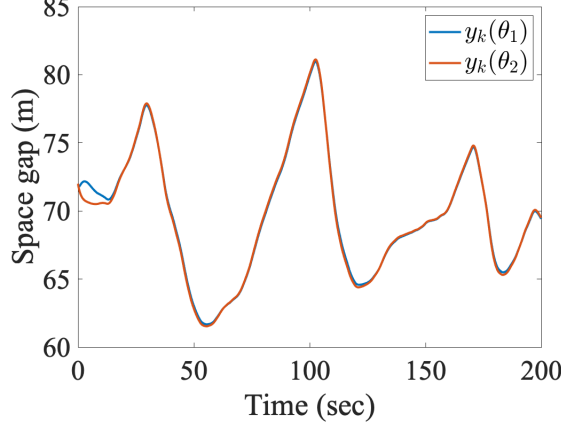
19

Figure 4: Visualize practically indistinguishable parameter sets for CTHRV.

CTH-RV The direct test result for CTH-RV by solving the optimization problem with $\epsilon = 10^{-6}$ is listed in Table 3. Table 3 suggests that given a non-equilibrium lead velocity profile (Figure 2) and an arbitrary, non-equilibrium initial condition that are known, the CTH-RV is unidentifiable. Solving problem (25) results in $\delta^* = 0.577$, meaning there exists two distinct parameters that generate the same output. Figure 3 visualizes this result: θ_1^* and θ_2^* are far apart but produce the same output. The value of $e(\theta, \theta_1^*)$ remains constant along the unidentifiable parameter. Furthermore, the unidentifiable parameters fall into a set such that $\tau = s_0/v_0$ and $k_2 = v_0/s_0$. This corresponds to Proposition 3 where k_1 cannot be uniquely identified.

Additionally, we discovered that when the initial conditions are in a neighborhood of the measure-zero set identified in Proposition 3, CTH-RV can still be practically unidentifiable, as shown in Figure 4. In this example, the initial condition is $s_0 = 72$, $v_0 = 32$, and the parameters are $\theta_1 = [1, 0.394, 2.35]^T$ and $\theta_2 = [0.1, 0.394, 2.35]^T$. A small output difference is obtained ($\epsilon = 0.0745$) with distinct parameters ($\delta^* = 0.5201$). Note that $k_2 \approx \dfrac{v_0}{s_0}$ and $\tau \approx \dfrac{s_0}{v_0}$, which mean that the initial conditions are not exactly in the measure-zero set as shown in Proposition 3, but in a neighborhood of a it.

Other models As for the other three models, the solution to the direct test (25) is summarized in Table 4. The results indicate that the FTL (8) is also unidentifiable under the specific experiment setup ($\delta^* = 0.677$). OV (6) and IDM (10) are practically identifiable ($\delta^* = 0.0117$ and 0.0476, respectively, both are small), which suggests that only parameters that are close to each other can produce the same output. The solutions found by numerical direct test are visualized in Figures 3-7 for the corresponding models.

20

Table 4: Direct test finds indistinguishable parameter sets for other models with $\epsilon =$1e-6

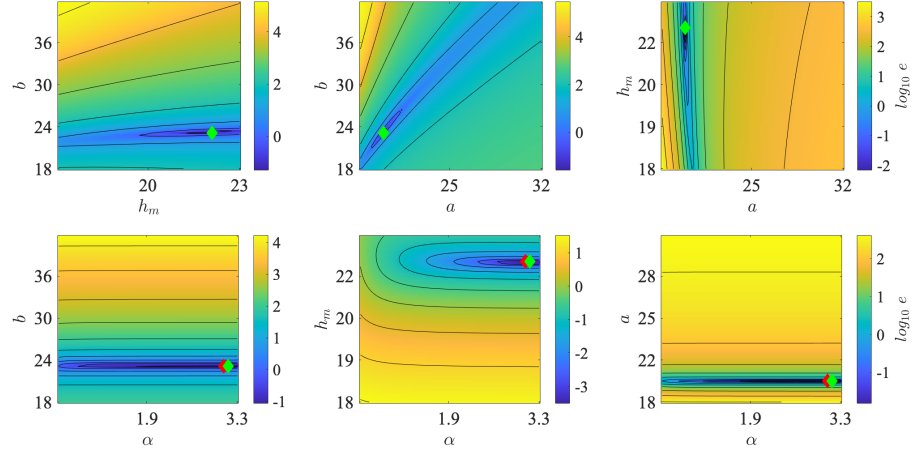| Model	Parameter	θ_1^*	θ_2^*	δ^*
OV	α	3.0772	3.1427	0.0117
	a	19.7485	19.7451	
	h_m	22.2094	22.2122	
	b	23.2986	23.2915	
FTL	C	130.0285	599.9699	0.6766
	γ	1	1.3582	
IDM	s_0	10.5615	8.3913	0.0476
	v_0	35.788	35.771	
	T	2.787	2.903	
	a	2.559	2.494	
	b	3.395	3.395	

Figure 5: Visualize indistinguishable parameter sets θ_1^* (red diamond) and θ_2^* (green diamond) for OV from Table 4.
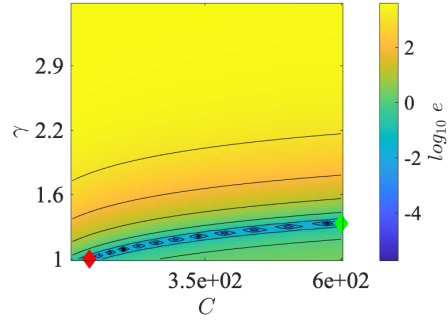
Figure 6: Visualize indistinguishable parameter sets θ_1^* (red diamond) and θ_2^* (green diamond) for FTL from Table 4.
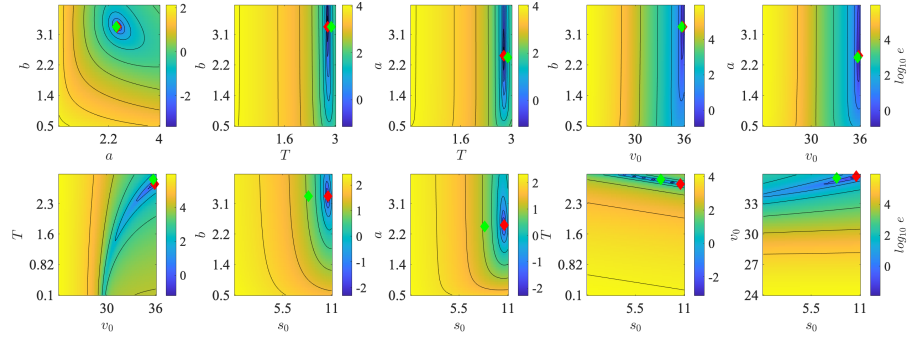

Figure 7: Visualize indistinguishable parameter sets θ_1^* (red diamond) and θ_2^* (green diamond) for IDM from Table 4.

(a) $e(\theta_1, \theta_2) = 0$

(b) $e(\theta_1, \theta_2) = 0.1$

(c) $e(\theta_1, \theta_2) = 1$

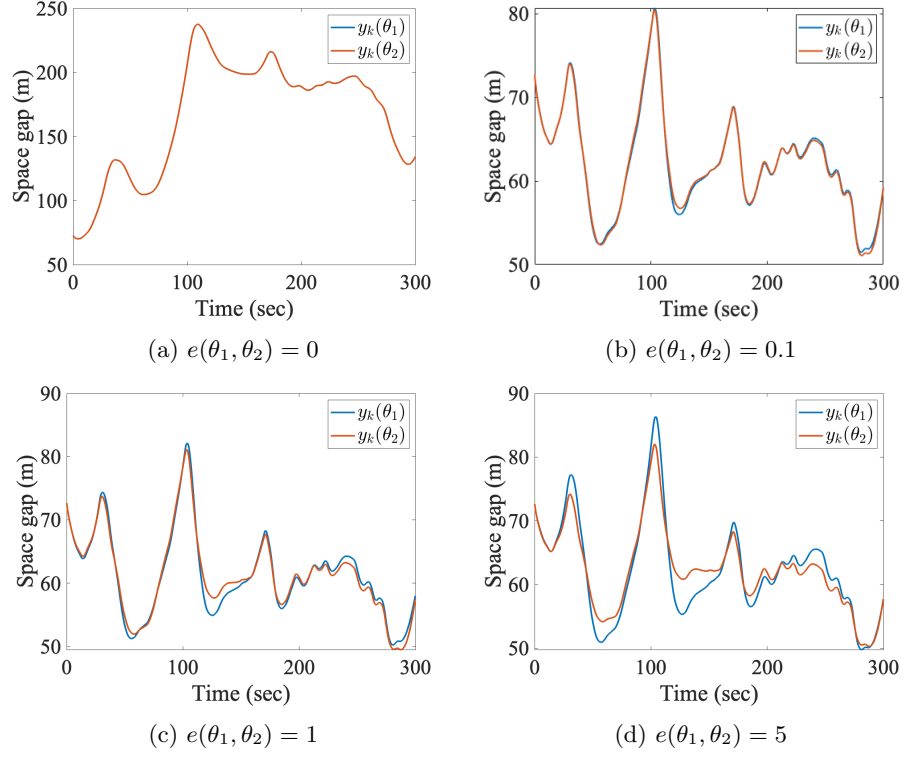(d) $e(\theta_1, \theta_2) = 5$

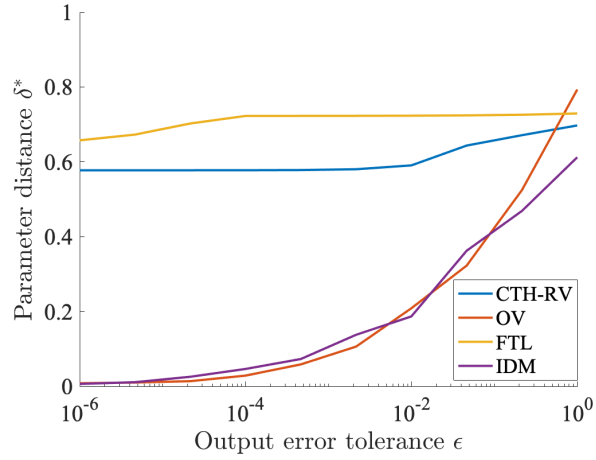Figure 8: A demonstration of various trajectory differences $e(\theta_1, \theta_2)$ of IDM.

Figure 9: The sensitivity of practical identifiability with respect to the error threshold ϵ. The specific setup is $x_0 = [72.7, 32.5]^T$ and input $u(t)$ shown in Figure 2 for all the models.

6.3 Practical identifiability when measurement noise is present

The practical identifiability is nontrivial to quantify because 1) the threshold for δ^* to be identifiable/unidentifiable cannot be easily generalized across all models. Parameters for each model also have different physical meanings. 2) The trajectory error ϵ has multiple possible formulations (RMSE, ARE, MSE, etc.), making a universal cut-off threshold for practical identifiability impossible. Therefore, we provide the sensitivity analysis instead. The analysis is to provide a sense of robustness of parameter estimation with the presence of imperfect measurements (e.g., with measurement errors).

The value of ϵ used in the constraint of the direct test problem is user-defined, and must be specified before solving the numerical problem. Therefore, we vary ϵ values to find the maximum parameter distances for each ϵ (see, for example, $e(\cdot, \cdot) = 0, 0.1, 1$ and 5m^2 for IDM in Figure 8). This is to analyze the sensitivity of the maximum distance between two indistinguishable parameters with respect to the tolerance on the output differences. Previous studies suggest that calibrating on empirical data achieves the lowest MSE of 1-2m^2 in space gap [20, 78], our analysis on practical identifiability will focus on $e(\cdot, \cdot)$ to be less than 1m^2, or $\epsilon \in [0, 1]$.

In particular, we select ϵ equally-spaced in log space between 10^{-6} and 10^0, and solve for δ^* and the corresponding (θ_1^*, θ_2^*) in problem (25). The resulting curves shown in Figure 9 are the solutions δ^* as a function of ϵ in the constraint for each model. The curve is the maximum distance in parameter space such that the output error does not exceed ϵ. Clearly, as the output difference ϵ increases, so does the maximum parameter distance δ^*. Intuitively, given a specific initial condition and lead vehicle trajectory, as measurement error increases, the best-fit parameters are more likely to be non-unique, and the estimated parameters are less robust in the worst case.

Furthermore, Figure 9 suggests that the CTH-RV and the FTL are practically unidentifiable no matter what output error threshold to chose, under the given initial and input conditions. I.e., there is no guarantee to find the unique parameters for the CTH-RV and the FTL even with perfect (error-free) data. On the other hand, the OV and the IDM can be practically identifiable when the threshold for output error is decreased. However, identifiability may not be achieved if the data is contaminated with noises (ϵ is large).

In summary, the *direct test* is a conceptually straightforward method to test identifiability given a particular experimental design (i.e., with a known initial condition and a known input). It is very helpful to know (for model that is structurally locally identifiable) if the experiment leads to practical unidentifiability. Our approach provides a tool that highlights where an experiment can fail in the sense that similar outputs are obtained by distinct parameters. This practical identifiability test is complementary to that of *structural identifiability* analysis, which provides theoretical identifiability results for almost all initial conditions and parameters, but does not preclude the existence of special initial conditions and parameters (those belong to a set of measure zero) that result in

non-identifiability (shown in Proposition 2 and 3). Note also that the *direct test* does not aim to find *all* initial conditions and parameters that fall into a set of measure zero, but simply tests if indistinguishable parameters exist given a full experimental setup, which can be useful in practice. Having a user-defined output difference threshold ϵ gives the *direct test* more flexibility to understand the impacts of output measurement errors.

However, numerical test also has some disadvantages. For example, depending on the optimization solver and the set of starting points, the numerical approach may not find the true global maximum, and thus may misleadingly provide a lower bound on δ^*, rather than the global maximum. In addition, the numerical direct test is performed based on a specific initial condition and input trajectory that are known, which limits its use to inform experimental designs. This approach does not consider practical experimental design, including how to design one or several experiments to best estimate the parameters, which is an interesting direction for further investigation. Nevertheless, it is insightful as a verification tool to diagnose the potential unidentifiability under a specific setup.

7   Conclusion

In this work, we study the structural and practical identifiability of four car-following models.

The structural identifiability analysis is carried through using a *differential geometry* approach. It provides theoretical identifiability results under error-free assumptions for almost all initial conditions and parameters, and informs the admissible input condition that enables structural identifiability. However, it does not preclude the existence of initial conditions and parameters that lose identifiability. The results from the *differential geometry* analysis show that all models are structurally locally identifiable, i.e., all models in theory have unique parameterizations when fitting with space gap data given almost all initial conditions and any admissible input. For some initial conditions such as an equilibrium initial condition, higher-order input is required for CTH-RV, OV and IDM to be identifiable. No admissible input exists for FTL when the initial condition is at equilibrium to enable identifiability.

As a complementary analysis, we use a numerical *direct test* to study practical identifiability given a specific experimental setup (a given known initial condition and input). It also provides insights on the sensitivity of parameter identifiability due to measurement errors on the output. The *direct test* finds that CTH-RV and FTL are practically unidentifiable, meaning that there exists multiple distinct parameters that produce the same output, given a specific experimental setup. It also suggests that OV and IDM are practically identifiable when the measurement error is small. Although only four models are investigated, the provided methods can be applied to other car following models.

The findings also open up questions for future research. For example, we are interested to design an experiment or sets of experiments that avoid problematic

initial conditions for identifiability. Design of reduced order models that allow robust parameter identification while reflecting the dynamics of car following behavior are also of interest.

8   Acknowledgement

A   Appendix

A.1   An example for structurally locally identifiable CTH-RV

Given the symbolic matrix (21), one can find numerical substitutions for all the symbolic variables such that (21) is full rank. Such substitutions belong to a generic initial condition and parameter set. One example is:

$$[k_1, k_2, \tau, u_0, v_0, s_0] = [0.01, 0.12, 1.4, 30, 33, 40]. \tag{26}$$

The corresponding identifiability matrix becomes:

$$\mathcal{O}_I(\tilde{x}_0, u) =$$
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ -0.0100 & 0.134 & 6.20 & 3.00 & 0.330 \\ 0.00134 & -0.00796 & 1.579 & -0.824 & -0.0484 \\ -0.0000796 & -0.000274 & -0.658 & 0.107 & 0.003456 \end{bmatrix}, \tag{27}$$

and rank($\mathcal{O}_I(\tilde{x}_0, u)$)=5.

A.2   Proof for proposition 3

Recall the system of equations for CTH-RV:

$$\dot{x}(t) = \begin{bmatrix} \dot{s}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} u(t) - v(t) \\ k_1(s(t) - \tau v(t)) + k_2(u(t) - v(t)) \end{bmatrix} \tag{28}$$
$$y(t) = x(t)$$

This system in fact can be written as a scalar differential equation:

$$\dot{v}(t) = k_1 \int_0^t (u(\xi) - v(\xi)) d\xi + (-k_1\tau - k_2)v(t) + k_2 u(t). \tag{29}$$

Take the derivative on both sides to get rid of the integral operator:

$$\ddot{v}(t) = k_1(u(t) - v(t))dt + (-k_1\tau - k_2)\dot{v}(t) + k_2\dot{u}(t). \tag{30}$$

Re-arrange to separate the input and output:

$$\ddot{v}(t) + (k_1\tau + k_2)\dot{v}(t) + k_1v(t) = k_1u(t) + k_2\dot{u}(t). \tag{31}$$

Now this is the standard form of a second-order non-homogeneous differential equation. The solution $v(t)$ has two parts, a homogeneous solution (or complementary solution $v_c(t)$, corresponds to the transient response), and a particular solution (or $V_p(t)$, corresponds to the steady state response). The general solution would be the sum of two parts:

$$v(t) = v_c(t) + V_p(t). \tag{32}$$

First let's see the homogeneous solution $v_c(t)$ by solving

$$\ddot{v}(t) + (k_1\tau + k_2)\dot{v}(t) + k_1v(t) = 0. \tag{33}$$

The solution is of the form:

$$v_c(t) = c_1e^{r_1t} + c_2e^{r_2t}, \tag{34}$$

where r_1 and r_2 are the roots of the characteristic equation for (33), and c_1 and c_2 are constants, which can be solved using initial values after the particular solution $V_p(t)$ is obtained.

Moving on to $V_p(t)$. We will show the condition such that the steady state response $V_p(t)$ does not depend on k_1.

For simplicity, let us assume the input (forcing) function as a sinusoidal wave, i.e., $u(t) = a\sin(\omega t)$. Because we can represent almost all functions by Fourier series, which is a superposition of sine and cosine waves, the choice of $u(t)$ does not change the solution. The forcing term, or the RHS of (31) is

$$g(t) = k_1u(t) + k_2\dot{u}(t) = k_1a\sin(\omega t) + k_2a\omega\cos(\omega t). \tag{35}$$

By undetermined coefficients method, we can solve for the particular solution and the corresponding coefficients:

$$V_p(t) = A\sin(\omega t) + B\cos(\omega t), \tag{36}$$

where A, B can be solved by plugging $V_p(t)$ into (31). A, B can be represented in terms of a, ω and the model parameters θ. We will simplify the expressions by denoting A, B as functions of some parameters:

$$\begin{aligned} A &= f_A(a, \omega, \theta) \\ B &= f_B(a, \omega, \theta). \end{aligned} \tag{37}$$

Recall the superposition principle for ODE since ODE is a linear operator, the particular solution will always be a combination of sine and cosine waves with

the same frequency but different magnitudes as the input $u(t)$. In this case, we just need to see if A, B depend on any parameter $p \in \{k_1, k_2, \tau\}$ to make a difference in the steady state response. This can be done by taking the partial derivatives $\dfrac{\partial f_A}{\partial p}$ and $\dfrac{\partial f_B}{\partial p}$ and examine the dependency. In other words:

find the solution $\mathcal{S}_p$ for
$$\frac{\partial f_A}{\partial p} = 0 \text{ and } \frac{\partial f_B}{\partial p} = 0, \ \forall \ p \in \{k_1, k_2, \tau\} \in \mathcal{P} \tag{38}$$

where $\mathcal{P}$ is the set of legal values for parameter p. If the solution set for each p, $\mathcal{S}_p$, does not depend on p, then p is unidentifiable. Solve for (38) we get

$$\begin{aligned}
\mathcal{S}_{k_1} &= \{k_2 = \frac{1}{\tau}\} \\
\mathcal{S}_{k_2} &= \emptyset \\
\mathcal{S}_\tau &= \emptyset.
\end{aligned} \tag{39}$$

This suggests that A, B are independent of k_1 if $k_2 = \dfrac{1}{\tau}$.

Now let us go back to the initial condition and the transient response. In order to use the information $s(0) = s_0$, we'll need to work on the first-order differential equation (29) instead of (31). Simply plug in the initial condition $s(0) = s_0$ and $v(0) = v_0$ in (29) to obtain:

$$\dot{v}(0) = k_1 s_0 + (-k_1 \tau - k_2) v_0 + k_2 u_0, \tag{40}$$

where $\int_0^t (u(\xi) - v(\xi)) d\xi|_{t=0} = s_0$, and u_0 is the initial value of the input. Now if we do the same as before (38) by taking partial derivatives of $\dot{v}(0)$ to find the unidentifiable parameter(s) and the corresponding "condition" in order for that parameter to be unidentifiable, i.e.,

for $p \in \{k_1, k_2, \tau\}$, find the solution $\mathcal{S}_p$ for
$$\frac{\partial \dot{v}(0)}{\partial p} = 0 \ \forall \ p \in \mathcal{P} \tag{41}$$

Solving (41) we get:
$$\begin{aligned}
\mathcal{S}_{k_1} &= \{s_0 = \tau v_0\} \\
\mathcal{S}_{k_2} &= \emptyset \\
\mathcal{S}_\tau &= \emptyset.
\end{aligned} \tag{42}$$

The solution shows that k_1 cannot be identified if $\tau = \dfrac{s_0}{v_0}$ from the transient

response. Substitute $\tau = \dfrac{s_0}{v_0}$ and $k_2 = 1/\tau = \dfrac{v_0}{s_0}$ into the $\mathcal{O}_I(\tilde{x}_0, u)$, we obtain:

$$
\mathcal{O}_I(\tilde{x}_0, u) =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 \\
-k_1 & v_0/s_0 + k_1 s_0/v_0 & 0 & v_0 - u(t) & k_1 v_0 \\
-k_1(v_0/s_0 + k_1 s_0/v_0) & (v_0/s_0 + k_1 s_0/v_0)^2 - k_1 & 0 & o_{44} & o_{45} \\
o_{51} & o_{52} & 0 & o_{54} & o_{55}
\end{bmatrix}
\tag{43}
$$

$o_{44} = -\dot{u}(t) - (v_0/s_0 + k_1 s_0/v_0)(v_0 - u(t)) - v_0(v_0 - u(t))/s_0$

$o_{45} = -k_1 v_0(v_0/s_0 + k_1 s_0/v_0) - k_1 v_0(v_0 - u(t))/s_0$

$o_{51} = k_1(k_1 - (v_0/s_0 + k_1 s_0/v_0)^2)$

$o_{52} = -(v_0/s_0 + k_1 s_0/v_0)(k_1 - (v_0/s_0 + k_1 s_0/v_0)^2) - k_1(v_0/s_0 + k_1 s_0/v_0)$

$o_{54} = \dot{u}(t)(2v_0/s_0 + k_1 s_0/v_0) - k_1(v_0 - u(t)) - (k_1 - (v_0/s_0 + k_1 s_0/v_0)^2)$
$\quad (v_0 - u(t)) - \ddot{u}(t) + (v_0(2v_0/s_0 + 2k_1 s_0/v_0))/s_0$

$o_{55} = k_1 v_0 \dot{u}(t)/s_0 - k_1 v_0(k_1 - (v_0/s_0 + k_1 s_0/v_0)^2) - k_1^2(v_0 - u(t)) +$
$\quad k_1 v_0(v_0/s_0 + k_1 s_0/v_0)(v_0 - u(t))^2/s_0.$

Matrix (43) clearly shows that the above substitution results in rank$(\mathcal{O}_I(\tilde{x}_0, u))$=4. The column corresponding to parameter k_1 is zero, meaning that k_1 is unidentifiable. This result is in agreement with the analytical proof that k_1 does not affect either the transient response or the steady state response. Note that we can choose $u_0 \neq v_0$ such that the initial condition of the system is at non-equilibrium. Please see Figure 1c for a visualization. Hence we complete the proof.

References

[1] F. Abodo, A. Berthaume, S. Zitzow-Childs, and L. Bobadilla. Strengthening the case for a bayesian approach to car-following model calibration and validation using probabilistic programming. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 4360–4367, 2019.

[2] ACC. Adaptive cruise control datasets. `https://acc-dataset.github.io/datasets/`, 2019.

[3] C. Audet and J. E. Dennis. Analysis of generalized pattern searches. *SIAM Journal on Optimization*, 13(3):889–903, 2002.

[4] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Structure stability of congestion in traffic dynamics. *Japan Journal of Industrial and Applied Mathematics*, 11:203–223, 1994.

[5] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E*, 51:1035–1042, Feb 1995.

[6] Z. Bareket, P. S. Fancher, Huei Peng, Kangwon Lee, and C. A. Assaf. Methodology for assessing adaptive cruise control behavior. *IEEE Transactions on Intelligent Transportation Systems*, 4(3):123–131, 2003.

[7] R. Bellman and K. Åström. On structural identifiability. *Mathematical Biosciences*, 7(3):329 – 339, 1970.

[8] G. Bellu, M. Saccomani, S. Audoly, and L. D'Angiò. Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine*, 88:52–61, 11 2007.

[9] O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, and J.-P. Steyer. Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology and Bioengineering*, 75(4):424–438, 2001.

[10] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom. nuScenes: A multimodal dataset for autonomous driving. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11618–11628, 2020.

[11] M. J. Chappell and K. R. Godfrey. Structural identifiability of the parameters of a nonlinear batch reactor model. *Mathematical Biosciences*, 108(2):241 – 251, 1992.

[12] X. Chen and L. Shi. *Stochastic evolutions of dynamic traffic flow: Modeling and applications.* Springer-Verlag Berlin Heidelberg, 2015.

[13] B. Ciuffo and V. Punzo. "No Free Lunch" theorems applied to the calibration of traffic simulation models. *IEEE Transactions on Intelligent Transportation Systems*, 15(2):553–562, 2014.

[14] G. C. Ferro and V. Gerdt. Improved kolchin–ritt algorithm. *Programming and Computer Software*, 29:83–87, 2004.

[15] M. Garavello, K. Han, and B. Piccoli. *Models for Vehicular Traffic on Networks.* American Institute of Mathematical Sciences, 2016.

[16] D. C. Gazis, R. Herman, and R. B. Potts. Car-following theory of steady-state traffic flow. *Operations Research*, 7(4):499–505, 1959.

[17] D. C. Gazis, R. Herman, and R. W. Rothery. Nonlinear follow-the-leader models of traffic flow. *Operations Research*, 9(4):545–567, 1961.

[18] S. T. Glad. Solvability of differential algebraic equations and inequalities: An algorithm. In *1997 European Control Conference (ECC)*, pages 1195–1200, 1997.

[19] M. Grewal and K. Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 21(6):833–837, 1976.

[20] G. Gunter, R. Stern, and D. B. Work. Modeling adaptive cruise control vehicles from experimental data: model comparison. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 3049–3054, 2019.

[21] Z. He, L. Zheng, and W. Guan. A simple nonparametric car-following model driven by field data. *Transportation Research Part B: Methodological*, 80:185 – 201, 2015.

[22] R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[23] Y. Hollander and R. Liu. The principles of calibrating traffic microsimulation models. *Transportation*, 35(3):347–362, May 2008.

[24] H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap. Global identifiability of differential models. *Communications on Pure and Applied Mathematics*, 73(9):1831–1879, 2020.

[25] S. P. Hoogendoorn and R. Hoogendoorn. Generic calibration framework for joint estimation of car-following models by using microscopic data. *Transportation Research Record*, 2188(1):37–45, 2010.

[26] X. Jin, Y. Zhang, F. Wang, L. Li, D. Yao, Y. Su, and Z. Wei. Departure headways at signalized intersections: A log-normal distribution model approach. *Transportation Research Part C: Emerging Technologies*, 17(3):318 – 327, 2009.

[27] J. Karlsson, M. Anguelova, and M. Jirstrand. An efficient method for structural identifiability analysis of large dynamic systems. *IFAC Proceedings Volumes*, 45(16):941 – 946, 2012. 16th IFAC Symposium on System Identification.

[28] A. Kesting and M. Treiber. Calibrating car-following models by using trajectory data: Methodological study. *Transportation Research Record*, 2088(1):148–156, 2008.

[29] A. Kesting, M. Treiber, and D. Helbing. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1928):4585–4605, 2010.

[30] P. K. Khosla and T. Kanade. Parameter identification of robot dynamics. In *1985 24th IEEE Conference on Decision and Control*, pages 1754–1760, 1985.

[31] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein. The highD dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2118–2125, 2018.

[32] C. Kreutz. An easy and efficient approach for testing identifiability. *Bioinformatics (Oxford, England)*, 34(11):1913—1921, June 2018.

[33] V. Kurtc and M. Treiber. Calibrating the local and platoon dynamics of car-following models on the reconstructed ngsim data. In V. L. Knoop and W. Daamen, editors, *Traffic and Granular Flow '15*, pages 515–522. Springer International Publishing, 2016.

[34] J. A. Laval, C. S. Toth, and Y. Zhou. A parsimonious model for the formation of oscillations in car-following models. *Transportation Research Part B: Methodological*, 70:228 – 238, 2014.

[35] L. Li and X. M. Chen. Vehicle headway modeling and its inferences in macroscopic/microscopic traffic flow theory: A survey. *Transportation Research Part C: Emerging Technologies*, 76:170 – 188, 2017.

[36] L. Li, X. M. Chen, and L. Zhang. A global optimization algorithm for trajectory data based car-following model calibration. *Transportation Research Part C: Emerging Technologies*, 68:311 – 332, 2016.

[37] L. Li, R. Jiang, Z. He, X. M. Chen, and X. Zhou. Trajectory data-based traffic flow studies: A revisit. *Transportation Research Part C: Emerging Technologies*, 114:225 – 240, 2020.

[38] C.-Y. Liang and H. Peng. Optimal adaptive cruise control with guaranteed string stability. *Vehicle System Dynamics*, 32(4-5):313–330, 1999.

[39] L. Ljung. *System Identification: Theory for the User*. Prentice Hall information and system sciences series. Prentice Hall PTR, 1999.

[40] L. Ljung and T. Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265 – 276, 1994.

[41] T. Ma and B. Abdulhai. Genetic algorithm-based optimization approach and generic tool for calibrating traffic microscopic simulation parameters. *Transportation Research Record*, 1800(1):6–15, 2002.

[42] T. Maiwald, H. Hass, B. Steiert, J. Vanlier, R. Engesser, A. Raue, F. Kipkeew, H. H. Bock, D. Kaschek, C. Kreutz, and J. Timmer. Driving the model to its limit: Profile likelihood based model reduction. *PLOS ONE*, 11(9):1–18, 09 2016.

[43] Matlab global optimization toolbox, R2020a. The MathWorks, Natick, MA, USA.

[44] N. Meshkat, C. E.-z. Kuo, and J. DiStefano, III. On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and combos: A novel web implementation. *PLOS ONE*, 9(10):1–14, 10 2014.

[45] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.

[46] H. Miao, X. Xia, A. S. Perelson, and H. Wu. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Review*, 53(1):3–39, 2011.

[47] V. Milanés and S. E. Shladover. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies*, 48:285–300, 2014.

[48] V. Milanés, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on Intelligent Transportation Systems*, 15(1):296–305, 2014.

[49] J. Monteil and M. Bouroche. Robust parameter estimation of car-following models considering practical non-identifiability. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 581–588, Nov 2016.

[50] J. Monteil, M. Bouroche, and D. J. Leith. $\mathcal{L}_2$ and $\mathcal{L}_\infty$ stability analysis of heterogeneous traffic with application to parameter optimization for the control of automated vehicles. *IEEE Transactions on Control Systems Technology*, 27(3):934–949, 2018.

[51] NGSIM. The next generation simulation program. `http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm`, 2006. U.S. Department of Transportation Federal Highway Administation.

[52] S. Panwai and H. Dia. Comparative evaluation of microscopic car-following behavior. *IEEE Transactions on Intelligent Transportation Systems*, 6(3):314–325, 2005.

[53] V. Papathanasopoulou and C. Antoniou. Towards data-driven car-following models. *Transportation Research Part C: Emerging Technologies*, 55:496 – 509, 2015. Engineering and Applied Sciences Optimization (OPT-i) - Professor Matthew G. Karlaftis Memorial Issue.

[54] H. Pohjanpalo. System identifiability based on the power series expansion of the solution. *Mathematical Biosciences*, 41(1):21 – 33, 1978.

[55] V. Punzo, B. Ciuffo, and M. Montanino. Can results of car-following model calibration based on trajectory data be trusted? *Transportation Research Record*, 2315:11 – 24, 2012.

[56] 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 Transactions on Intelligent Transportation Systems*, 16(1):184–193, Feb 2015.

[57] V. Punzo and F. Simonelli. Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. *Transportation Research Record*, 1934(1):53–63, 2005.

[58] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 06 2009.

[59] J. F. Ritt. Differential algebra. *American Mathematical Society Colloquium Publications*, 33, 1950.

[60] M. P. Saccomani, S. Audoly, and L. D'Angiò. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica*, 39(4):619–632, 2003.

[61] J. D. Stigter and J. Molenaar. A fast algorithm to assess local structural identifiability. *Automatica*, 58:118 – 124, 2015.

[62] M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E*, 62:1805–1824, Aug 2000.

[63] M. Treiber and A. Kesting. *Calibration and Validation*, pages 303–338. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[64] M. Treiber and A. Kesting. Microscopic calibration and validation of car-following models – a systematic approach. *Procedia - Social and Behavioral Sciences*, 80:922 – 939, 2013. 20th International Symposium on Transportation and Traffic Theory (ISTTT 2013).

[65] N. Tuncer and T. T. Le. Structural and practical identifiability analysis of outbreak models. *Mathematical Biosciences*, 299:1 – 18, 2018.

[66] S. Vajda, K. R. Godfrey, and H. Rabitz. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical Biosciences*, 93(2):217 – 248, 1989.

[67] A. F. Villaverde. Observability and structural identifiability of nonlinear biological systems. *Complexity*, 2019:1–12, 2019.

[68] A. F. Villaverde and J. R. Banga. Structural properties of dynamic systems biology models: Identifiability, reachability, and initial conditions. *Processes*, 5(4):29, Jun 2017.

[69] A. F. Villaverde and A. Barreiro. Identifiability of large nonlinear biochemical networks. *Match*, 76(2), 2016.

[70] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou. Structural identifiability of dynamic systems biology models. *PLOS Computational Biology*, 12(10):1–22, 10 2016.

[71] A. F. Villaverde, N. D. Evans, M. J. Chappell, and J. R. Banga. Input-dependent structural identifiability of nonlinear systems. *IEEE Control Systems Letters*, 3(2):272–277, 2019.

[72] A. F. Villaverde, N. Tsiantis, and J. R. Banga. Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models. *Journal of The Royal Society Interface*, 16(156):20190043, 2019.

[73] E. Walter, I. Braems, L. Jaulin, and M. Kieffer. Guaranteed numerical computation as an alternative to computer algebra for testing models for identifiability. In R. Alt, A. Frommer, R. B. Kearfott, and W. Luther, editors, *Numerical Software with Result Verification*, pages 124–131, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[74] E. Walter and Y. Lecourtier. Unidentifiable compartmental models: what to do? *Mathematical Biosciences*, 56(1):1 – 25, 1981.

[75] E. Walter and Y. Lecourtier. Global approaches to identifiability testing for linear and nonlinear state space models. *Mathematics and Computers in Simulation*, 24(6):472–482, 1982.

[76] E. Walter, J. Norton, and L. Pronzato. *Identification of Parametric Models: From Experimental Data*. Communications and Control Engineering. Springer, 1997.

[77] H. Wang, W. Wang, J. Chen, and M. Jing. Using trajectory data to analyze intradriver heterogeneity in car-following. *Transportation Research Record*, 2188(1):85–95, 2010.

[78] Y. Wang, G. Gunter, M. Nice, M. L. D. Monache, and D. B. Work. Online parameter estimation methods for adaptive cruise control systems. *IEEE Transactions on Intelligent Vehicles*, 2020.

[79] X. Xia and C. H. Moog. Identifiability of nonlinear systems with application to hiv/aids models. *IEEE Transactions on Automatic Control*, 48(2):330–336, 2003.