Data-driven Analysis for Disturbance Amplification in Car-following Behavior of Automated Vehicles

Yang Zhou
Assistant Professor
Zachry Department of Civil and Environmental Engineering
Texas A&M
199 Spence Street, Room 301A
College Station, TX, 53706
Email: yangzhou295@tamu.edu

Xinzhi Zhong
Graduate Student
Department of Civil and Environmental Engineering
University of Wisconsin, Madison
1415 Engineering Drive
Madison, WI 53706
Email: xzhong54@wisc.edu

Qian Chen
Ph.D. Candidate
School of Automation, Southeast University
Department of Civil and Environmental Engineering, University of Wisconsin, Madison
Madison, WI 53706
Email: qianchen2010@seu.edu.cn

Soyoung Ahn (Corresponding Author)
Professor
Department of Civil and Environmental Engineering
University of Wisconsin, Madison
1415 Engineering Drive
Madison, WI 53706
Email: sue.ahn@wisc.edu; Phone: (608) 265-9067

Jiwan Jiang
Graduate Student
Department of Civil and Environmental Engineering
University of Wisconsin, Madison
1415 Engineering Drive
Madison, WI 53706
Email: jjiang295@wisc.edu

Ghazaleh Jafarsalehi
Graduate Student
Department of Civil and Environmental Engineering
University of Wisconsin, Madison
1415 Engineering Drive
Madison, WI 53706
Email: jafarsalehi@wisc.edu

Abstract

This paper presents a data-driven framework to quantitatively analyze the disturbance amplification behavior of automated vehicles in car-following (CF). The data-driven framework can be applied to unknown CF controllers based on the concept of empirical frequency response function (FRF). Specifically, a well-known signal processing method, Welch's method, together with a short time Fourier transformation is developed to extract the empirical transfer functions from vehicle trajectories. The method is first developed assuming a generic linear controller with time-*invariant* CF control features (e.g., control gains) and later extended to capture time-*variant* features. The proposed methods are evaluated for estimation consistencies via synthetic data-based simulations. The evaluation includes the performances of the linear approximation accuracy for a linear time-invariant controller, a nonlinear controller, and a linear time-variant controller. Results indicate that our framework can provide reasonably consistent results as theoretical ones in terms of disturbance amplification. Further it can perform better than a linear theoretical analysis of disturbance amplification, particularly when nonlinearity in CF behavior is present. The methods are applied to existing field data collected from vehicles with adaptive cruise control (ACC) on the market. Findings reveal that all tested vehicles tend to amplify disturbances, particularly in low frequency (< 0.5 Hz). Further, the results demonstrate that these ACC vehicles exhibit time-varying features in terms of disturbance amplification ratio depending on the leading vehicle trajectories.

1. Introduction

Adaptive cruise control (ACC) is one of the most well-known automation functions that can achieve Society of Automotive Engineers (SAE) Level 1 and 2 automation. Numerous studies suggest that ACC can potentially improve traffic safety, traffic throughput, and energy efficiency (e.g., Hu et al., 2016; Milanés & Shladover, 2014). Due to its potential, ACC has been widely studied in the last few decades. Based on the approach, ACC can be largely divided into three groups: (1) analytical form linear/nonlinear controllers, (2) hard-constrained optimal control based controllers, and (3) deep reinforcement learning based controllers. Analytical linear/nonlinear controllers are usually represented in a feedback fashion to regulate the gap and speed difference with the preceding vehicle. They can be further categorized into linear feedback controllers (Arem et al., 2006; Bian et al., 2019; Li et al., 2018; Zhou and Ahn, 2019; Zhou et al., 2020; Gunter et al., 2021), proportional derivative controllers (Gong et al., 2019; Milanés et al., 2014; Wang et al., 2019), proportional integrated derivative controllers (Wu et al., 2016), and other types of nonlinear controllers (Talebpour and Mahmassani, 2016; Jin and Orosz, 2018; Qin and Orosz, 2017). Due to the simplicity in their analytical form, the analytical linear/nonlinear controllers have received great attention for both theoretical analysis and experimental validation. However, these controllers lack explicit objective functions and physical constraints such as acceleration/deceleration limits.

In contrast, optimal control based controllers, which rely on optimization techniques, can achieve multi-objective control and directly incorporate physical constraints. Furthermore, it can be implemented in a rolling horizon fashion, as in model predictive control (MPC), to handle time-varying disturbances (Gong and Du, 2018; Gong et al., 2016; Wang et al., 2014; Zhou et al., 2017; Zhou et al., 2019a). Nevertheless, the computation time of MPC depends on the complexity of the objective function and constraints (e.g., nonlinearities), which can hinder real-time implementations.

To overcome these limitations, some studies proposed deep learning based controllers. This type of controller, particularly the ones based on reinforcement learning, can easily handle different forms of objective function and constraints through reward functions (Cheng et al., 2019; Qu et al., 2020; Shi et al., 2021). Further, these controllers can be computationally efficient, achieved by an offline training process, which makes it suitable for real-time implementations. However, the performance depends greatly on the accuracy and scenario coverage of the training dataset. A more systematic review of ACC controllers can be found in Zhou et al., (2017) and Zhou et al., (2019a). In summary, the aforementioned approaches have distinct advantages and disadvantages, and the best choice of controllers remains an open question.

Regardless of controller type, disturbance amplification through vehicle string is an important and desirable property. Particularly, traffic oscillations, characterized by a recurring pattern of deceleration and acceleration in congested traffic, can amplify in traffic streams, resulting in traffic inefficiency (Chen et al., 2014; Zheng et al., 2011) and potential safety issues (Zheng et al., 2010).

Due to the importance, researches of autonomous vehicle longitudinal control The disturbance amplification can be described by a mathematically derived transfer function/frequency response function (FRF) of the leading vehicle speed and the following vehicle speed in the frequency domain. By interpreting a norm of the transfer function, we can clearly understand the degree of disturbance amplification over each frequency. Numerous studies (Naus & Ploeg, 2010; Ploeg et al., 2013; Swaroop, 1996; Swaroop et al., 1994; Wilson & Ward, 2011) have investigated the 'string stability' of an automated vehicle platoon by requiring the infinity norm of the transfer function to be less than or equal to one to ensure the non-amplification of disturbances. Though elegant, this analytical approach requires exact knowledge of the car-following (CF) model considered. Based on that, the transfer function can be theoretically derived based on the linearization of the CF model and a Laplacian (or Fourier) transformation (Bian et al., 2019b; Wang et al., 2018; Zhou et al., 2020).

The current literature lacks an approach to derive a norm of the transfer function to examine disturbance amplification if the CF controller is unknown or is not in closed form. This is a major disadvantage because well-known controllers, such as constrained optimal controllers and deep learning based controllers, do not have closed-form formulations. Furthermore, for partially automated vehicles (AVs) with ACC (ACC vehicles hereon) on the market, the control algorithm is proprietary and thus unknown to the public. In recent years, ACC vehicles have become more available for field testing, which made it possible to evaluate their disturbance amplification behavior (Basselink et al., 2017; Naus et al., 2010; Wu et al., 2017; Zhao et al., 2020). However, these evaluations were largely model-based given a customized closed-form linear controller (Milanés and Shladover, 2014; Naus et al., 2010; Stern et al., 2018), or by fitting a linear controller or a CF model for human-driven vehicles (e.g., Intelligent Drivers Model (Kesting et al., 2010) and Optimal Velocity Model (Bando et al., 1998)) for an unknown controller using field data (Gunter et al, 2020). The latter approach may lead to model mismatch that can render an inaccurate analysis of disturbance amplification, especially after linearization. Further, calibration error resulting from model mismatch can further induce inaccuracy in disturbance amplification analysis.

To remedy the aforementioned problems, this paper presents a data-driven disturbance amplification analysis to systematically evaluate a wide variety of CF controllers, including unknown controllers. Rather than assuming a certain controller type, we extract the frequency domain characteristics of the speed and acceleration of CF vehicles. Specifically, by treating the speed and acceleration as compounding signals consisting of multiple sinusoidal waves of different frequencies, we adopt Welch's method to robustly estimate the empirical transfer function based on a linear time invariant CF law assumption. Based on the empirical transfer function, we can readily calculate the norm of transfer function, which can be directly used to evaluate the disturbance amplification ratio. Considering noisy measurements and stochasticity that may be introduced by the estimation method itself, we further propose a stochastic framework to evaluate the disturbance amplification in a probabilistic fashion. Further, considering the potential time-variant behaviors of CF law, we also extend our framework

integrating with short term Fourier transformation, to enable disturbance amplification analysis in a joint temporal-frequency domain, based on the linear-variant CF law assumption.

This paper is organized as follows. Section 2 provides the general philosophy of our approach based on the concept of empirical transfer function with a linear time-invariant system assumption. Section 3 provides a numerical algorithm to estimate the empirical transfer function under noisy measurement in a stochastic and time-variant fashion. Section 4 conducts multiple synthetic experiments to test the consistencies of our framework with theoretical derivations. Section 6 applies our framework for field-collected commercial ACC data, and Section 6 provides concluding remarks and future research directions.

2. Disturbance Amplification Analysis based on a Linear Time-invariant Behavior Approximation

In this section, we begin with a well-known model-based transfer function for a linear CF controller to illustrate how the transfer function is used to describe disturbance amplification for CF behaviors, which can be further used for string stability analysis. Based on this concept, we establish a data-driven approach to analyze disturbance amplification.

2.1 Model-based FRF

A model-based FRF, also known as a transfer function, has been widely applied in control theory, aiming to theoretically derive the input-output relationship of different frequencies. For CF analysis, the transfer function between the leading and the following vehicle speeds has been widely applied to evaluate string stability; i.e., how a disturbance evolves through each CF pair. This is done based on the exact knowledge of the CF control model (Wang, 2018; Zhou and Ahn, 2019b) or by assuming a CF control form and calibrating the parameters (Gunter et al., 2020; Stern et al., 2018). Based on the control law, linearization together with Laplacian transformation is usually applied to derive the theoretical transfer function, $G_i(j\omega)$, between leading vehicle's speed $v_{i-1}(j\omega)$ and following vehicle's speed $v_i(j\omega)$, $\forall i$, in the frequency domain, as given in Eq.(1).

$$G_i(j\omega) = \frac{v_i(\mathrm{s})}{v_{i-1}(\mathrm{s})} \tag{1}$$

where j is an imaginary unit, and ω is the frequency. Details of the derivation can be found in several studies (Naus & Ploeg, 2010; Ploeg et al., 2013; Swaroop, 1996; Wang, 2018).

$G_i(j\omega)$ delivers rich information for analyzing disturbance amplification. Specifically, $G_i(j\omega) = |G_i(j\omega)| \measuredangle G_i(j\omega)$, where $|G_i(j\omega)|$ is the norm of $G_i(j\omega)$, denoting the disturbance amplification ratio over each frequency ω, and $\measuredangle G_i(j\omega)$ denotes the phase shift. Thus, $|G_i(j\omega)|$ can be used to evaluate string stability: the CF pair is string stable iff $\|G_i(j\omega)\|_\infty = sup_\omega |G_i(j\omega)| \leq 1$ (Naus et al., 2012, Swaroop et al., 1996, Zhou et al., 2020), and we can conclude that the disturbance is not amplified under all frequency. Though string stability has been widely investigated by different criteria such as L_∞ norm and L_P norm (Shuo et al., 2019),

H_∞ norm (Jin & Oroz, 2014, Vittoria et al., 2020) is adopted in our paper considering its wide application and direct interpretation of disturbance amplification over different frequencies.

The model-based analysis shown above has the following limitations: 1) The CF control law must have a closed form; 2) A vehicle dynamics equation and its parameters need to be exactly known; 3) Even if the CF control law has a closed-form, it must be differentiable and linearizable. These precludes the application of the model-based analysis for certain CF control approaches including model predictive control and reinforcement learning approaches. Further, for ACC vehicles on the market, control laws are usually unknown, and assuming a closed-form control law risks introducing significant errors to the frequency domain behavior and thus an inaccurate string stability evaluation. To remedy these limitations, we propose a data-driven method to analyze disturbance amplification by extracting the frequency domain behavior of ACC CF controllers. Details follow.

2.2 Data-driven FRF

This subsection describes the principle, in which the frequency domain characteristics of CF behavior are extracted from measured data. Specifically, rather than deriving $G_i(j\omega)$, we aim to approximate $G_i(j\omega)$ from data. If we measure noise-free speeds for a CF pair with a high sampling frequency (approaching infinity) for a long period (approaching infinity), we can estimate the theoretical $G_i(j\omega)$ by a data-driven transfer function based on a best linear time-invariant approximation, $\hat{G}_i(j\omega)$, as follows:

$$\hat{G}_i(j\omega) = \frac{\hat{v}_i(j\omega)}{\hat{v}_{i-1}(j\omega)} \tag{2}$$

where $\hat{v}_i(j\omega)$ denotes the Fourier transformation of the measured speed for vehicle i, and $\hat{G}_i(j\omega)$ is the estimated transfer function (usually known as FRF) using measurements. However, measurements usually contain noise as shown in Fig. 1, and the sampling frequency can also be limited.

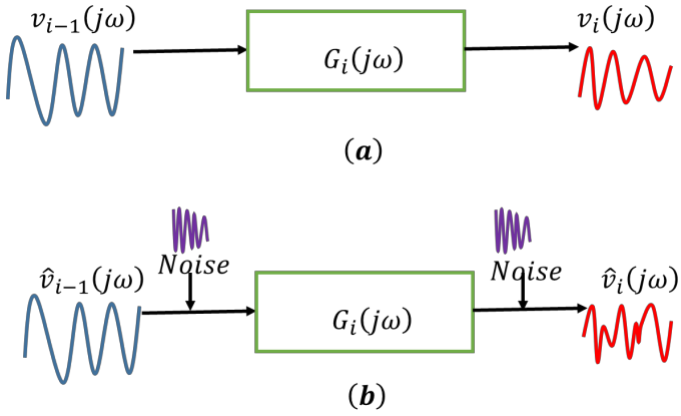

Fig. 1 FRF representation (a) theoretical case; (b) noisy case.

To address the issue of measurement noise, a spectral analysis technique (Randall, 2008) is adopted to estimate an empirical transfer function more robustly. Specifically, $G_{i+1}(j\omega)$ can be asymptotically estimated by:

$$\hat{G}_i(j\omega) = \frac{R_{i-1,i}(j\omega)}{R_{i-1,i-1}(j\omega)} \tag{3}$$

where $R_{i,i+1}(j\omega)$ is the cross spectral density function describing the cross 'energy' distribution over each frequency component:

$$R_{i-1,i}(j\omega) = \int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty} \hat{v}_{i-1}(\tau) \times \hat{v}_i(t+\tau)\, d\tau]e^{-j2\pi\omega t}dt \tag{4}$$

$R_{i-1,i-1}(j\omega)$ is the auto-spectral density function for vehicle $i-1$, describing the self 'energy' distribution over each frequency component:

$$R_{i-1,i-1}(j\omega) = \int_{-\infty}^{+\infty}[\int_{-\infty}^{+\infty} \hat{v}_i(\tau) \times \hat{v}_i(t+\tau)\, d\tau]e^{-j2\pi\omega t}dt \tag{5}$$

Based on $\hat{G}_i(j\omega)$, we can readily evaluate the disturbance amplification behavior without assuming a CF control law. Though the linearity and time-invariance property of CF without assuming details of CF laws and the corresponding parameter values. This approach can be further treated as a data-driven linear time-invariant approximation for the transfer function by simply utilizing vehicle speed.

3. Numerical Estimation Algorithm and Time-variant Extension

This section describes how to estimate the auto/cross spectral density, and the corresponding FRF in an unbiased fashion using discrete data with noise. We further extend the framework to a stochastic treatment of the FRF by basic statistics, considering the estimation error caused by algorithm itself and noisy measurements to facilitate a robust evaluation of disturbance amplification. Based on that, we extend the method to a time-variant fashion to better capture the time-variant behavior of ACC control.

3.1 Spectrum numerical estimation algorithm and data-driven string stability

Estimating the auto-spectrum and cross-spectrum requires some caution, as directly applying discrete Fourier transformation (DFT) is not consistent (Schoukens and Godfrey, 2018). The main reason is that when the number of measurements increases, DFT will only increase the resolution of spectrum (sampling frequency) but fail to 'average' the spectrum over the increased measurements. (Note that the frequency upper bound is determined by the sampling interval, and the frequency resolution is determined by the number of measurements. Hence, we apply a method that can properly average the estimations of auto-spectrum and cross-spectrum. Specifically, we apply the Welch method (Welch, 1967), a powerful tool to estimate auto-spectrum and cross spectrum, as illustrated in Fig. 2. The main idea of the Welch method is to decompose the signal into multiple overlapped segments, a windowed DFT is applied to extract the spectrum and phase angle in each segment given potential measurement noise. Finally, we average the spectra and phase angles to reduce the variance of the estimate.

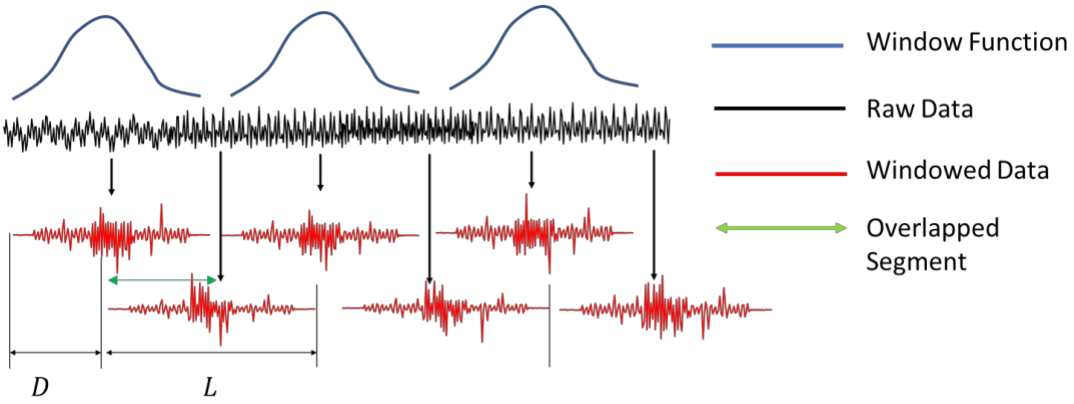

Fig. 2 Illustration of the Welch method

The detailed steps of the Welch method are given below:

Step 1: Input the (discrete) speed measurements of vehicle i for set l (e.g., from the l^{th} field test run) denoted by $\hat{v}_{l,i}[j]$, where $j = 0, \cdots, N_l - 1$, and N_l is the total number of measurements for set l.

Step 2: For each set l, we divide the data sequence into K_l segments, where each segment includes L data points, and successive segments are offset by D points (see Fig. 2). Then, the k^{th} segment can be described as: $\hat{v}_{l,i,k}[n] = \hat{v}_{l,i}[n + kD], \forall n = 1, \dots, L, k = 1,2, \dots, K_l$. Notice that overlapping segments have $L - D$ points, and thus, $N_l = L + D(K_l - 1)$. The main intention for overlapping segments is to increase the number of segments and reduce the variance of spectrum analysis (Schoukens and Godfrey, 2018). As suggested by Welch (1967), D is usually selected from $L/2$ to $2L/3$ (by trial and error).

Step 3: For each segment $k = 1,2, \cdots K_l$, a windowed finite Fourier transform is conducted as $V_k[h] = \frac{1}{L}\sum_{m=0}^{L-1} v_k[m]w[m]\, e^{-j2kmh/L}$, where j is the imaginary unit; $w[m]$ is the window (weight) function; and $h = 0, \cdots, L/2$. Some typical windows are rectangular, Kaiser, Harm or Hanning, Hamming, Blackman, and Blackman-Harris. Specifically, the rectangular window has a clear peak in the graph showing the power spectrum estimation. The spectral resolution of the rectangular window and Kaiser window is higher, but the noise level near the frequency of the analyzed signal is also higher. It is considered that both rectangular window and Kaiser window are suitable for high precision spectrum estimation of signals with a high signal-to-noise ratio. Hanning window is more effective for spectrum leakage suppression, but the frequency resolution is relatively low. Thus, it is suitable for the general frequency estimation of signals with a low signal-to-noise ratio. Interested readers are referred to Çakrak and Loughlin (2001) and Eberhard (1973) for more details on these windows. As suggested by Ponn et al., (2019) the signal-to-noise ratio can be relatively low (e.g., < 0.1) based on current sensors instrumented on autonomous vehicles. Hence we adopt the Hanning window, whose detailed form is given below:

$$w[m] = sin^2(\frac{\pi m}{L}) \quad (6)$$

The main reason to employ a window function is two-fold: (1) to reduce the 'spectral leakage effect' caused by dividing the data sequence into multiple segments and (2) to increase the performance of finite Fourier transform. The spectral leakage effect is the smearing of power across a frequency spectrum, which occurs when the signal is not periodic in the sample interval. A well-designed window function should increase the frequency range of the main lobe (usually manifested by higher power) for precision while reducing the power of the sidelobe to suppress noisy measurement contamination.

Step 4: For each segment $k = 1,2,\cdots K_l$ in each set, compute the corresponding modified auto-spectrum value by Eq. (7):

$$I_k(\omega_h) = \frac{L}{U}|V_k[h]|^2, k = 1,2,\cdots K_l \quad (7)$$

where $\omega_h = \frac{h}{L} \times f_s, h = 0,\cdots, L/2$. f_s is the sampling frequency and $U = \frac{1}{L}\sum_{m=0}^{L-1} w^2(m)$ is the mean value of the squared window weights. According to Eq. (7), the resolution of the spectrum estimation is positively correlated with the segment length L.

Step 5: Average the modified spectrum for all segments within each set l to obtain the average auto-spectrum estimate:

$$\hat{R}_l(\omega_h) = \frac{1}{K_l}\sum_{k=1}^{K_l} I_k(\omega_h) \quad (8)$$

Further, we take the average of the estimated auto-spectrum for all filed experiment runs M:

$$E[\hat{R}_l(\omega_h)] = \frac{1}{M}\sum_{l=1}^{\mathrm{M}} \hat{R}_l(\omega_h) \quad (9)$$

It is worth noting that bias is due to the window associated with truncation and increases as the segment length decrease. For a given window and fixed data length, increasing the number of segments reduces the variance but increases the bias (Welch, 1967).

Similarly, we can approximate the average cross-spectrum within each set l through Steps 1 to 5, while replacing Eq. (7) by calculating the cross-spectrum for each segment as below:

$$J_k(\omega_h) = \frac{L}{U}|C_k[h]|^2, k = 1,2,\cdots K_l \quad (10)$$

Then, the cross-spectrum for all segments within each set l can be approximated as:

$$\hat{S}_l(\omega_h) = \frac{1}{K_l}\sum_{k=1}^{K_l} J_k(\omega_h) \quad (11)$$

And the general average of the cross-spectrum can be approximated as:

$$E[\hat{S}_l(\omega_h)] = \frac{1}{M}\sum_{l=1}^{\mathrm{M}} \hat{S}_l(\omega_h) \quad (12)$$

According to Eq. (9), we can now calculate the FRF of each set l and the general average FRF for

an experiment:

$$\hat{G}_i(\omega_h) = \frac{\hat{S}_l(\omega_h)}{\hat{R}_l(\omega_h)} \tag{13}$$

$$E[\hat{G}_i(\omega_h)] = \sum_{l=1}^{M} \frac{\hat{S}_l(\omega_h)}{\hat{R}_l(\omega_h)} \tag{14}$$

Though the above data-driven transfer function highly relies on the assumption that the system is linear and time-invariant, it can still be applied to nonlinear systems as a best linear and time-invariant approximation for practicality. To account for nonlinearities and potential time-variant behaviors, we take a stochastic approach by describing an empirical cumulative density function (ECDF) for the data-driven transfer function based on M experiment runs, as below:

$$\Pr\big(\hat{G}_i(\omega_h) \le \gamma\big) = \sum_{l=1}^{M} \mathbf{1}_{\hat{G}_l(\omega_h)\le\gamma}/M \tag{15}$$

where $\mathbf{1}_{G_l(\omega_h)\le\gamma}$ is the indicator function, for which the event $\hat{G}_l(\omega_h) \le \gamma$ happens ($\mathbf{1}_{\hat{G}_l(\omega_h)\le\gamma} = 1$) or not.

Based on that, we use the $\bar{F}(\gamma)$, the ECDF of the maximum value of $\hat{G}_l(\omega_h)$ over all frequency components ω_h (ECDFM) to characterize the H-infinity norm of the estimated transfer function in a probabilistic way. The detailed equation is defined as:

$$\bar{F}(\gamma) = \Pr\big(\cup_{\omega_h}\{\hat{G}_l(\omega_h) \le \gamma\}\big) = \Pr\,(max_{\omega_h}[\hat{G}_l(\omega_h)] \le \gamma) \tag{16}$$

Given the lack of evidence showing dependence among different frequencies, we assume independence of each frequency ω_h. Hence, we have the following relationship between $\bar{F}(\gamma)$ and $F_{\omega_h}(\gamma)$

$$\bar{F}(\gamma) = \prod_{\omega_h} F_{\omega_h}(\gamma) \tag{17}$$

Eq. (20) provides a tool to analyze the string stability in a stochastic data-driven fashion. By the definition of the theoretical string stability given in Section 2.1, a homogenous vehicular platoon of length N with the averaged FRF $\hat{\mathrm{G}}_{\mathrm{l}}(\omega_{\mathrm{h}})$ and its ECDFM $\bar{\mathrm{F}}(\gamma)$ is strong string stable with a probability greater than or equal to α, if $\bar{\mathrm{F}}(1) \ge \alpha$.

In many existing controllers, the transfer function $G_i(j\omega) \to 1$ when $\omega \to 0$, which may make $\mathrm{F}_{\omega_{\mathrm{h}}}(1)$ very sensitive to the measurement error and estimation error of the proposed technique. To address this possible sensitivity, we give the estimation of $\mathrm{F}_{\omega_{\mathrm{h}}}(\gamma)$ a small buffer β (e.g., 0.06) when $\gamma = 1$. With the buffered estimation of empirical cumulative distribution, denoted by $\mathrm{F}^{\beta}_{\omega_{\mathrm{h}}}(\gamma)$, it is less likely to draw the unstring stable conclusion due to those errors. The buffered ECDF over each frequency ω_{h} can be systematically defined as:

$$F^{\beta}_{\omega_h}(\gamma) = \sum_{l=1}^{M} \mathbf{1}_{\hat{G}_l(\omega_h)\le\gamma+\beta}\,/M \tag{18}$$

And we can define the buffered ECDFM as:

$$\bar{F}^{\beta}(\gamma) = \prod_{\omega_h} F^{\beta}_{\omega_h}(\gamma) \tag{19}$$

Based on that, we get the buffered string stability: a homogenous vehicular platoon is practically buffered string stable with a probability greater than or equal to α, if $\bar{F}^{\beta}(1) \geq \alpha$.

The time-invariant analysis above is more suited to describe general CF behavior with different leading vehicle trajectories by the statistical maneuver given by Eqs. (15)-(19). To capture more nuanced CF behavior that varies over time, a time-variant extension is given in Section 3.2.

3.2 Time-variant extension of FRF

Here we extend the framework above and assume that the CF controller is time-variant under different leading vehicle trajectories. To incorporate the time-variant behavior, we extend the FRF from the frequency domain, $G_i(j\omega)$, into a response function in the joint time-frequency domain, $G_i(t, j\omega)$. Analogous to the time-invariant approximation, $\hat{G}_i(j\omega)$ in Eq. (6), the time-variant extension of the theoretical FRF $G_i(t, j\omega)$ could be approximated as $\hat{G}_i(t, j\omega)$ in the time-frequency domain:

$$\hat{G}_i(t, j\omega) = \frac{\hat{v}_i(t, j\omega)}{\hat{v}_{i-1}(t, j\omega)} \tag{20}$$

Where $\hat{v}_i(\tau, j\omega)$ is the short-time Fourier transformation of measured speed for vehicle i, providing both the temporal and frequency resolutions and frequency resolution. Different from the time-invariant case, we apply the discrete Short Time Fourier Transformation (STFT) to estimate the auto spectrum in the joint time-frequency domain to capture the time-variant features of speed energy change:

$$\hat{v}_i(t, j\omega) = \sum_{-\infty}^{\infty} \hat{v}_i(t)\, w(t-\tau) e^{-j\omega t} \tag{21}$$

where $w(t-\tau)$ is the time window with the center at τ, as discussed in Section 3.1. Note that there is trade-off in selecting the STFT time window function (also named block length). A short block length would provide a finer time resolution but degrade the frequency resolution. In contrast, when the block length is larger, more frequency information will be averaged over the time interval. Generally, there is no optimal STFT window. To numerically approximate $\hat{v}_i(t, j\omega)$, we similarly apply the discrete STFT based Welch method while considering the time-variant feature. Details follow.

Similar to Step 2 in Section 3.1, for a l, we divide the speed data for vehicle i into K_l segments, where each segment includes L data points, and successive segments are offset by D (usually $\mathrm{D} = \frac{L}{2}$) points for robustness as mentioned in Section 3.1. For each segment, we conduct a DFT transformation for the measured speed as:

$$V_k[h] = \frac{1}{L} \sum_{m=0}^{L-1} v_k[m] w[m]\, e^{-j2kmh/L} \tag{22}$$

and the corresponding auto-spectrum as:

$$I_k(\omega_h) = \frac{L}{U} |V_k[h]|^2, k = 1,2, \cdots K_i \tag{23}$$

where $\omega_h = \frac{h}{L} \times f_s, h = 0, \cdots, L/2$. f_s is the sampling frequency and $U = \frac{1}{L}\sum_{m=0}^{L-1} w^2(m)$ is the mean value of the squared window weights as described in Section 3.1. Then, the norm of time variant transfer function can be approximated as:

$$|\hat{v}_i(t, j\omega)|^2 \approx |\tilde{v}_i(t, j\omega_h)|^2 = \begin{cases} I_k(\omega_h) & when\ 0 \le t \le \frac{L}{2f_s} \\ \frac{I_k(\omega_h) + I_k(\omega_h)}{2} & when\ \frac{Lk}{2f_s} \le t \le \frac{L(k+1)}{2f_s} \\ I_{K_i}(\omega_h) & when\ \frac{Lk}{2f_s} \le t \le \frac{L(k+1)}{2f_s} \end{cases} \quad (23)$$

Similarly, we can approximate auto-spectrum of vehicle $i-1$, and then approximate $\left|\hat{G}_i(t, j\omega)\right|$ by:

$$\left|\hat{G}_i(t, j\omega)\right| \approx \sqrt{\frac{|\tilde{v}_i(t, j\omega_h)|^2}{|\tilde{v}_{i-1}(t, j\omega_h)|^2}} \quad (24)$$

Similar to the time-invariant case, $\left|\hat{G}_i(t, j\omega)\right|$ denotes the disturbance amplification ratio at time t and is thus string stable if $\sup_{t,\omega}\left|\hat{G}_i(t, j\omega)\right| \le 1$. Note that $\left|\hat{G}_i(t, j\omega)\right|$ is directly related to t, which means that the analysis fits for analyzing the time-varying CF behavior under non-steady state leading vehicle trajectories (e.g., with speed change, etc.).

The analysis above suggests a trade-off in the temporal and frequency domain resolutions in relation to L: a larger value of L increases the frequency domain resolution while sacrificing the temporal domain resolution. There is no optimal setting for L, and it needs to be determined based on the analysis resolution. Similar to the window function selection in Step 3 in Section 3.1, we choose the Hanning window and add the 50% overlap between adjacent blocks to make the STFT smoother, satisfy the completeness condition, and avoid the spectral leakage (Wexler and Raz 1990, Avargel and Cohen 2007). Note that discrete wavelet transformation is an alternative solution to a multi-resolution analysis in the time-frequency domain. However, it is sensitive to minute variations, especially for noisy speed signals (Faust et al. 2015).

For the above analysis, we select the speed variation $v_i(s)$ as a surrogate variable for stability since it is widely used in traffic flow analysis. For a homogenous vehicular platoon (i.e., same control law and control parameters), it can be proved that $\frac{v_i(s)}{v_{i-1}(s)} = \frac{a_i(s)}{a_{i-1}(s)} = \frac{d_i(s)}{d_{i-1}(s)} = \frac{\Delta v_i(s)}{\Delta v_{i-1}(s)}$ regardless of the control law, where $a_i(s)$, $d_i(s)$, $v_i(s)$, and $\Delta v_i(s)$ respectively represent the acceleration, spacing, speed, and relative speed for vehicle i in the frequency domain. For a heterogenous platoon, $\frac{v_i(s)}{v_{i-1}(s)} = \frac{a_i(s)}{a_{i-1}(s)} = G_i(s)$ and $\frac{d_i(s)}{d_{i-1}(s)} = \frac{\Delta v_i(s)}{\Delta v_{i-1}(s)} = F_i(s)$, but usually $F_i(s) \neq G_i(s)$. Thus, $F_i(s)$ and $G_i(s)$ need to be estimated separately by changing the surrogate variable for stability. Nevertheless, our framework can be still applied.

4. Evaluation of the Method via Numerical Simulation

In this section, we systematically validate the proposed analysis for data-driven stochastic string stability and estimation consistency of transfer function using numerical simulation. We generate synthetic data from a known linear controller to check whether the data-driven approach can provide a transfer function consistent with the mathematical derivation in both time-invariant and time-variant fashions. We also check whether the time-invariant data-driven FRF serves as a good approximation for a nonlinear controller. Since the describing function analysis can only be applied to controllers with certain types of nonlinearities, we select a linear controller with bounded acceleration and deceleration as a special case for the evaluation.

4.1 Synthetic test for a linear controller

We first generate leading vehicle trajectories with the equilibrium speed of 15 m/s, and to reproduce the oscillatory feature, we add compounding oscillatory speed as an excitation (a disturbance to trigger the frequency response) for the CF simulation. These features are selected based on the reconstructed NGSIM data (Montanino and Punzo, 2015; Punzo et al., 2011). Given the leading vehicle trajectories, we simulate the following vehicle trajectories based on an assumed controller. Particularly, we adopt the linear controller by Zhou et al. (2019b) as an illustrative example, whose transfer function can be mathematically derived as below:

$$G_i(\mathrm{s}) = \frac{(k_s + k_v s + k_f s^2)}{T_L s^3 - (k_a - 1)s^2 + (\tau_i^* k_s + k_v)s + k_s} \tag{25}$$

where k_s, k_v, k_a are feedback gains for the deviation from equilibrium spacing, speed difference, and acceleration. k_f is the feedforward gains. T_L and τ_i^* are the time engine coefficient and constant time gap for vehicle i. The corresponding transfer function norm can be obtained as in Eq. (26):

$$|G_i(j\omega)| = \frac{k_s^2 + k_v^2\omega^2 + k_f^2\omega^4 - 2k_s k_f \omega^2}{T_L^2\omega^6 + \left[-2T_L(k_s\tau_i^* + k_v) + (k_a - 1)^2\right]\omega^4 + \left[(k_s\tau_i^* + k_v)^2 + 2k_s(k_a - 1)\right]\omega^2 + k_s^2} \tag{26}$$

To simulate real-world sensor measurements, we add normally distributed white noise with the standard deviation of 0.1 m/s (NHTSA, 2013 March) for both leading and following vehicle trajectories. Using these synthetic data, we apply the Welch method to extract the transfer function in a probabilistic fashion and check the accuracy of the data-driven transfer function estimation. Considering the trade-off between the resolution and the number of segments to average the estimation, the parameter values for the Welch method are set as shown in Table 1. The detailed process of the numerical simulation analysis is given in Fig. 3.

Table 1. Default parameters for the Welch method

f_s	0.1 Hz
L	12 sec
M	1000
D	6 sec
N_l	6 sec

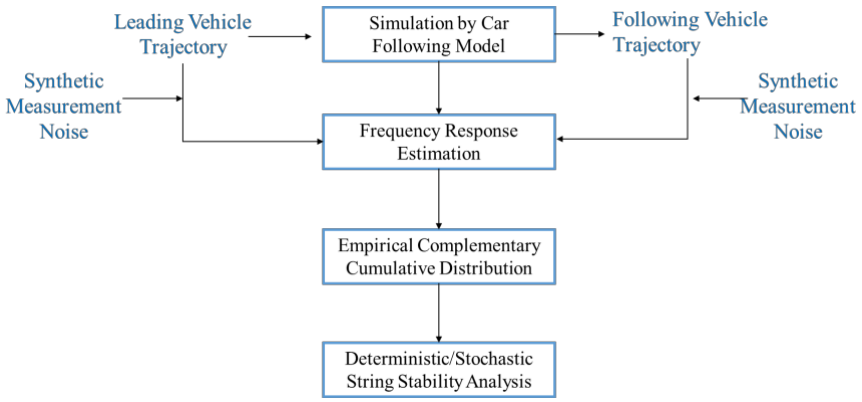

Fig. 3 Evaluation process based on synthetic data simulation

To test whether the proposed method can distinguish different string stability behaviors by different controller parameters, we design three scenarios: (1) string stable, (2) string unstable, and (3) slightly string unstable. The parameter settings and the corresponding theoretical infinity norms for the three scenarios are given in Table 2, and other parameters follow Zhou et al. (2020). For the three scenarios, we evaluate the accuracy of the estimated norm of transfer functions against the empirical two-sided 90 percent boundaries extracted from the synthetic data. We conducted 1000 simulation runs, with each run containing 3.5 minutes of trajectories for the proceeding and following vehicles. The discretization step of the simulation is set as 0.1 sec.

Table 2. Controller parameters and H_∞ norm for different scenarios

Scenario	Feedback gains $[k_s, k_v, k_a]$	H_∞ norm
1	[0.3, 1.5, -0.6]	1
2	[1.5, 0.3, -0.6]	1.11
3	[1,1,-1]	1.03

For the three scenarios, we evaluate the accuracy of the estimated norm of transfer functions against the empirical [5%, 95%] boundaries extracted from the synthetic data. We conducted 1000 simulation runs, in which each run contains 3.5 minutes of trajectories for the proceeding and following vehicles. The discretization step of the simulation is set as 0.1 sec. The estimated norm of the transfer function for each run under three scenarios is given in Fig. 4(a)-(c). Furthermore, Fig. 4(d)-(f) show the average norm and the empirical two-sided 90 percent boundaries, compared against the theoretical norm. They show that overall, the average norms based on the proposed method (red curves) are nearly superimposed with the theoretical norms (blue curves) in all three scenarios. This suggests that the proposed method can produce accurate and consistent results in various scenarios by averaging the estimations, which demonstrates the robustness of our proposed method. Note that though negligible, the proposed method is less accurate in lower frequency regions (e.g., < 0.06 Hz), and results in a 0.01 discrepancy with the theoretical value.

The main reason is that estimating an empirical transfer function over a low-frequency region requires a long period of measurements and thus a relatively large L in our framework. The number of low-frequency signals given limited measurements can be small. Further, a large L can significantly reduce the number of segments used to average the transfer functions, leading to potential inaccuracy. Another factor is the magnitude of measurement noise. Nonetheless, our results indicate that when the standard deviation of the speed measurement is less than 0.05m/s, our estimation can produce results close to the theoretical ones.

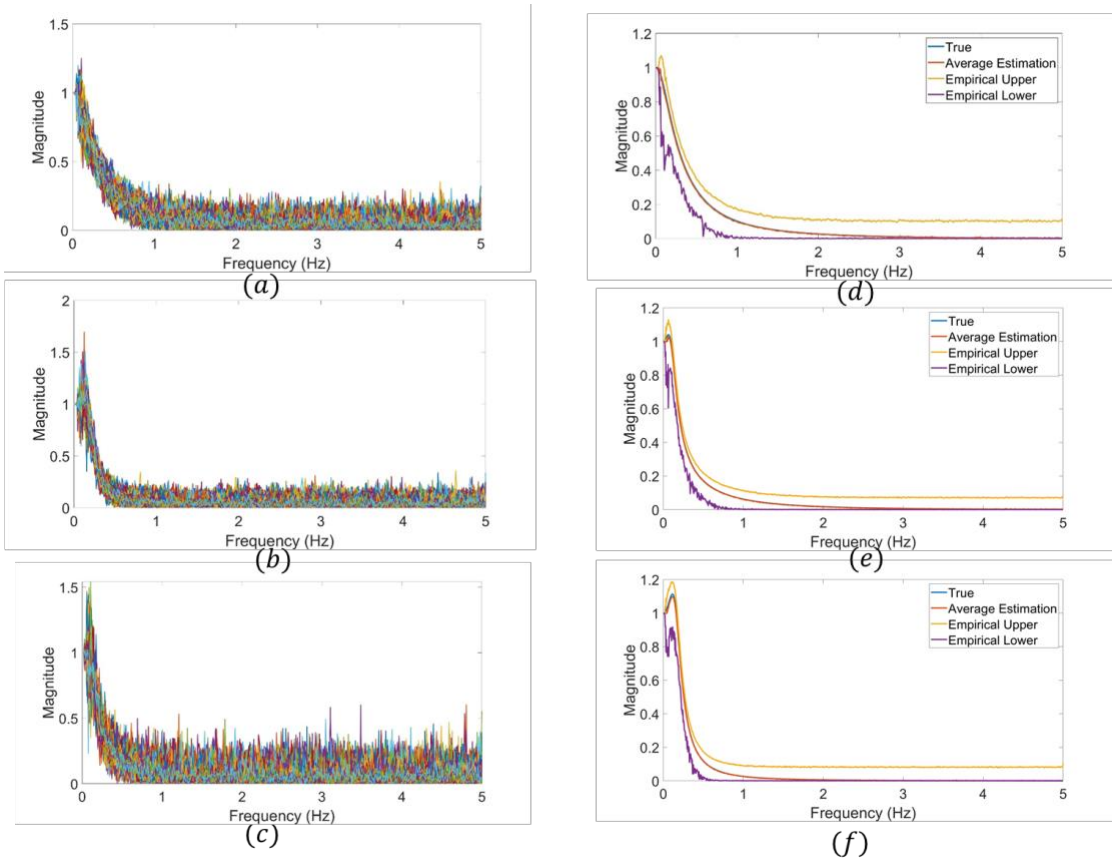

Fig. 4 Empirical norm of estimated transfer functions (a-c) all experimental empirical norm of transfer functions for Scenario 1 to 3; (d-f) norm of the average empirical transfer function and empirical boundaries

We further analyzed the empirical [5%, 95%] confidence interval given in Fig. 4(d)-(f). Specifically, we adopt Proposition 3 to check whether the proposed method can properly determine string stability and instability for the three scenarios designed above. The details are provided in Fig. 5. As in Section 3.1, we first analyze the buffered string instability probability for each frequency, ω_h, i.e., the probability that the norm of the estimated transfer function is greater than $1+\beta$ (with $\beta = 0.06$), $1 - F^{\beta}_{\omega_h}(1)$ in Eq. (19); see the blue bars in the figure. Then we compute the buffered complementary cumulative string stability probability $\bar{F}^{\beta}(\gamma)$; see the red curve in the figure. The buffered string instability probability (blue bars) mainly lies in the low-frequency part. Further,

scenario 1 (strong string stable case) has the least likelihood of string instability across different frequencies, scenario 2 (strong string unstable case) has the highest likelihood, and scenario 3 (slightly string stable case) has a likelihood in between. For the general string stability conclusion, the red curve shows that the complementary cumulative string stability probabilities for scenarios 1, 2, and 3 are respectively 76%, 0%, and 30%. The results demonstrate that the proposed method can accurately distinguish strong string stability and instability cases, and further identify which frequency domain the ACC might be string unstable. However, the proposed method is more likely to produce a higher probability of string instability in the low-frequency regions as expected based on Remark 1. The probability of strong string stability depends on the choice of hyper-parameter β. With proper tuning of the parameter, the proposed method can produce justifiable results.

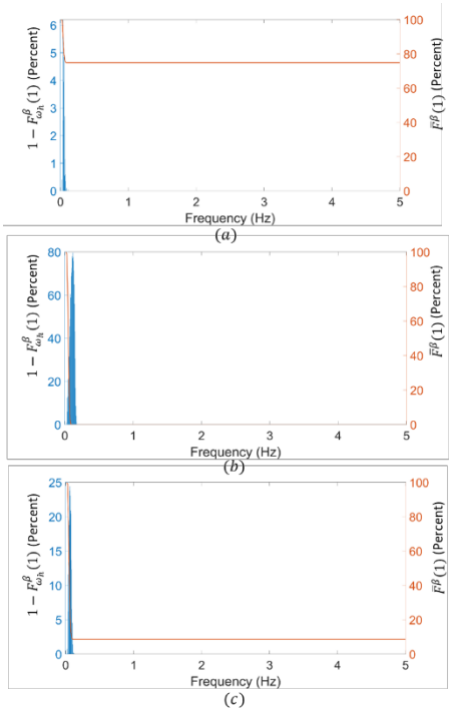

Fig. 5 $1-F^{\beta}_{\omega_h}(1)$ and $\bar{F}^{\beta}(1)$ (a) scenario 1; (b) scenario 2; (c) scenario 3

4.2 Synthetic test for a non-linear controller

Since deriving the theoretical transfer function of a nonlinear controller is extremely challenging, we can only evaluate our method in limited scope. The complexity depends on the nonlinearity form of CF law and if the CF law is even analytical. Inspired by Li et al. (2012), we apply a describing function analysis to provide a theoretical transfer function derivation for a simple nonlinear control law. Specifically, in this section, we analyze a nonlinear controller, which is piecewise linear, given as below:

$$u_i(t) = mid\,\{a_{min},\ k_i x_i(t),\ a_{max}\} \tag{27}$$

where a_{min} and a_{max} denote the deceleration and acceleration limits, respectively. Given the simplicity, the above controller can be theoretically derived by a describing function analysis

method if the leading vehicle acceleration portfolio is dominated by a frequency ω with a magnitude A_{i-1}, rendering $a_{i-1}(t) = A_{i-1}sin(\omega t)$. The core idea is to decompose the signal of the following vehicle trajectories by Fourier Transformation as:

$$a_i(t) = Y_{1,1}sin(\omega t) + Y_{1,2}cos(\omega t) + Y_{2,1}sin(2\omega t) + Y_{2,2}cos(2\omega t) + \cdots \quad (28)$$

By assuming the low-pass property of CF law (i.e., $Y_{1,1}$, $Y_{1,2} \gg Y_{k,1}$, $Y_{k,2}$, for $k \geq 2$), $a_i(t)$ can be approximated as:

$$a_i(t) \approx Y_{1,1}sin(\omega t) + Y_{1,2}cos(\omega t) \quad (29)$$

where

$$Y_{1,1} = \frac{1}{\pi}\int_{-\theta}^{2\pi-\theta} a_i(t)\sin(\omega t)\, d(\omega t) \quad (30)$$

$$Y_{1,2} = \frac{1}{\pi}\int_{-\theta}^{2\pi-\theta} a_i(t)\cos(\omega t)\, d(\omega t) \quad (31)$$

Then the relationship between $a_{i-1}(t)$ and $a_i(t)$ can be described by a describing function, $H_{nl}(j\omega)$, defined as multiplication of a norm function $|H_{nl}(j\omega)|$:

$$H_{nl}(j\omega) = |H_{nl}(j\omega)|\measuredangle H_{nl}(j\omega) \quad (32)$$

where,

$$|H_{nl}(j\omega)| = \frac{\sqrt{Y_{1,1}^2 + Y_{1,2}^2}}{A_{i-1}} \quad (33)$$

$$\measuredangle H_{nl}(j\omega) = arctan(\frac{Y_{1,2}}{Y_{1,1}}) \quad (34)$$

For simplicity, we set $a_{max} = -a_{min} = a_{bound}$. Hence, in an oscillation cycle $[-\theta, 2\pi - \theta]$, we can mathematically represent the following vehicle acceleration based on two conditions:

Condition 1 (Inactive boundary): when $\frac{a_{bound}}{A_{i-1}K_L(\omega)} \leq 1$

$$a_i(t) = |G_i(j\omega)|v_i\big(\omega t + \measuredangle G_i(j\omega)\big) \quad (35)$$

Condition 1 follows an exact linear control law that has already been evaluated.

Condition 2 (Active boundary): when $\frac{a_{bound}}{A_{i-1}|G_i(j\omega)|} > 1$

$$a_i(t) = \begin{cases} |G_i(j\omega)|v_i\big(\omega t + \measuredangle G_i(j\omega)\big)\ if -\beta - \measuredangle G_i(j\omega) < \omega t \leq \beta - \measuredangle G_i(j\omega) \\ a_{bound}\ if\ \beta - \measuredangle G_i(j\omega) < \omega t \leq \pi - \beta - \measuredangle G_i(j\omega) \\ |G_i(j\omega)|v_i\big(\omega t + \measuredangle G_i(j\omega)\big)\ if\ \pi - \beta - \measuredangle G_i(j\omega) < \omega t \leq \pi + \beta - \measuredangle G_i(j\omega) \\ -a_{bound}\ if\ \pi + \beta - \measuredangle G_i(j\omega) < \omega t \leq 2\pi - \beta - \measuredangle G_i(j\omega) \end{cases} \quad (36)$$

The details are given in Fig. 6.

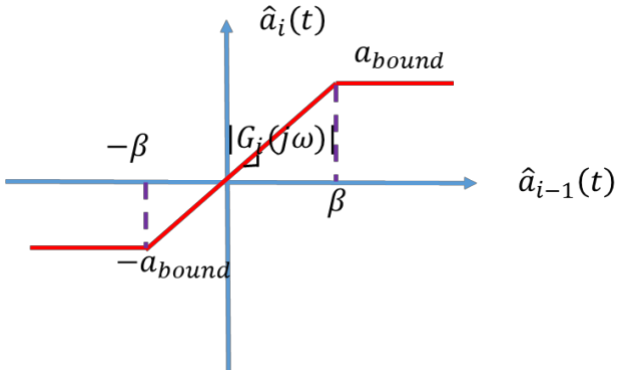

Fig. 6 $\hat{a}_{i-1}(t)$ vs. $\hat{a}_i(t)$ plane

By letting $\beta = \sin^{-1}\left(\frac{a_{bound}}{A_{i-1}|G_i(j\omega)|}\right)$, we have:

$$Y_{1,1} = \frac{1}{\pi}\int_{-\theta}^{2\pi-\theta} a_i(t)\sin(\omega t)\,d(\omega t) = \frac{4a_{max}\cos(\beta)\cos(\theta)+A_{i-1}|G_i(j\omega)|\cos(\theta)(2\beta-sin(2\beta))}{\pi} \quad (37)$$

$$Y_{1,2} = \frac{1}{\pi}\int_{-\theta}^{2\pi-\theta} a_i(t)\cos(\omega t)\,d(\omega t) = \frac{4a_{max}\cos(\beta)\sin(\theta)+A_{i-1}|G_i(j\omega)|\sin(\theta)(2\beta-sin(2\beta))}{\pi} \quad (38)$$

Further, based on (40), we have:

$$|H_{nl}(j\omega)| = \frac{\sqrt{Y_{1,1}^2+Y_{1,2}^2}}{A_{i-1}} = \frac{|4a_{bound}\cos(\measuredangle|G_i(j\omega)|)-A_{i-1}|G_I(j\omega)|\sin(2\beta)+2A_{i-1}|G_i(j\omega)\beta|}{A_{i-1}\pi} \quad (39)$$

$$\measuredangle H_{nl}(j\omega) = \measuredangle G_i(j\omega) \quad (40)$$

As notable by Eq. (40), the theoretical describing function analysis based on a FRF norm is apparently affected by the leading vehicle oscillation magnitude A_{i-1} and acceleration/deceleration boundaries a_{bound}, in contrast to the theoretical linear controller. As a comparison, we further analyze if the data-driven linear (time-invariant) FRF can provide a reasonable approximation for the theoretical non-linear FRF. As an example, we select a leading vehicle's acceleration oscillating at 5Hz with its magnitude varying from 6 to 8 m/s^2 to make the acceleration and deceleration boundaries active. As shown in Fig. 7, the data-driven linear FRF is very similar to the nonlinear theoretical approximation under different leading vehicle oscillation magnitudes. However, there is an apparent discrepancy between the linear theoretical one and nonlinear theoretical one, and the discrepancy increases as the magnitude of the leading vehicle oscillations grows. The example provided highlights the limitation of the linear theoretical string stability analysis and the potential of our proposed data-driven method for disturbance amplification analysis. Furthermore, ACC algorithms of vehicles on the market are usually unknown. Even if the CF algorithm is fully known and well calibrated, the linear theoretical string stability still renders unsatisfied results.

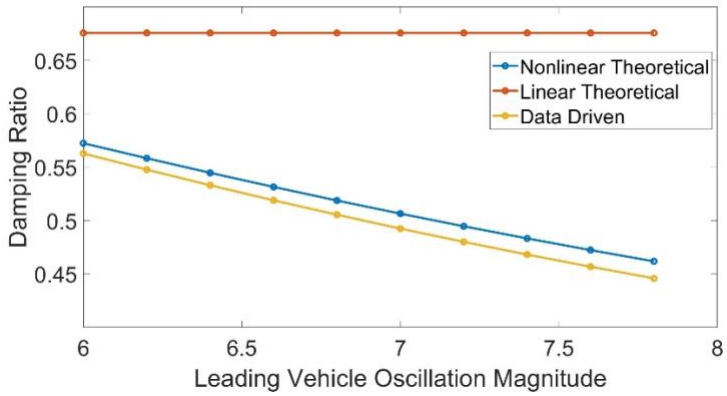

Fig. 7 Disturbance amplification ratio under different leading vehicle acceleration oscillation magnitude

4.3 Synthetic test for a linear time-varying controller

We further validate our time-variant method using synthetic data similar to the time-invariant case. We design the leading vehicle trajectory to travel at $15\ \mathrm{m/s}$ and oscillate at the frequency of 0.25 Hz and 0.5Hz, with the magnitude of 1.5 $\mathrm{m/s}$ and 1 $\mathrm{m/s}$. We apply the same linear control law with time-varying control gains to reflect the potential time-varying features. The detailed theoretical time-variant transfer function is given as:

$$G_i(\mathrm{t},\mathrm{s}) = \frac{(k_s(t)+k_v(t)s+k_f(t)s^2)}{T_{i,L}s^3-(k_a(t)-1)s^2+\left(\tau_i^* k_s(t)+k_v(t)\right)s+k_s(t)} \tag{41}$$

and

$$[k_s(t), k_v(t), k_a(t)] = \begin{cases} [0.3, 1.5, -0.6] & when\ t \le 90s \\ [1.5, 0.3, -0.6] & when\ t \ge 90s \end{cases} \tag{42}$$

Considering the trade-off between the time domain and frequency domain resolutions, we let $\mathrm{L} = 60$ s, assuming the CF behavior remains unchanged for 30 s. Following the method shown in Section 3.2, the auto-spectrum of leading vehicle speed, CF pair speed evolvement, norm of theoretical FRF and norm of empirical FRF are given in Fig. 8. The result given in Fig. 8 shows that when the control gains suddenly change at $\mathrm{t} = 90$ s, there is a discrepancy between theoretical and empirical FRF as anticipated, due to the nature of STFT assuming that the behavior remains unchanged over each L. Other than the discrepancy around the change, our method provides consistent estimations.

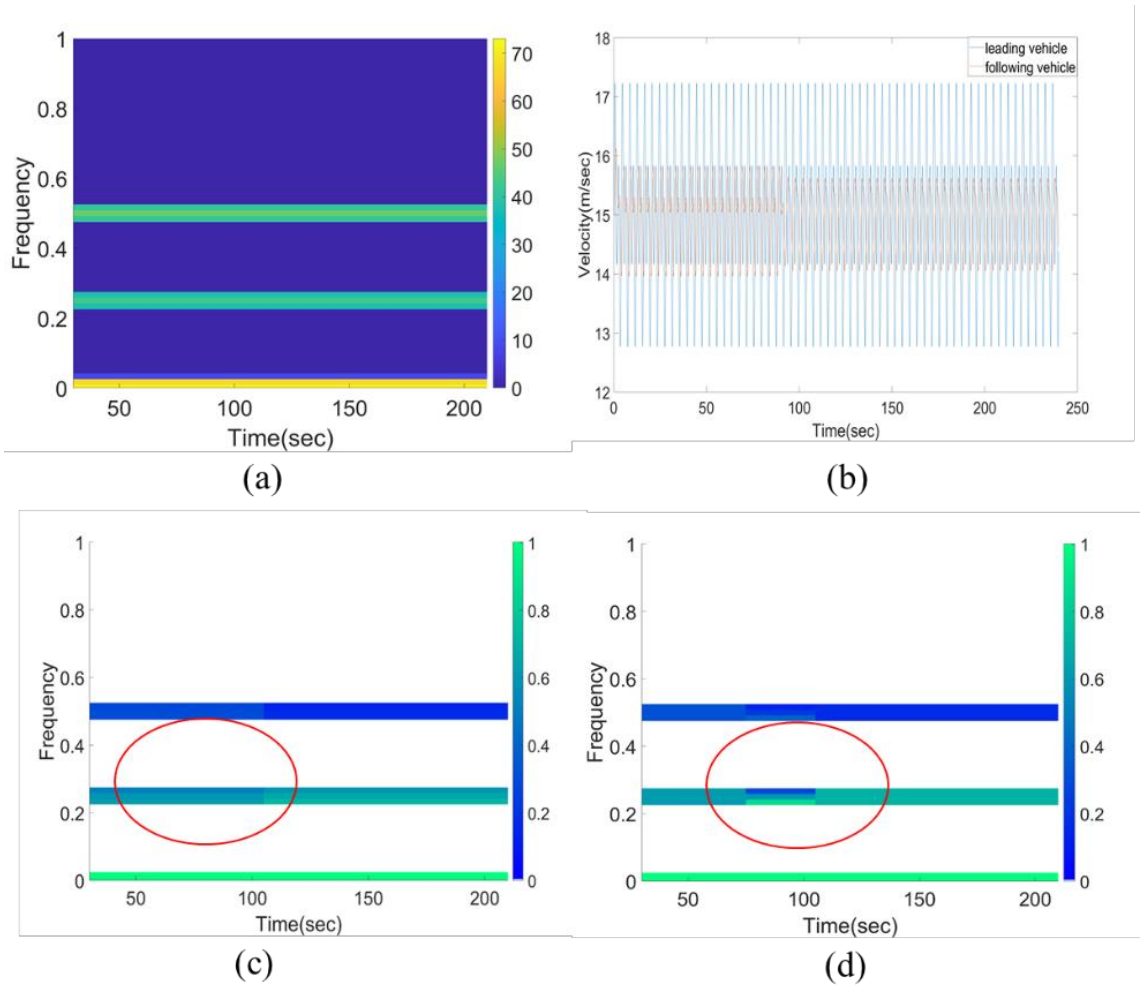

Fig. 8 (a) Auto-spectrum of leading vehicle speed; (b) speed portfolio; (c) Norm of Theoretical FRF (d) Norm of Empirical FRF

5. Analysis of commercial ACC Vehicles

In this section, we apply our proposed method to systematically quantify the disturbance dampening behaviors of ACC vehicles on the market. Considering the experiment duration and scenarios design, we apply the linear time-invariant FRF approach to describe the general behaviors using the OpenACC dataset (Makridis et al., 2020). Further, we also apply the linear time-variant FRF approach to describe the nuanced behavior concerning specific leading vehicle trajectories using the Massachusetts experiment dataset (Li et al., 2021).

5.1 Linear Time-Invariant FRF based Analysis

The OpenACC database includes trajectory and speed data of five ACC vehicles in three CF test campaigns with different driving paths, study durations, and platoon formations. We analyze the data from the third campaign with fewer deployment issues, which was conducted in 2019 on a 5.7-km rural road section of the AstaZero test track in a protected environment. The details of the experimental set-up are provided in (Makridis et al., 2020). The path layout is shown in Fig. 9. In

this campaign, a platoon of five vehicles (referred to as vehicle types 1-5 in this paper) was chosen as study objects to test the ACC capability during traffic oscillations. The vehicle position and speed data were measured through high-fidelity differential Global Navigation Satellite Systems (GNSSs) by recording the GPS coordinates (latitude, longitude, altitude), ENU coordinates, instantaneous inter-vehicle distances, and speeds of each vehicle at a certain frequency. The initial frequency provided by the data acquisition system is more than 100 Hz. However, the frequency has been reduced to 10Hz in the open-source data to reduce the noise caused by high-frequency sampling. A low-pass filter and linear interpolation were also applied to deal with invalid and missing data.

Fig. 9. Path layout of the third experiment in OpenACC dataset (Makridis et al., 2020)

Field testing has been conducted for the above-mentioned vehicle platoon under speed disturbances of different frequencies around the target speed. To estimate the transfer function over the entire potential frequency range, we first examine vehicle trajectories during traffic oscillations, which tend to exhibit wider frequency ranges. An example of vehicle trajectories and the corresponding speeds are shown in Fig. 10. All vehicles started from rest. The first leading vehicle accelerated from rest to the designed target speed of 60 mph at the average rate of 0.95 $\mathrm{m/s^2}$. After that, the first vehicle intentionally decelerated and accelerated around the target speed to create traffic oscillations of different magnitudes and frequencies. It is observed that the speed oscillations were amplified through the vehicle string, which intuitively indicates string instability.

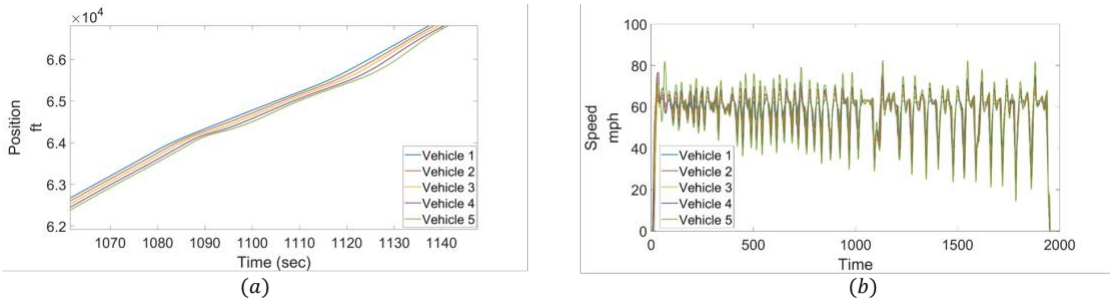

Fig. 10. Examples of (a) vehicles trajectories and (b) speeds

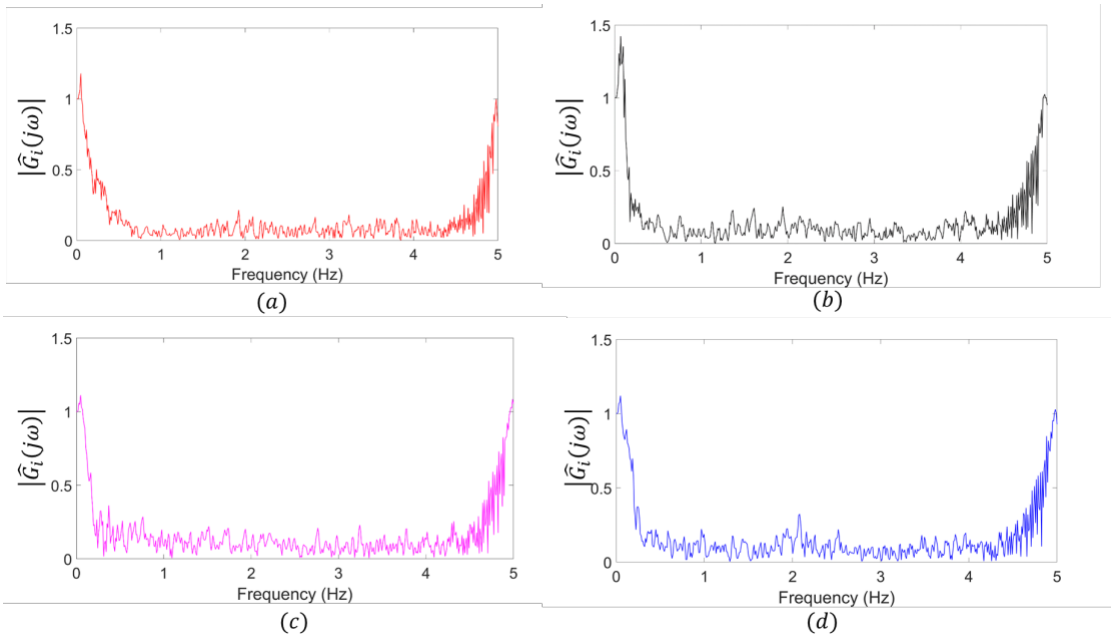

Fig. 11 Norm of average empirical transfer functions for ACC vehicles (a) vehicle 2; (b) vehicle 3; (c) vehicle 4; (d) vehicle 5

To further investigate the frequency-domain properties, we used trajectory and speed data over the duration of 2.6 hours to estimate the empirical transfer function and the corresponding empirical and practical cumulative distribution of practical string stability. The results in Fig. 11 show that the norm of the average empirical transfer function varies by vehicle model/manufacturer, suggesting varying ACC performance in terms of string stability. Furthermore, we find that all vehicle models exhibit large norm values in the low-frequency region. Contrary to the linear controllers tested in Section 4.1, the ACC vehicles on the market also exhibit large norm values in high-frequency regions (e.g., > 4Hz). We suspect that with a high-frequency disturbance, vehicle dynamics such as actuation delay may exacerbate the follower response. In comparison, vehicle 3 is most string unstable, especially in the low-frequency region below 0.5 Hz, where the norm reaches nearly 1.5. Vehicles 2, 4, and 5 are all string unstable, though vehicle 4 performs better with the lowest H_∞ norm value of transfer function and the smallest string unstable regions. To further evaluate the strong string stability in a probabilistic manner, we calculate the string stability probability for each frequency component and cumulative string stability probability. The results are given in Fig. 12.

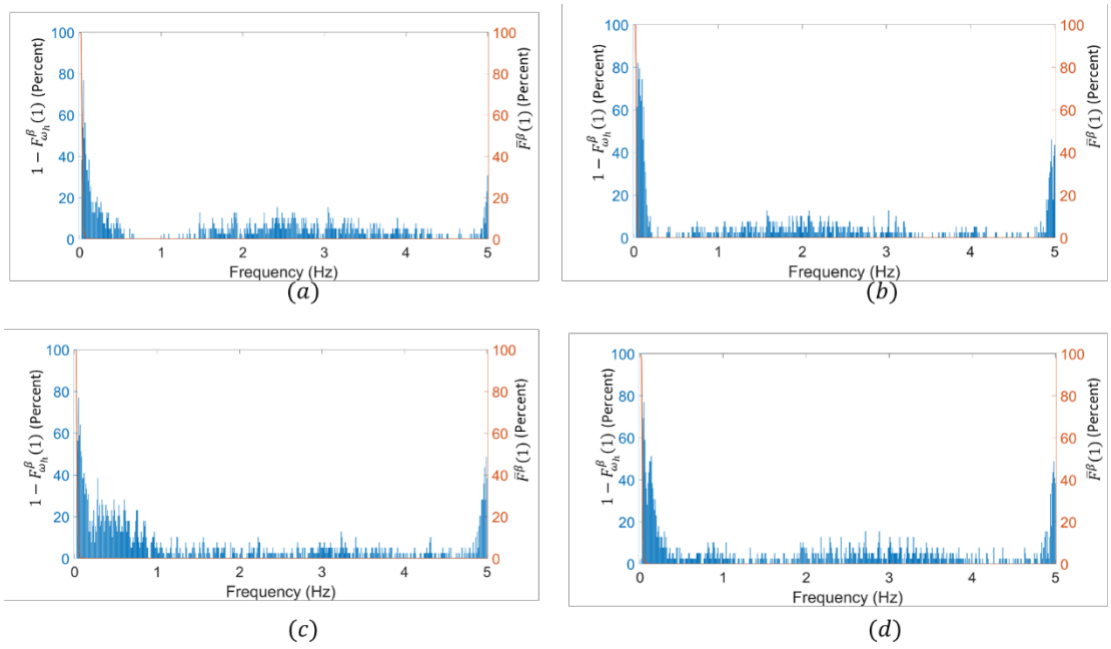

Fig. 12 $1 - F^{\beta}_{\omega_h}(1)$ and $\bar{F}^{\beta}(1)$: (a) vehicle 2; (b) vehicle 3; (c) vehicle 4; (d) vehicle 5

From the probability perspective, the cumulative probability of string instability in the low-frequency region is very high for all vehicles, while the *buffered ECDFM* approaches zero. The results suggest that the ACC vehicles do not perform well in terms of string stability with a low-frequency disturbance, which supports the results of the deterministic string stability conclusion in Fig. 11. Furthermore, there is still some possibility of string instability in the high-frequency region (~ 5Hz). While the general conclusion that the tested ACC vehicles are string unstable is consistent with Gunter et al. (2020), our analysis reveals more detailed results – the frequency ranges in which the vehicles are string unstable and the probabilistic degree of string instability.

We are further interested in the time-variant behavior of commercial ACC. For this, we analyze the Massachusetts ACC data collected by Li et al. (2021); see Fig. 13(a) for an example trajectory. In that experiment, the leading vehicle trajectory is designed to travel at a constant speed initially, then decelerate to create a disturbance (at time 25 s), and eventually accelerate to resume the initial speed (at time 50 s). As shown in Fig. 13(b) and 13(c), the largest STFT norms with the magnitude of 50-60 m/s can be found in the low-frequency domain below 0.2 Hz, when the leading vehicle maintains a constant speed. The autospectrum of the higher frequency range (0.2 to 0.5 Hz) becomes larger when the vehicles experience a disturbance from 25 s to 45 s in the time domain. We further applied the time-variant extension of the FRF to explore the time-frequency domain characteristics; see Fig. 13(d). Rather than analyzing over all frequency ranges, we focus on the range with significant time-frequency features. The result suggests that the commercial ACC in the experiment exhibits significant time-variant features in terms of disturbance dampening ratio, especially when the leading vehicle decelerates. The disturbance dampening

ratio can reach up to 2.5 in the frequency range from 0.3 to 0.5 Hz, which suggests that the disturbance get amplified significantly during deceleration process compared with the relative steady state process. Though the disturbance amplification is to a less extent, we can also find that commercial ACC is not string stable from 0.1 Hz to 0.3 Hz in general, which is consistent with the finding from the OpenACC dataset. The above analysis in Section 5, provided a detailed analysis of time-frequency range of disturbance amplification for commercial ACC. These finding can be applied to leading CAV's trajectory design (e.g., Ma et al., 2017) as well as mixed traffic stabilization via adaptive adjustment of CAV control parameters (e.g., Zhou et al., 2020).

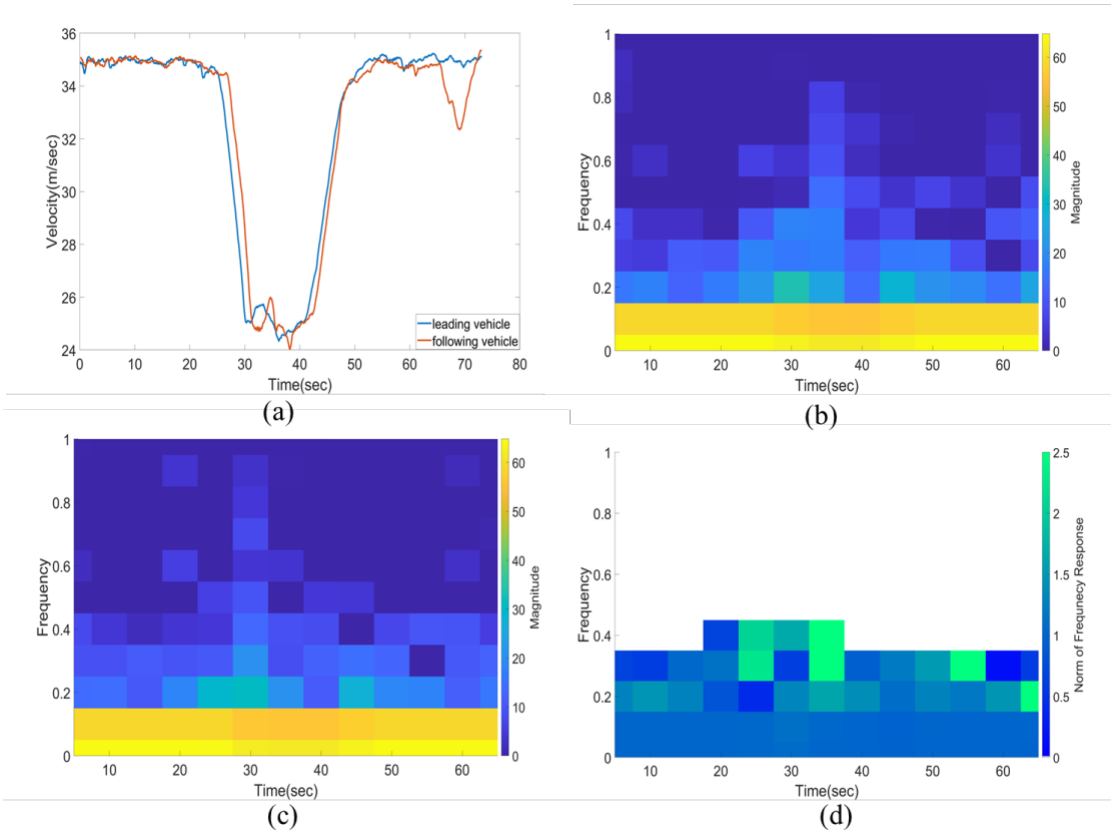

Fig. 13 Time-Variant Extension of the FRF (a) Speed portfolio of the analyzed CF pair; (b) Spectrogram of the leading vehicle; (c) Spectrogram of the following vehicle; (d) Norm of time-varying FRF

6. Conclusion

Disturbance attenuation is an important property for CF control of AVs. This paper presented a data-driven method to approximate the empirical FRF of CF control, whose norm can be used to evaluate the degree of disturbance amplification. Specifically, a data-driven stochastic analysis method was developed based on a linear time-invariant approximation of CF control law to

describe the general CF behavior. We further relaxed the time-invariant assumption and provided a linear time-variant approximation of CF law to describe more detailed time-varying CF features. To facilitate the analysis above, a well-known signal processing method, Welch's method, together with a STFT are developed to extract the empirical transfer function from vehicle trajectories. Based on the empirical FRF, we developed a deterministic and time-invariant data-driven string stability criterion for a qualitative evaluation of string stability classification (i.e., whether a vehicle is string stable or not). Our stochastic and time-variant framework considers measurement noise, estimation error, and time-varying CF features. Further, beyond the qualitative assessment of string stability (i.e., whether a vehicle is string stable or not), our method provides a quantitative assessment of to what extent the system is string stable over frequency and time.

The evaluation using synthetic data showed that the proposed method can estimate transfer function norms close to the theoretical norms and effectively distinguish different levels of string stable/unstable cases. Further, the result also suggests that our data-driven linear approximation can reproduce the disturbance amplification behavior of nonlinear controllers better than the linear theoretical string stability analysis. We further applied the proposed method to evaluate the string stability of ACC vehicles currently on the market using field data. Our results indicate that the tested vehicles are all string unstable in both deterministic and stochastic senses, particularly for low-frequency disturbances (< 0.5 Hz). Further, the results also demonstrate that commercial ACC vehicles exhibit time-varying CF features, especially when the leading vehicle decelerates. This result can be used to tune control parameters, such that the ACC system can better dampen disturbances in this range.

The main contribution of the present paper lies in facilitating a disturbance amplification analysis using data even when the CF controller is unknown or too complex for theoretical analysis. Furthermore, the probabilistic and time-variant extension provides a deeper quantitative insight into the disturbance amplification level over each frequency and time, rather than giving a qualitative conclusion as previously done. Results also demonstrate the potential of our approach to better estimate disturbance amplification for nonlinear CF laws in a data-driven way. The proposed method should be further verified using real field data with known controllers, and a more advanced transfer function method is desirable to better describe the disturbance amplification with high-order CF characteristics. Furthermore, for a more general analysis framework, different norm criteria for a data-driven disturbance amplification analysis should be investigated in the future. Finally, a more in-depth investigation of ACC vehicles on the market is desired to identify the sources of instability in the high-frequency region. Despite these shortcomings, the present study significantly expands the use domain of string stability analysis beyond the simple, known linear controllers.

Acknowledgement

This research is supported by the National Science Foundation (Award CNS 1739869). We also sincerely thank Dr. Michalis Markridis and Dr. Danjue Chen for sharing the first-hand data.

Appendix 1 Disturbance Amplification Analysis for DRL-based Controller

To further show the wide application and potential of the proposed designed framework, we applied our framework to DRL based ACC. Specifically, we use the NGSIM data filtered by Montanino and Punzo (2015) as leading vehicle trajectories, and the following vehicles trajectories are generated by the method developed by Shi et al., (2021). We further split the dataset into two sets: a training set and test set. Each set contains 14 different leader-follower trajectories whose example leader-follower trajectories and speed are given in Fig A1 (a-b). The mean and the standard deviation of the estimated FRF norm for both sets are given in Fig. A1(c-d). As shown, our method can estimate the disturbance amplification for DRL based ACC in a decent manner.

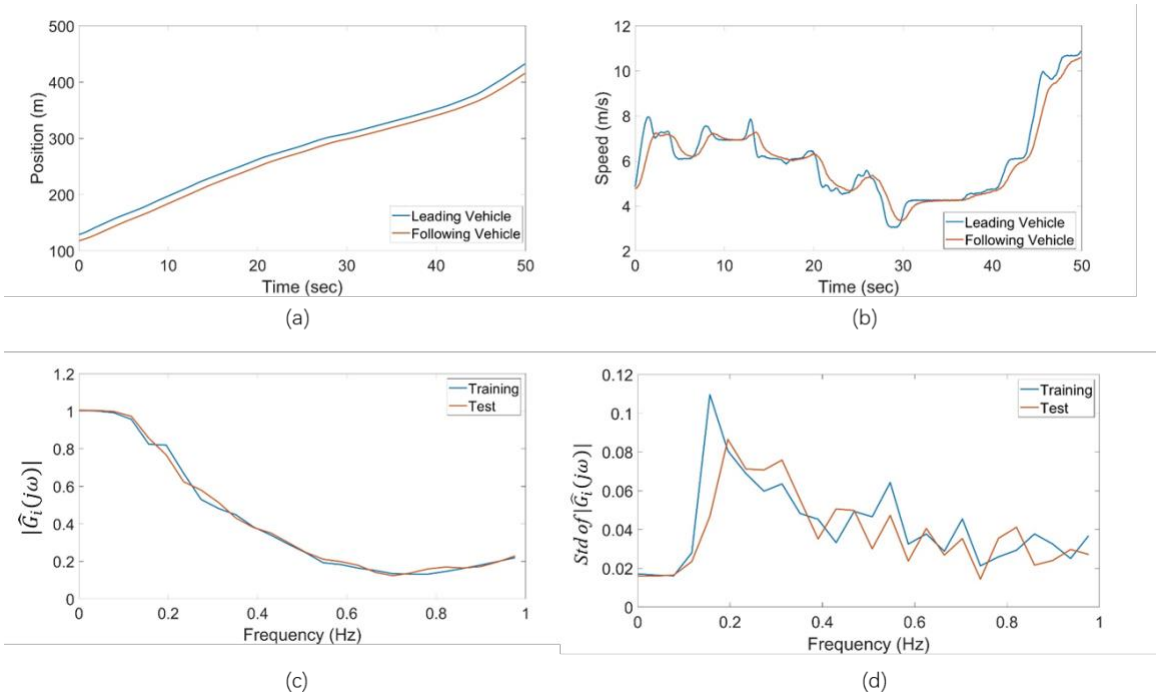

Fig. A1 DRL based ACC Analysis (a) Example CF pair trajectories; (b) Example CF pair trajectories; (c) Estimated average FRF Norm; (d) Standard deviation of estimated FRF Norm

Reference

Administration, N. H. T. S. (2013). LIDAR speed-measuring device performance specifications. *NHTSA, Washington, DC*.

Arem, B. Van, Driel, C. J. G. Van, & Visser, R. (2006). *The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics*. *7*(4), 429–436.

Avargel, Y. and I. Cohen (2007). "On multiplicative transfer function approximation in the short-time Fourier transform domain." *IEEE Signal Processing Letters.* **14**(5): 337-340.

Bando, M., Hasebe, K., Nakanishi, K., & Nakayama, A. (1998). Analysis of optimal velocity model with explicit delay. *Physical Review E*, *58*(5), 5429.

Bian, Y., Zheng, Y. pd., Ren, W., Eben, S., & Wang, J. (2019a). Reducing time headway for platooning of connected vehicles via V2V communication. *Transportation Research Part C*, *102*(March), 87–105. https://doi.org/10.1016/j.trc.2019.03.002

Bian, Y., Zheng, Y., Ren, W., Eben, S., & Wang, J. (2019b). Reducing time headway for platooning of connected vehicles via V2V communication. *Transportation Research Part C*, *102*(March), 87–105. https://doi.org/10.1016/j.trc.2019.03.002

Çakrak， Ferhat, Loughlin, P. J. (2001). *Multiple Window Time-Varying Spectral Analysis*. *49*(2), 448–453.

Chen, D., Ahn, S., Laval, J., & Zheng, Z. (2014). On the periodicity of traffic oscillations and capacity drop : The role of driver characteristics. *TRANSPORTATION RESEARCH PART B*, *59*, 117–136. https://doi.org/10.1016/j.trb.2013.11.005

Cheng, R., Orosz, G., Murray, R. M., & Burdick, J. W. (2019, July). End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks. *In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 3387-3395).*

Eberhard, A. (1973). An Optimal Discrete Window for the Calculation of Power Spectra. *IEEE TRANSACTIONS ON AUDIO AND ELECTROACOUSTICS*, (9).

Faust, O., U. R. Acharya, H. Adeli and A. Adeli (2015). "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis." *Seizure.26*, 56-64.

Gong, S., & Du, L. (2018). Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles. *Transportation Research Part B*, *116*, 25–61. https://doi.org/10.1016/j.trb.2018.07.005

Gong, S., Shen, J., & Du, L. (2016). *Constrained optimization and distributed computation based car following control of a connected and autonomous vehicle platoon*. *Transportation Research Part B*, *94*, 314–334. https://doi.org/10.1016/j.trb.2016.09.016

Gong, S., Zhou, A., & Peeta, S. (2019). *Cooperative Adaptive Cruise Control for a Platoon of Connected and Autonomous Vehicles considering Dynamic Information Flow Topology*. *2673*(10), 185–198. https://doi.org/10.1177/0361198119847473

Gunter, G., Janssen, C., Barbour, W., Stern, R. E., & Work, D. B. (2020). *Model-Based String Stability of Adaptive Cruise Control Systems Using Field Data*. *5*(1), 90–99.

Hu, J., Shao, Y., Sun, Z., Wang, M., Bared, J., & Huang, P. (2018). Integrated optimal eco-driving on rolling terrain for hybrid electric vehicle with vehicle-infrastructure communication.

TRANSPORTATION RESEARCH PART C, *68*(2016), 228–244. https://doi.org/10.1016/j.trc.2016.04.009

Jin, I. G., & Orosz, G. (2014). Dynamics of connected vehicle systems with delayed acceleration feedback. *Transportation Research Part C: Emerging Technologies*, *46*, 46-64.

Jin, I. G., & Orosz, G. (2018). Connected cruise control among human-driven vehicles: Experiment-based parameter estimation and optimal control design. *Transportation research part C: emerging technologies, 95,* 445-459.

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

Li, K., Gao, F., Li, S. E., Zheng, Y., & Gao, H. (2018). *Robust cooperation of connected vehicle systems with eigenvalue-bounded interaction topologies in the presence of uncertain dynamics*.

Li, X., Wang, X., & Ouyang, Y. (2012). Prediction and field validation of traffic oscillation propagation under nonlinear car-following laws. *Transportation research part B: methodological*, *46*(3), 409-423.

Ma, J., Li, X., Zhou, F., Hu, J., & Park, B. B. (2017). Parsimonious shooting heuristic for trajectory design of connected automated traffic part II: Computational issues and optimization. *Transportation Research Part B: Methodological, 95,* 421-441.

Makridis, M., Mattas, K., Anesiadou, A., & Ciuffo, B. (2020). *openACC . An open database of car-following experiments to study the properties of commercial ACC systems .* 1–27.

Milanés, V., & Shladover, S. E. (2014). *Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data*. *48*, 285–300. https://doi.org/10.1016/j.trc.2014.09.001

Milanés, V., Shladover, S. E., Spring, J., & Nowakowski, C. (2014). *Cooperative Adaptive Cruise Control in Real Traffic Situations*. *15*(1), 296–305.

Montanino, M., & Punzo, V. (2015). Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns. *TRANSPORTATION RESEARCH PART B*, *80*, 82–106. https://doi.org/10.1016/j.trb.2015.06.010

Montanino, M., Monteil, J., & Punzo, V. (2021). From homogeneous to heterogeneous traffic flows: Lp String stability under uncertain model parameters. *Transportation Research Part B: Methodological*, *146*, 136-154.

Montanino, M., & Punzo, V. (2021). On string stability of a mixed and heterogeneous traffic flow: A unifying modelling framework. *Transportation Research Part B: Methodological*, *144*, 133-154.

Naus, G. J. L., Vugts, R. P. A., Ploeg, J., & Molengraft, M. R. J. G. Van De. (2010). *String-Stable CACC Design and Experimental Validation : A Frequency-Domain Approach*. *59*(9), 4268–4279.

Naus, G., & Ploeg, J. (2010). *Cooperative adaptive cruise control , design and experiments*. (1), 6145–6150.

Ploeg, J., Semsar-kazerooni, E., Lijster, G., Wouw, N. Van De, & Nijmeijer, H. (2013). *Graceful*

Degradation of CACC Performance Subject to Unreliable Wireless Communication. (Itsc), 1210–1216.

Ponn, T., Müller, F., & Diermeyer, F. (2019). Systematic analysis of the sensor coverage of automated vehicles using phenomenological sensor Models. *2019 IEEE Intelligent Vehicles Symposium (IV)*, 1000–1006. IEEE.

Punzo, V., Teresa, M., & Ciuffo, B. (2011). *On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data*. *19*, 1243–1262. https://doi.org/10.1016/j.trc.2010.12.007

Qin, W. B., & Orosz, G. (2017). Scalable stability analysis on large connected vehicle systems subject to stochastic communication delays. *Transportation Research Part C: Emerging Technologies, 83,* 39-60.

Qu, X., Yu, Y., Zhou, M., Lin, C., & Wang, X. (2020). Jointly dampening tra ffi c oscillations and improving energy consumption with electric , connected and automated vehicles : A reinforcement learning based approach. *Applied Energy*, *257*(October 2019), 114030. https://doi.org/10.1016/j.apenergy.2019.114030

Randall, R. B. (2008). Spectral Analysis and Correlation. In D. Havelock, S. Kuwano, & M. Vorländer (Eds.), *Handbook of Signal Processing in Acoustics* (pp. 33–52). https://doi.org/10.1007/978-0-387-30441-0_3

Samiee, K., P. Kovacs and M. Gabbouj (2014). "Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform." *IEEE transactions on Biomedical Engineering, 62*(2), 541-552.

Schoukens, J., & Godfrey, K. (2018a). Nonparametric. *IEEE Control Systems*, *38*(August), 49–88. https://doi.org/10.1109/MCS.2018.2830080

Schoukens, J., & Godfrey, K. (2018b). Nonparametric data-driven modeling of linear systems: Estimating the frequency response and impulse response function. *IEEE Control Systems*, *38*(August), 49–88. https://doi.org/10.1109/MCS.2018.2830080

Shi, H., Zhou, Y., Wu, K., Wang, X., Lin, Y., & Ran, B. (2021). Connected automated vehicle cooperative control with a deep reinforcement learning approach in a mixed traffic environment. *Transportation Research Part C: Emerging Technologies*, *133*, 103421.

Stern, R. E., Cui, S., Laura, M., Monache, D., Bhadani, R., Bunting, M., … Work, D. B. (2018). Dissipation of stop-and-go waves via control of autonomous vehicles : Field experiments. *Transportation Research Part C*, *89*(April 2017), 205–221. https://doi.org/10.1016/j.trc.2018.02.005

Swaroop, D. (1996). *String Stability of Interconnected Systems*. *41*(3).

Swaroop, D., Hedrick, K., Chien, C., & IOANNOU, P. (1994). A Comparision of Spacing and Headway Control Laws for Automatically Controlled Vehicles1. *Vehicle System Dynamics*, *23*(1), 597–625. https://doi.org/10.1080/00423119408969077

Talebpour, A., & Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C*, *71*, 143–163. https://doi.org/10.1016/j.trc.2016.07.007

Giammarino, V., Baldi, S., Frasca, P., & Delle Monache, M. L. (2020). Traffic flow on a ring with a

single autonomous vehicle: An interconnected stability perspective. *IEEE Transactions on Intelligent Transportation Systems*, *22*(8), 4998-5008.
Wang, C., Gong, S., Zhou, A., Li, T., & Peeta, S. (2018). *Cooperative Adaptive Cruise Control for Connected Autonomous Vehicles by Factoring Communication-Related Constraints*. *00*(July), 24–26.
Wang, C., Gong, S., Zhou, A., Li, T., & Peeta, S. (2019). Cooperative adaptive cruise control for connected autonomous vehicles by factoring communication-related constraints ☆. *Transportation Research Part C*, (November 2018), 1–22. https://doi.org/10.1016/j.trc.2019.04.010
Wang, M. (2018a). Infrastructure assisted adaptive driving to stabilise heterogeneous vehicle strings. *Transportation Research Part C*, *91*(April), 276–295. https://doi.org/10.1016/j.trc.2018.04.010
Wang, M. (2018b). Infrastructure assisted adaptive driving to stabilise heterogeneous vehicle strings. *Transportation Research Part C*, *91*(April 2017), 276–295. https://doi.org/10.1016/j.trc.2018.04.010
Wang, M., Daamen, W., Hoogendoorn, S. P., & Arem, B. Van. (2014). Rolling horizon control framework for driver assistance systems . Part II : Cooperative sensing and cooperative control. *Transportation Research Part C*, *40*, 290–311. https://doi.org/10.1016/j.trc.2013.11.024
Welch, P. D. (1967). *The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Aver. aging Over Short, Modified Periodograms*. (2), 70–73.
Wexler, J. and S. Raz (1990). "Discrete gabor expansions." *Signal processing, 21(3)*, 207-220.
Wilson, R. E., & Ward, J. A. (2011). Car-following models: fifty years of linear stability analysis–a mathematical perspective. *Transportation Planning and Technology*, *34*(1), 3-18.
Wu, F., Stern, R., Churchill, M., Laura, M., Monache, D., Piccoli, B., … Han, K. (2017). *Measuring trajectories and fuel consumption in oscillatory traffic : experimental results To cite this version : HAL Id : hal-01516133 Measuring trajectories and fuel consumption in oscillatory traffic : experimental results*.
Wu, X., Qin, G., Yu, H., Gao, S., Liu, L., & Xue, Y. (2016). *Optik Using improved chaotic ant swarm to tune PID controller on cooperative adaptive cruise control*. *127*, 3445–3450. https://doi.org/10.1016/j.ijleo.2015.12.014
Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., & Qu, X. (2020). Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle ☆. *Transportation Research Part C*, *114*(September 2018), 205–224. https://doi.org/10.1016/j.trc.2020.02.018
Zheng, Z., Ahn, S., Chen, D., & Laval, J. (2011). Applications of wavelet transform for analysis of freeway traffic : Bottlenecks , transient traffic , and traffic oscillations. *Transportation Research Part B*, *45*(2), 372–384. https://doi.org/10.1016/j.trb.2010.08.002
Zheng, Z., Ahn, S., & Monsere, C. M. (2010). *Impact of traffic oscillations on freeway crash occurrences*. *42*, 626–636. https://doi.org/10.1016/j.aap.2009.10.009
Zhou, Y., Ahn, S., Wang, M., & Hoogendoorn, S. (2020). Stabilizing mixed vehicular platoons with

connected automated vehicles: An H-infinity approach. *Transportation Research Part B: Methodological*, *132*, 152–170.

Zhou, Y., & Ahn, S. (2019a). Robust local and string stability for a decentralized car following control strategy for connected automated vehicles. *Transportation Research Part B*, *125*, 175–196. https://doi.org/10.1016/j.trb.2019.05.003

Zhou, Y., Ahn, S., Chitturi, M., & Noyce, D. A. (2017). Rolling horizon stochastic optimal control strategy for ACC and CACC under uncertainty. *Transportation Research Part C*, *83*, 61–76. https://doi.org/10.1016/j.trc.2017.07.011

Zhou, Y., Wang, M., & Ahn, S. (2019b). Distributed model predictive control approach for cooperative car-following with guaranteed local and string stability. *Transportation Research Part B*, *128*, 69–86. https://doi.org/10.1016/j.trb.2019.07.001