

Stochastic Modeling and Simulations of Car-Following Behaviors

Chengyuan Zhang (enzozcy@gmail.com)

Ph.D. Candidate (advised by Prof. Lijun Sun)

Department of Civil Engineering

McGill University

<https://chengyuan-zhang.github.io/>

Feb. 13, 2025

JTL Research Seminar @ MIT (online)



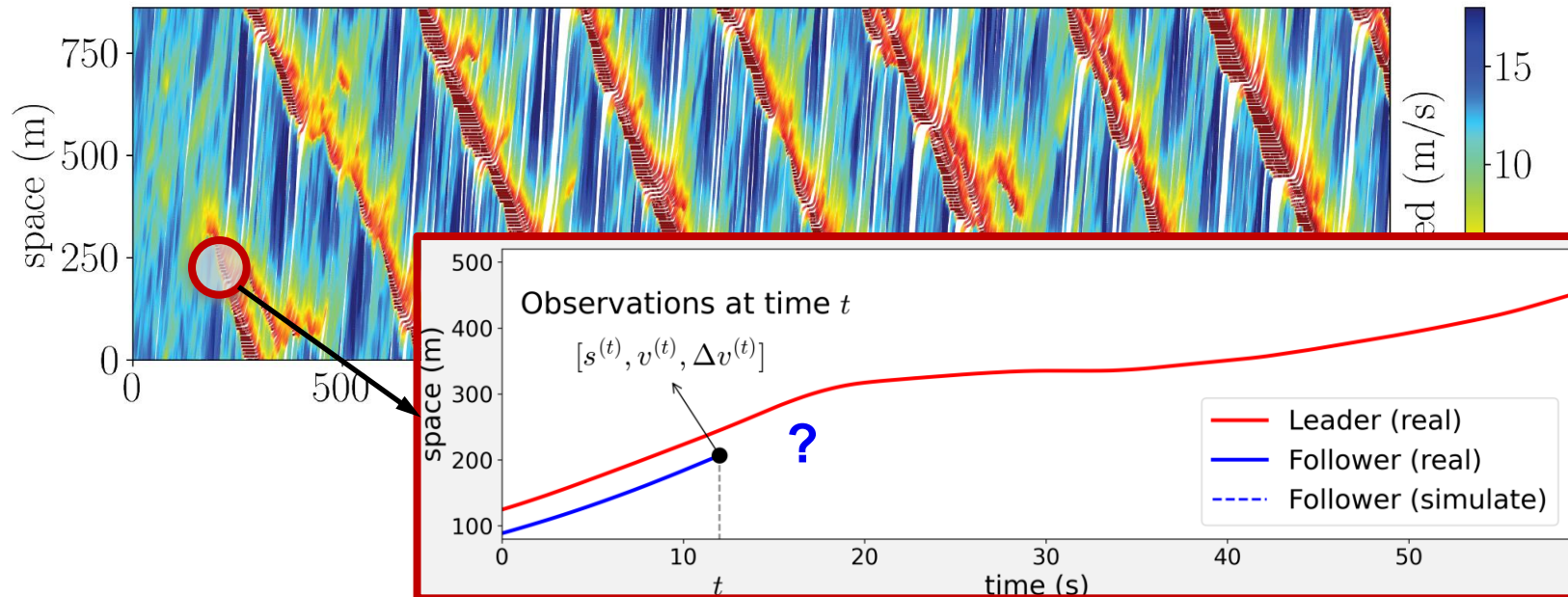
McGill

About me

- Smart Transportation Lab at McGill University
- Research interests:
 - **Traffic flow theory;**
 - **Traffic Simulation;**
 - **Human driving behavior;**
 - **Bayesian learning;**
 - **Multi-agent interaction modeling.**
- I aim to bridge the gap between **theoretical modeling** and **practical traffic simulation** using advanced **statistical techniques**. Driven by a passion for understanding human driving behavior, my work seeks to enhance microscopic traffic simulations, ultimately contributing to safer and more efficient transportation systems.

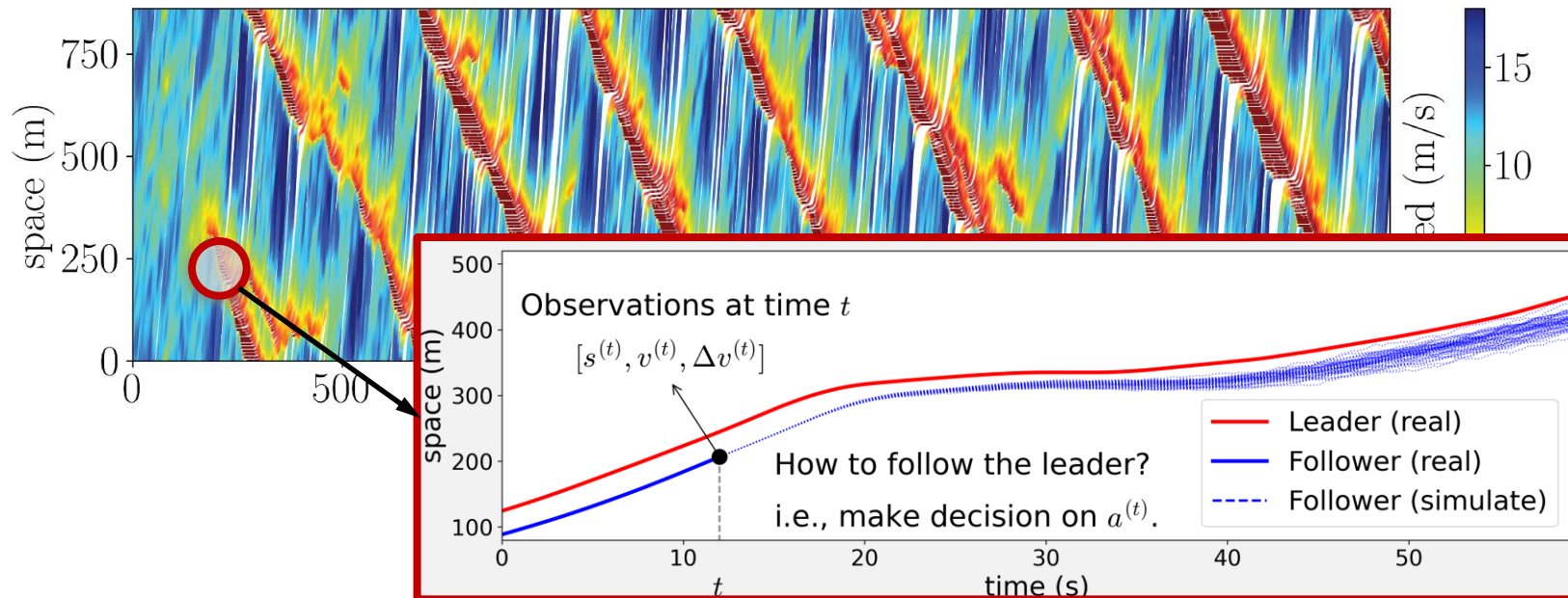
Motivation / background

- How would the vehicle react in response to the leading vehicle?



Motivation / background

- How would the vehicle react in response to the leading vehicle?



- What do we need for traffic simulations?

Motivation / background

- The goal of traffic simulations:
 - **Past:** reproduce traffic phenomenon.
 - **Future:** support the development and test of control algorithms:
 - Connected and Automated Vehicle;
 - Reinforcement learning for traffic control/management;
 - Human drivers still involved;
 - Safety, predictability, and uncertainty;



(How well are the blue cars performing in this simulator?)

Motivation / background

- The goal of traffic simulations:
 - **Past:** reproduce traffic phenomenon.
 - **Future:** support the development and test of control algorithms:
 - Connected and Automated Vehicle;
 - Reinforcement learning for traffic control/management;
 - Human drivers still involved;
 - Safety, predictability, and uncertainty;
- How do we introduce **realistic** randomness?
 - × Deterministic car-following models;
 - ✓ **Probabilistic** car-following models with **uncertainty quantification**.

In this presentation, we are interested in:

- How do we model the human-driver car-following behaviors?
- How do we simulate human-like car-following behaviors?

Outline

- **Background**

- Temporal correlations in driving behaviors
- The general form of car-following models

- **Probabilistic Modeling Framework**

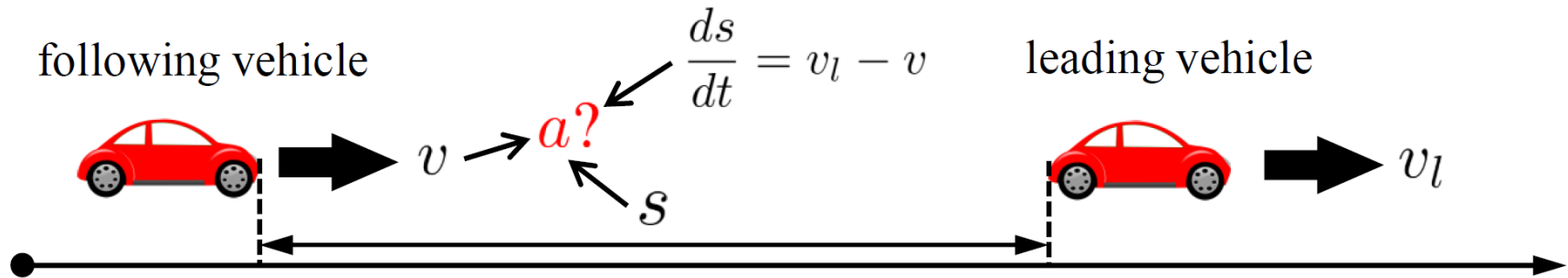
- Stationary: IDM (mean model) + stochastic processes (GP)
- Nonstationary: NN (mean model) + stochastic processes (GP)

- **Stochastic Simulation**

- Short-term single car-following pair
- Long-term multiple car-following pairs

- **Discussions**

Intelligent driver model



Intelligent Driver Model (IDM) (Treiber et al. 2000)

$$a_{\text{IDM}} = \alpha \left(\boxed{1 - \left(\frac{v}{v_0} \right)^\delta} - \boxed{\left(\frac{s^*(v, \Delta v)}{s} \right)^2} \right)$$

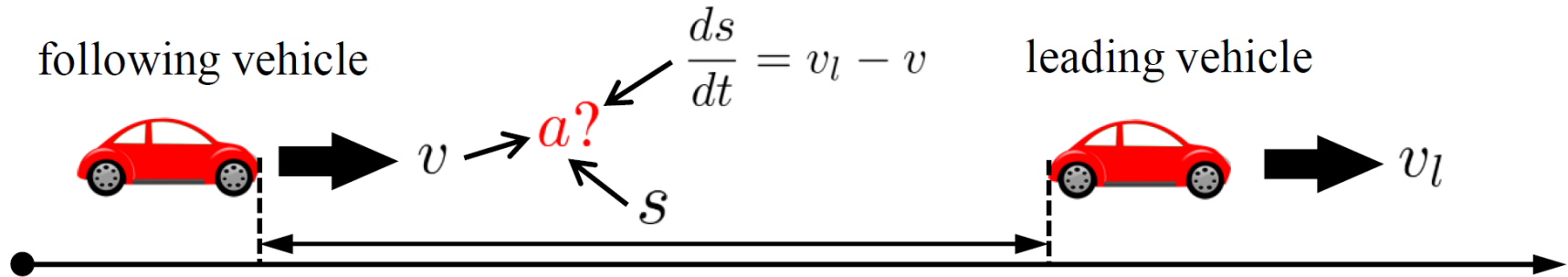
free-flow interaction

$$s^*(v, \Delta v) = s_0 + s_1 \sqrt{\frac{v}{v_0}} + vT + \frac{v \Delta v}{2 \sqrt{\alpha \beta}}$$

- v_0 : desired speed;
- s_0 : jam spacing;
- T : time headway;
- α : maximum acceleration;
- β : comfortable deceleration rate.

$$\theta = \underbrace{[v_0, s_0, T, \alpha, \beta]}_{\text{parameter set}}$$

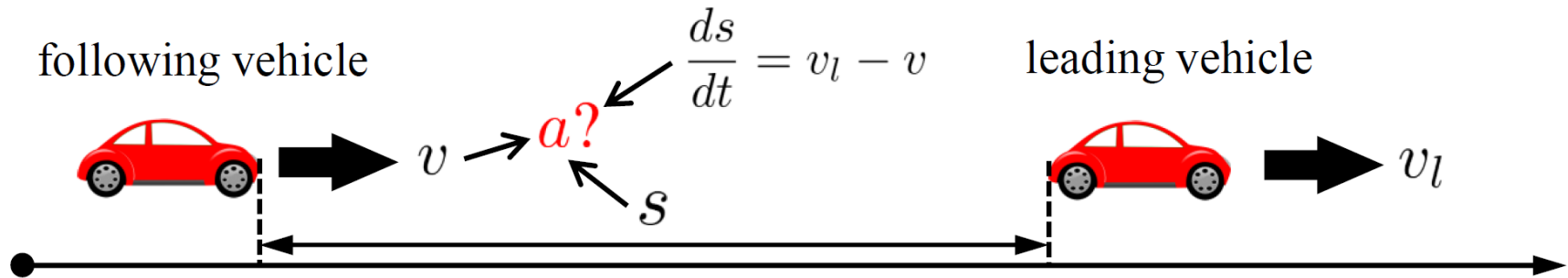
Intelligent driver model



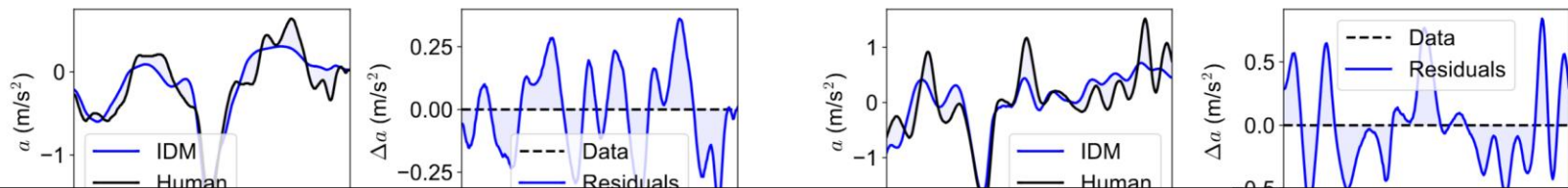
- Gaussian IDM assumes: $a^{(t)} \approx a_{\text{IDM}}^{(t)}$. \Rightarrow Likelihood: $\mathcal{N}(\hat{a}^{(t)} | a_{\text{IDM}}^{(t)}, \sigma_\epsilon^2)$
- Calibration by MLE: $\max_{\theta} \prod_{t=1}^T \text{likelihood}$ and $\theta = \underbrace{[v_0, s_0, T, \alpha, \beta]}_{\text{parameter set}}$
- Loss function in literature ([Punzo et al. 2021](#)):

$$\min_{\theta} \frac{1}{T} \sum_{t=1}^T \underbrace{(a_{\text{IDM}}^{(t)} - \hat{a}^{(t)})^2}_{\text{acceleration}} + \frac{\alpha}{T} \sum_{t=1}^T \underbrace{(v_{\text{IDM}}^{(t)} - \hat{v}^{(t)})^2}_{\text{speed}} + \frac{\beta}{T} \sum_{t=1}^T \underbrace{(x_{\text{IDM}}^{(t)} - \hat{x}^{(t)})^2}_{\text{position}}$$

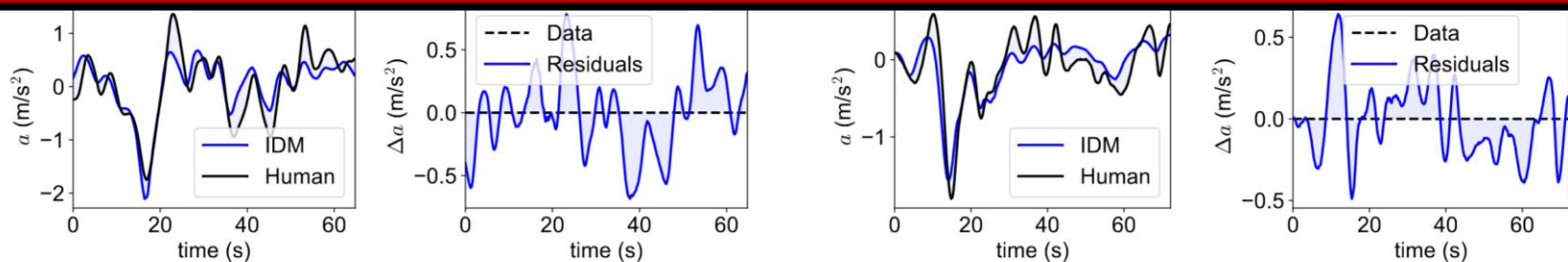
Intelligent driver model



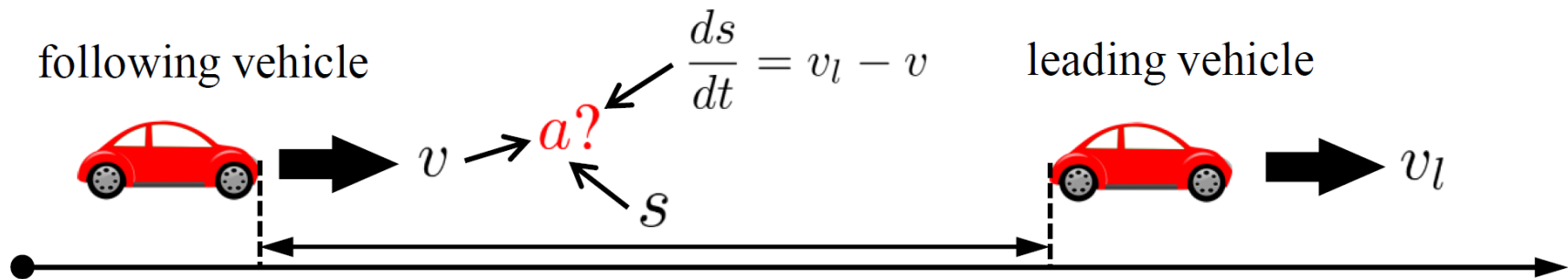
- IDM assumes: $a^{(t)} \approx a_{\text{IDM}}^{(t)}$. Let's visualize the residuals $a^{(t)} - a_{\text{IDM}}^{(t)}$.



**IDM captures much information,
but *some are still left in the residuals!***

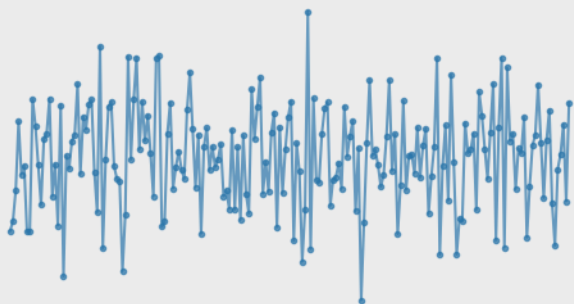


Temporal correlations in driving behaviors



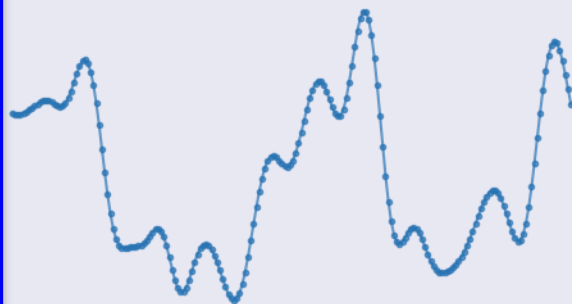
- For a human-driver CF model, what do we miss?
 - **Temporally correlated errors / Time delay**
 - **Heteroscedasticity**

i.i.d. Noise (No Correlation, Constant Variance)



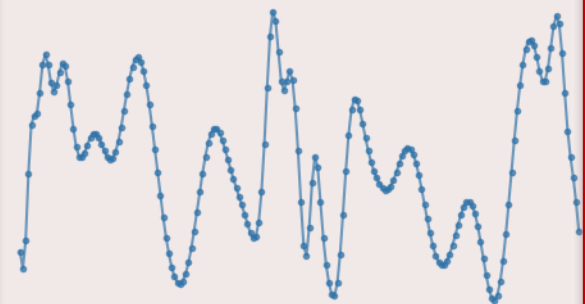
$$\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

Homoscedastic Noise (GP with RBF Kernel)



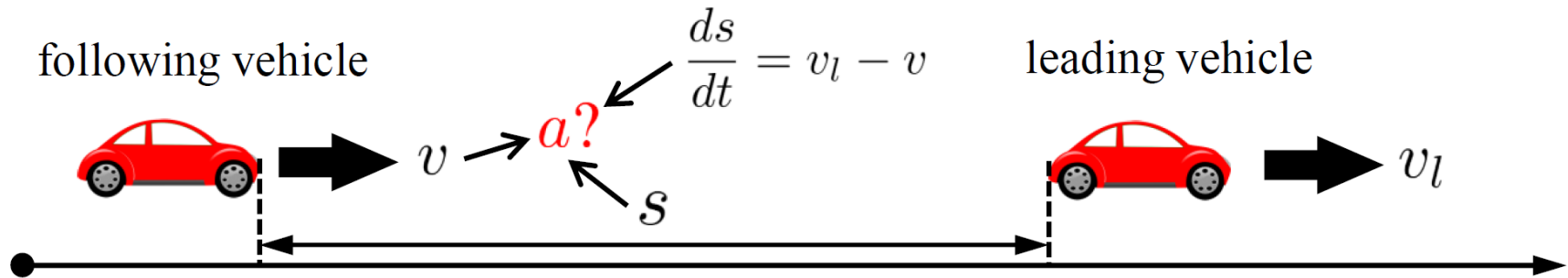
$$\mathcal{N}(\mathbf{0}, \Sigma)$$

Heteroscedastic Noise (GP with Gibbs Kernel)



$$\mathcal{N}(\mathbf{0}, \Sigma_t)$$

Temporal correlations in driving behaviors

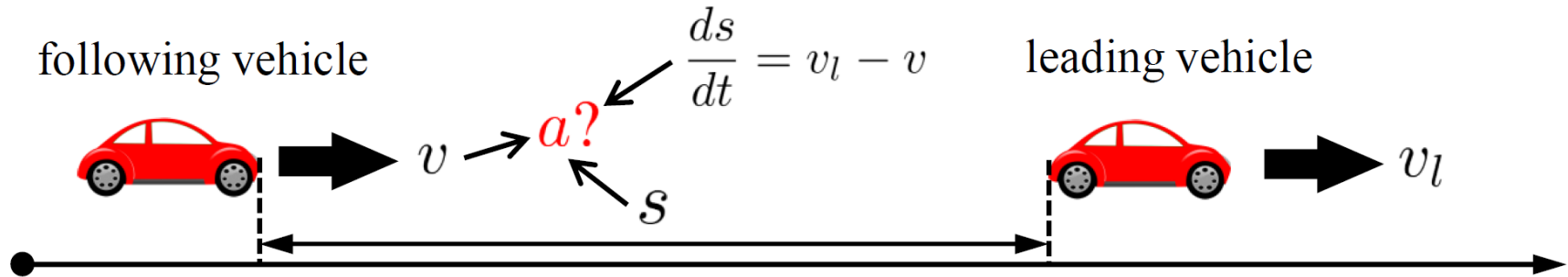


- For a human-driver CF model, what do we miss?
 - **Temporally correlated errors / Time delay**
 - **Heteroscedasticity**

Core contribution:

- Therefore, we assume $a(x, t) \approx \underbrace{f_{\text{CFM}}(x; \theta)}_{\text{mean car-following model}} + \underbrace{\delta(t)}_{\text{temporal correlations (capture how past behaviors influence future behaviors)}}$ (Zhang et al. 2025)

The general form of car-following models



- We assume: $a(x, t) \approx \boxed{f_{\text{CFM}}(x; \theta)} + \boxed{\delta(t)}$ (Zhang et al. 2025)
- IDM assumes: $a(x, t) \approx \boxed{a_{\text{IDM}}(x; \theta)}$. (Treiber et al. 2000)

Missed the temporal part $\boxed{\delta(t)}$

TO-DO:

- Consider $\delta(t)$ in modeling;
- Model $\delta(t+1)|\delta(t)$ in simulation.

Memory-Augmented IDM (MA-IDM) (Zhang and Sun 2024)

How to model $\delta(t)$ and $\delta(t+1)|\delta(t)$?

- We assume:

$$a(x, t) = f_{\text{CFM}}(x; \theta) + \delta(t) + \epsilon_t, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

- IDM assumes:

$$a^{(t)} = a_{\text{IDM}}^{(t)} + \epsilon_t, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2) \quad \Rightarrow \quad a|i, \theta \sim \mathcal{N}(a_{\text{IDM}}, \sigma_\epsilon^2 I)$$

- MA-IDM assumes:

Vector form with Multivariate Normal

$$a^{(t)} = a_{\text{IDM}}^{(t)} + a_{\text{GP}}^{(t)} + \epsilon_t \quad \Rightarrow \quad a|i, \theta \sim \mathcal{N}(a_{\text{IDM}}, K + \sigma_\epsilon^2 I)$$

mean model residuals i.i.d. error where K is a kernel matrix .

[Chengyuan Zhang and Lijun Sun. (2024). Bayesian calibration of the intelligent driver model. *IEEE Transactions on Intelligent Transportation Systems*.]

Memory-Augmented IDM (MA-IDM) (Zhang and Sun 2024)

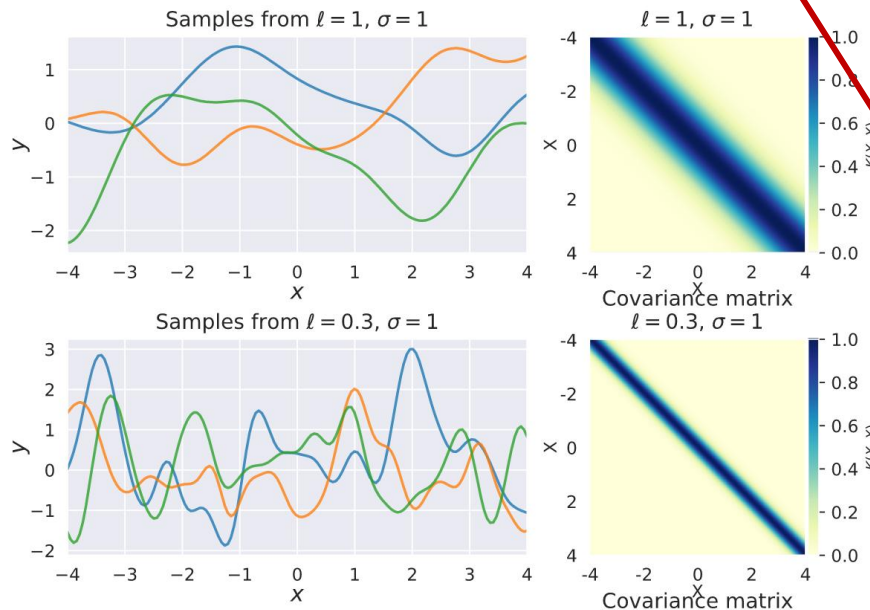
- IDM assumes:

$$a^{(t)} = a_{\text{IDM}}^{(t)} + \epsilon_t, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2) \quad \Rightarrow \mathbf{a} | \mathbf{i}, \boldsymbol{\theta} \sim \mathcal{N}(\mathbf{a}_{\text{IDM}}, \sigma_\epsilon^2 \mathbf{I})$$

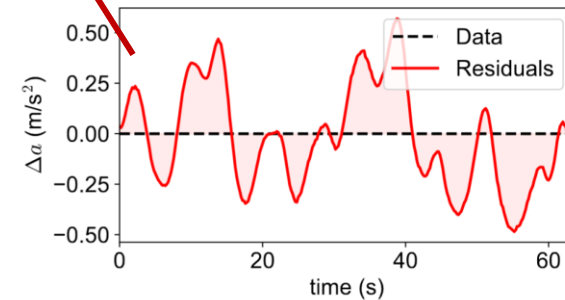
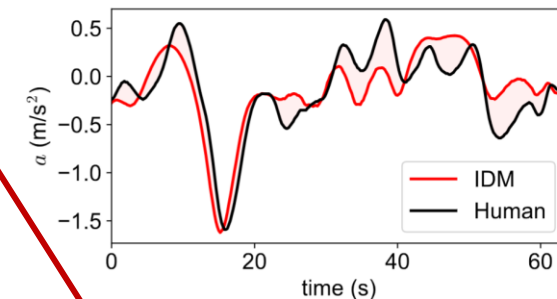
- MA-IDM assumes:

$$a^{(t)} = a_{\text{IDM}}^{(t)} + \boxed{a_{\text{GP}}^{(t)}} + \epsilon_t \quad \Rightarrow \mathbf{a} | \mathbf{i}, \boldsymbol{\theta} \sim \mathcal{N}(\mathbf{a}_{\text{IDM}}, \boxed{\mathbf{K}} + \sigma_\epsilon^2 \mathbf{I})$$

- Gaussian processes



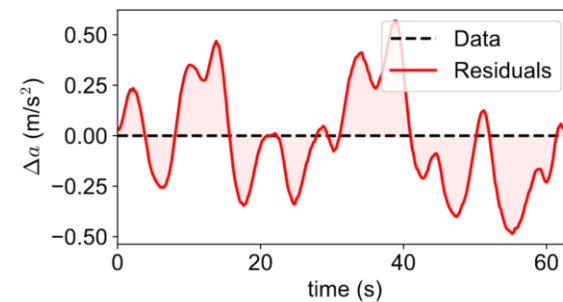
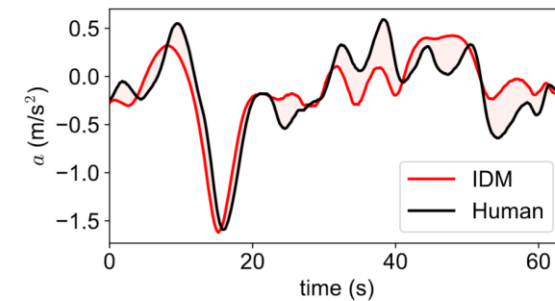
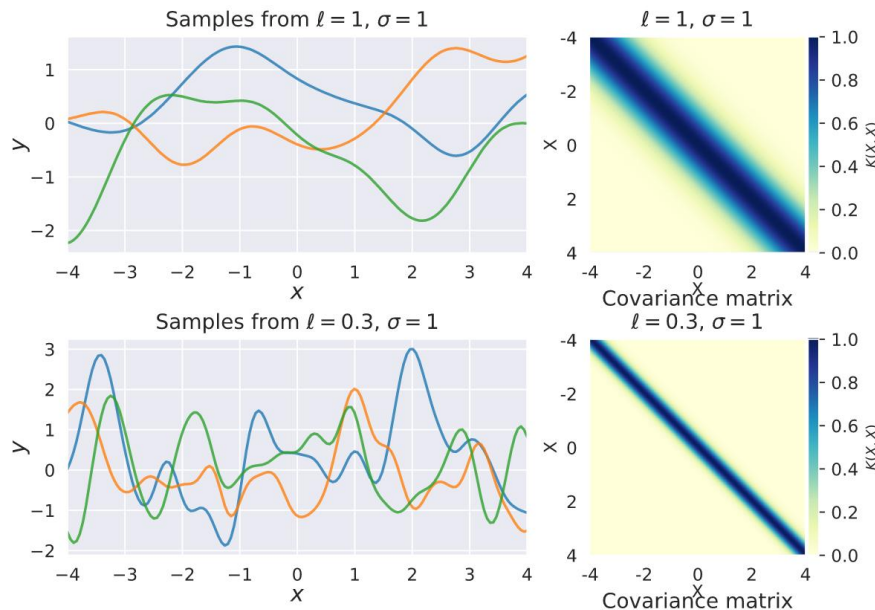
where \mathbf{K} is a kernel matrix .



Memory-Augmented IDM (MA-IDM) (Zhang and Sun 2024)

Construct the kernel matrix K :

- SE kernel: $k_{\text{SE}}(t, t'; \lambda) := \sigma^2 \exp\left(-\frac{d(t, t')^2}{2\ell^2}\right),$
- Matern kernel: $k_{\text{Matérn}}^\nu(t, t'; \lambda) := \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{d(t, t')}{\ell}\right)^\nu K_\nu\left(\sqrt{2\nu} \frac{d(t, t')}{\ell}\right),$
- Gaussian processes



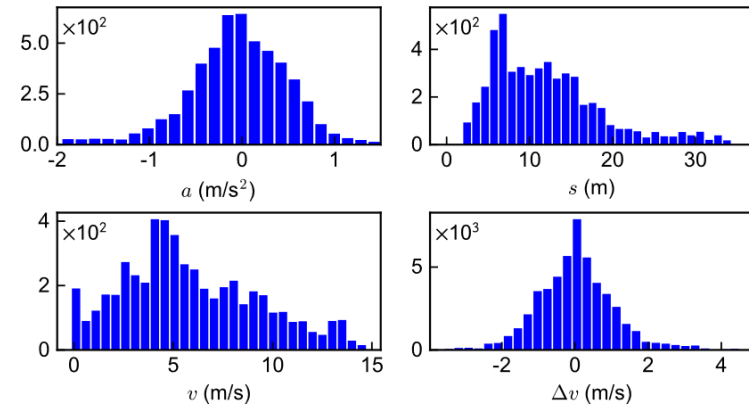
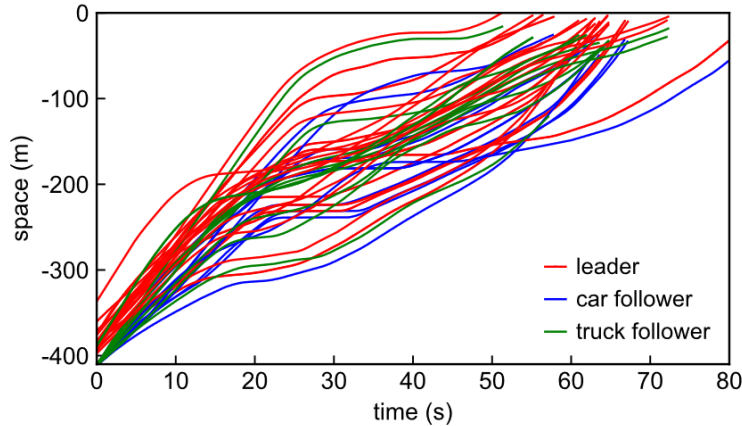
Experiments – Car-Following Data Extraction

- **HighD** dataset:

(Krajewski et al. 2018)

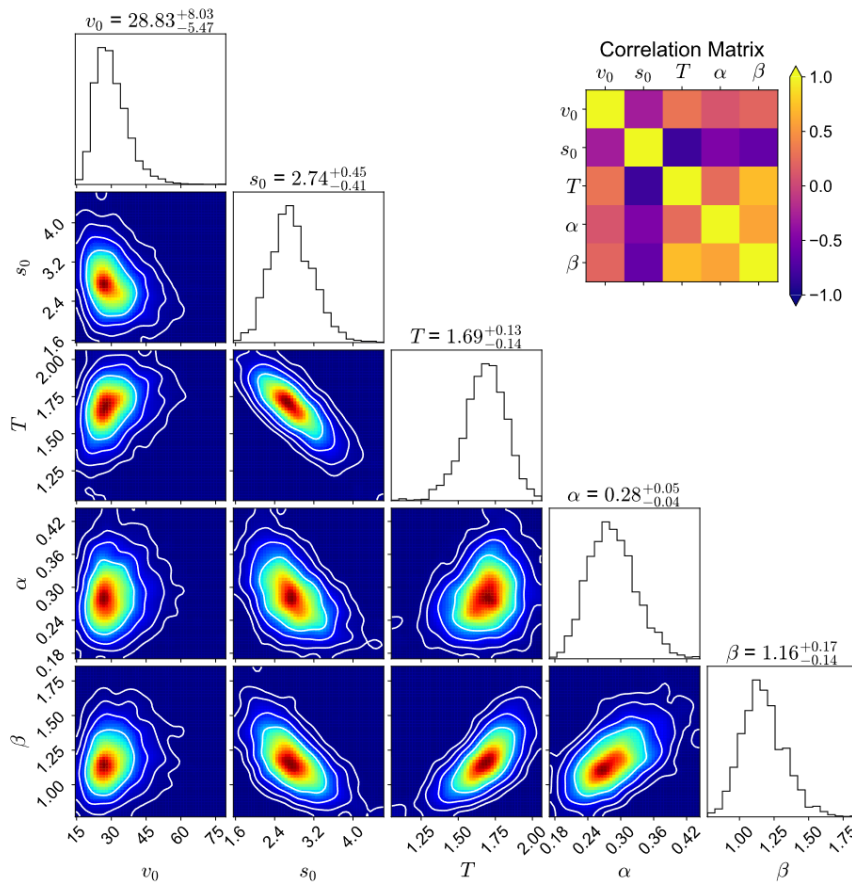
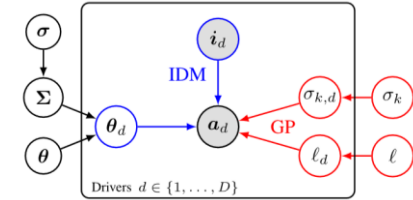
<https://levelxdata.com/highd-dataset/>

- naturalistic vehicle trajectories → leader-follower pairs.

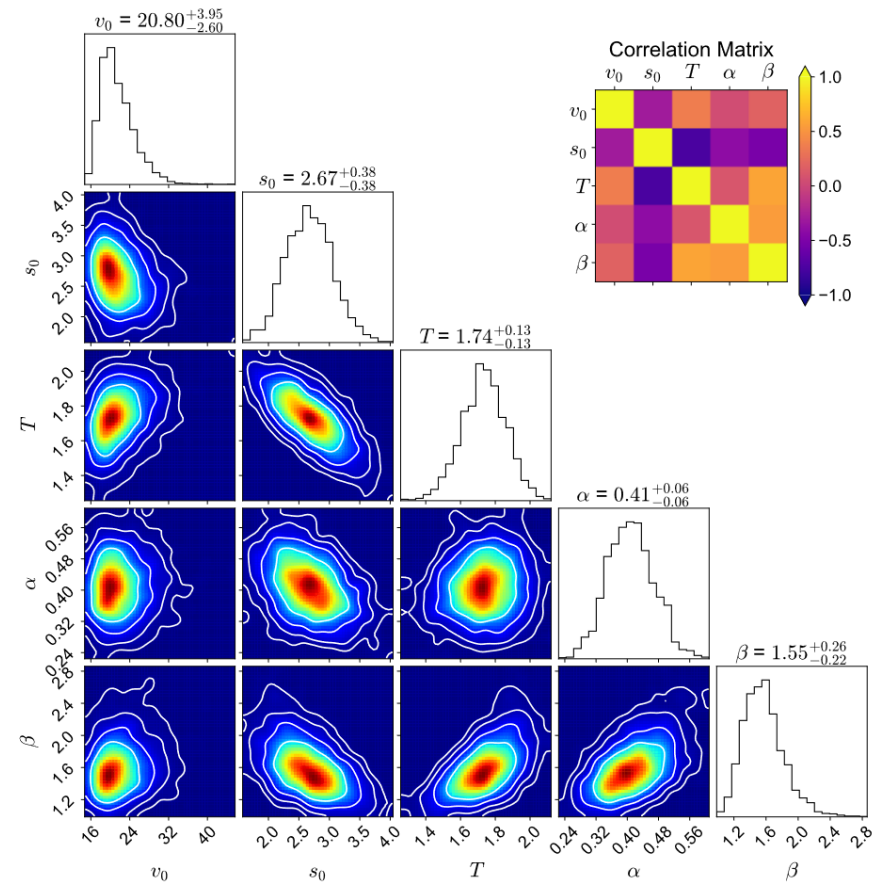


Experiments – Identified IDM Parameters

Similar posterior distribution shape but **more concentrated**



(a) Hierarchical B-IDM posteriors of truck #211.



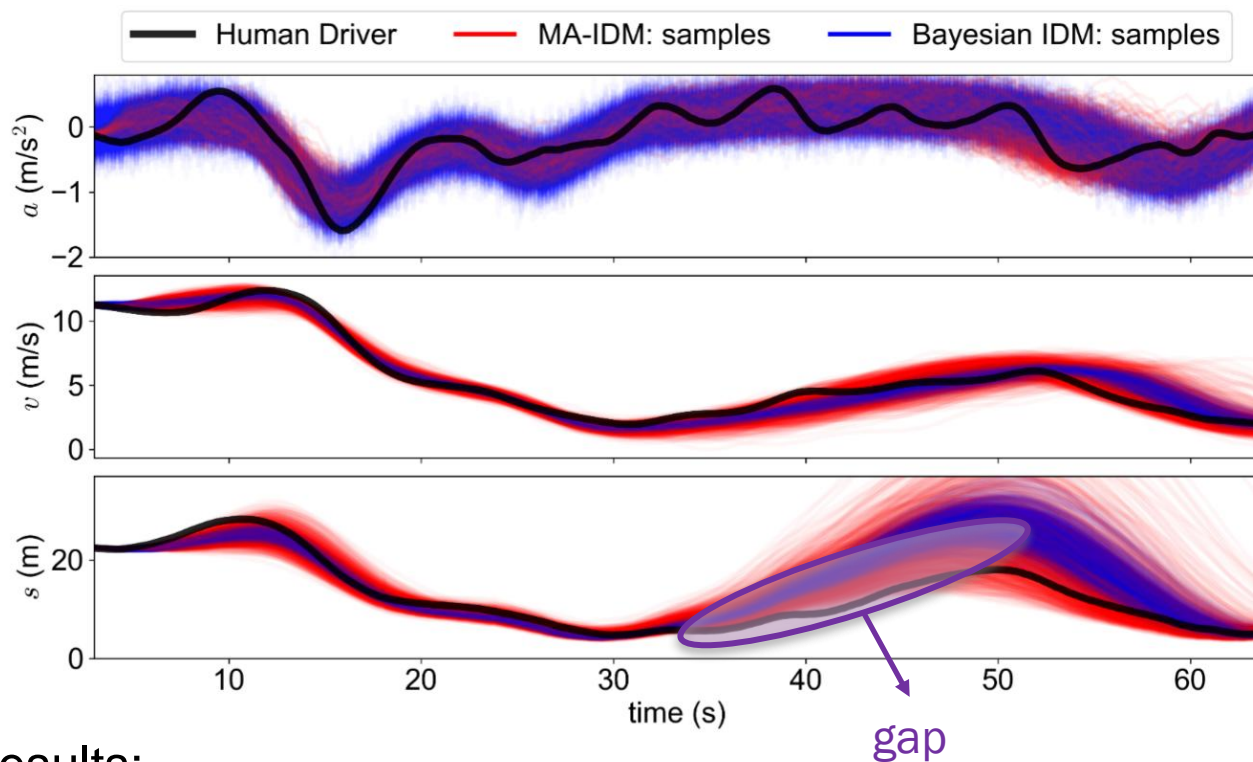
(b) Hierarchical MA-IDM posteriors of truck #211.

We can draw samples (IDM parameters) from the posterior distributions!!

Simulations – Deterministic v.s. Stochastic

- **MA-IDM assumes:** $a_d^{(t)} \approx a_{\text{IDM},d}^{(t)} + a_{\text{GP},d}^{(t)}$
- **Stochastic simulation for step t_0 :**
 - 1) Obtain the first term $a_{\text{IDM},d}^{(t)}$ by feeding θ_d and inputs into the IDM function;
 - 2) Draw a sample $a_{\text{GP},d}^{(t)} | \mathbf{a}_{\text{GP},d}^{(t-T:t-1)}$ at time t from the GP to obtain the temporally correlated information $a_{\text{GP},d}^{(t)}$;

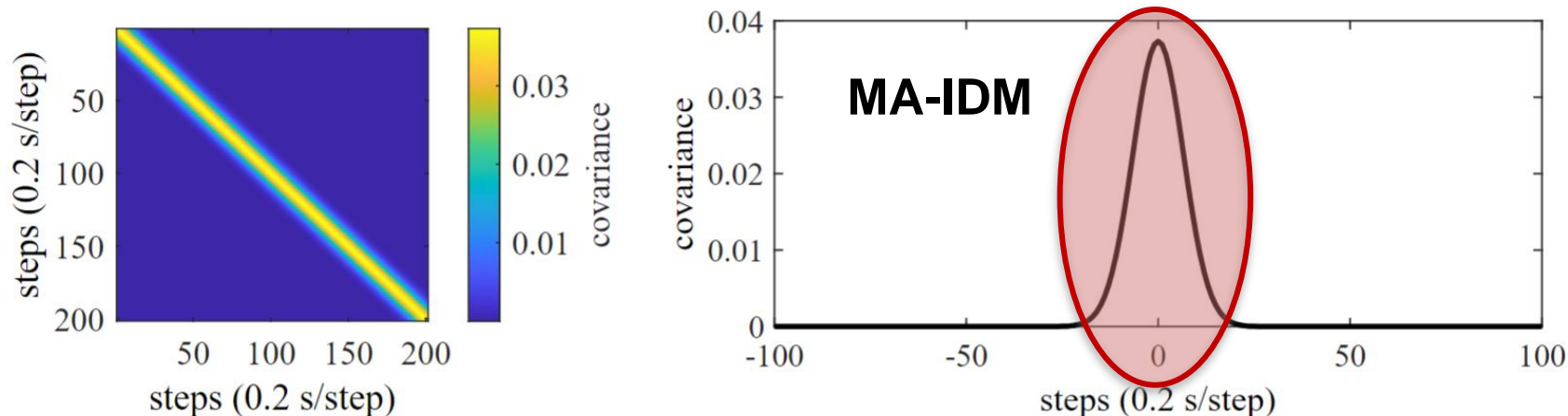
Simulations – Stochastic Simulation (**MA-IDM** v.s. **B-IDM**)



Brief results:

- **Action uncertainty is scenario specific:** When the leading vehicle is braking, all drivers must decelerate; But when the leading vehicle accelerates, actions are more uncertain at their own will.
- **MA-IDM** has a better calibration result than **B-IDM**. Even **B-IDM** is with a large noise variance, it still cannot bridge the **gap** (*i.e.*, with bad uncertainty quantification.)

Stationary kernel and nonstationary behaviors



The lengthscale is about 1.5 sec → **capture correlations within 4~5 sec**
(3-sigma in Normal distribution).

But it assumes that the temporal correlations are **stationary**.

- Stationary temporal correlations: The correlations between time steps are assumed to be **constant** over time.
- Human driving behavior is **dynamic**. Drivers might react differently under congested traffic conditions compared to open road driving.

Nonstationary temporal correlations (Zhang et al. 2025)

- We assume:**

$$a(x, t) = \boxed{f_{\text{CFM}}(x; \theta)} + \boxed{\delta(t)} + \boxed{\epsilon_t}, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

- MA-IDM assumes:**

$$a^{(t)} = \boxed{a_{\text{IDM}}^{(t)}} + \boxed{a_{\text{GP}}^{(t)}} + \boxed{\epsilon_t} \quad \Rightarrow \quad a|i, \theta \sim \mathcal{N}(\boxed{a_{\text{IDM}}}, \boxed{K} + \boxed{\sigma_\epsilon^2 I})$$

Homoscedasticity assumption with a stationary kernel. (inappropriate)

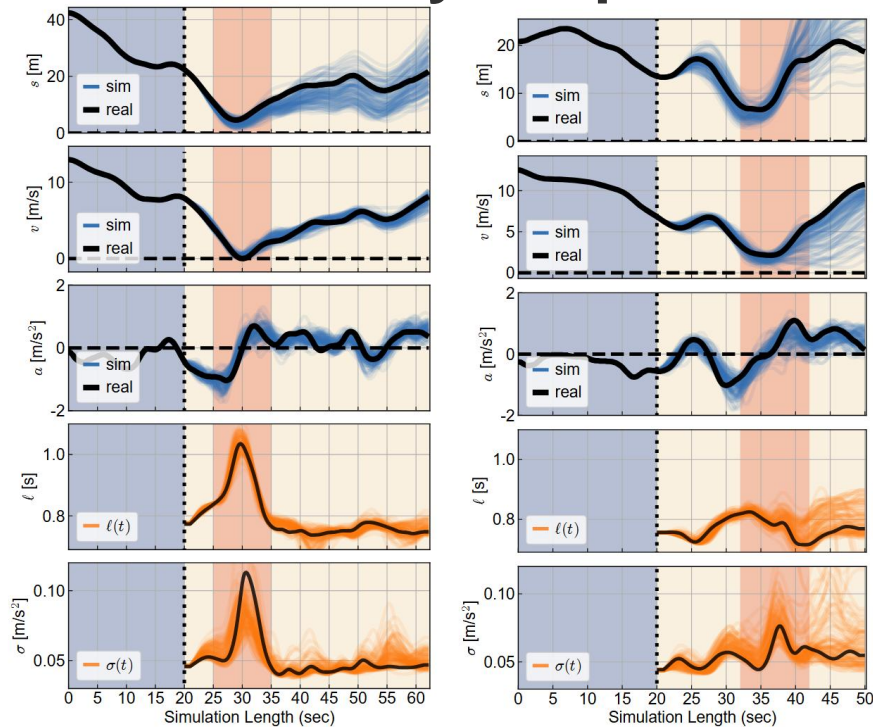
- Nonstationary model assumes:**

$$a^{(t)} = \boxed{a_{\text{NN}}^{(t)}} + \boxed{a_{\text{GP}}^{(t)}} + \boxed{\epsilon_t} \quad \Rightarrow \quad a|i, \theta_{\text{NN}} \sim \mathcal{N}(\boxed{a_{\text{NN}}}, \boxed{K} + \boxed{\sigma_\epsilon^2 I})$$

Heteroscedasticity assumption with a nonstationary kernel (Gibbs kernel)

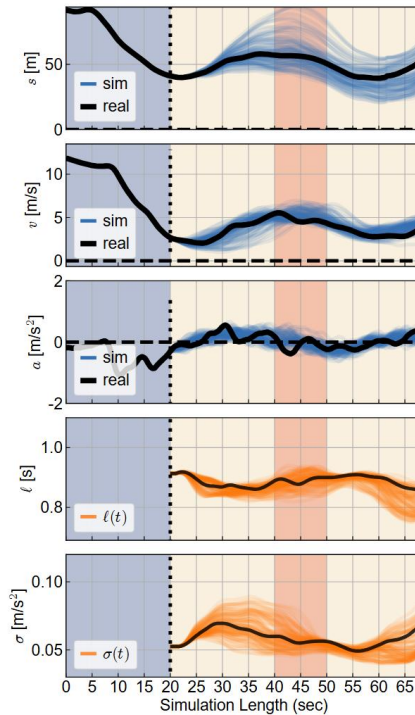
$$k_{\text{Gibbs}}(t, t'; \lambda) := \sigma(t)\sigma(t') \sqrt{\frac{2\ell(t)\ell(t')}{\ell(t)^2 + \ell(t')^2}} \exp\left(-\frac{(t - t')^2}{\ell(t)^2 + \ell(t')^2}\right)$$

Nonstationary temporal correlations (Zhang et al. 2025)



(a) Car-following Pair #1.

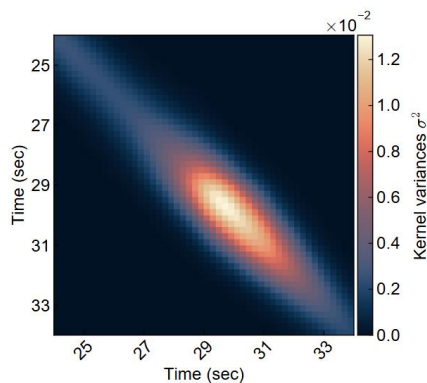
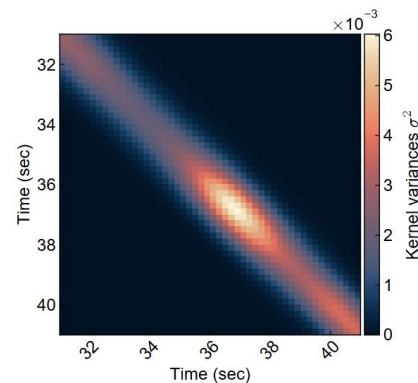
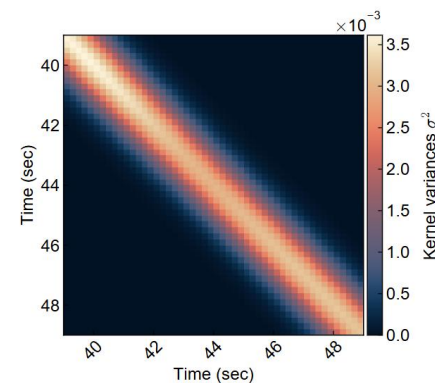
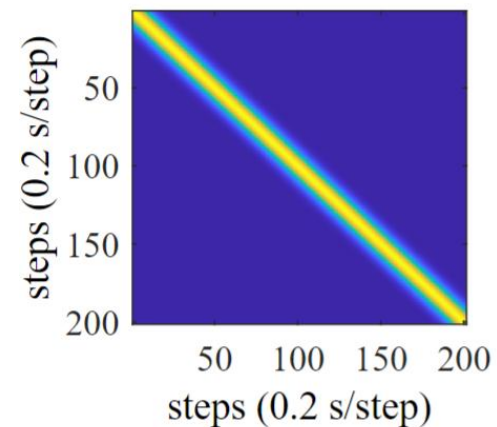
(b) Car-following Pair #2.



(c) Car-following Pair #3.

Lengthscale:
 Smooth driving \uparrow
 Abrupt transition \downarrow

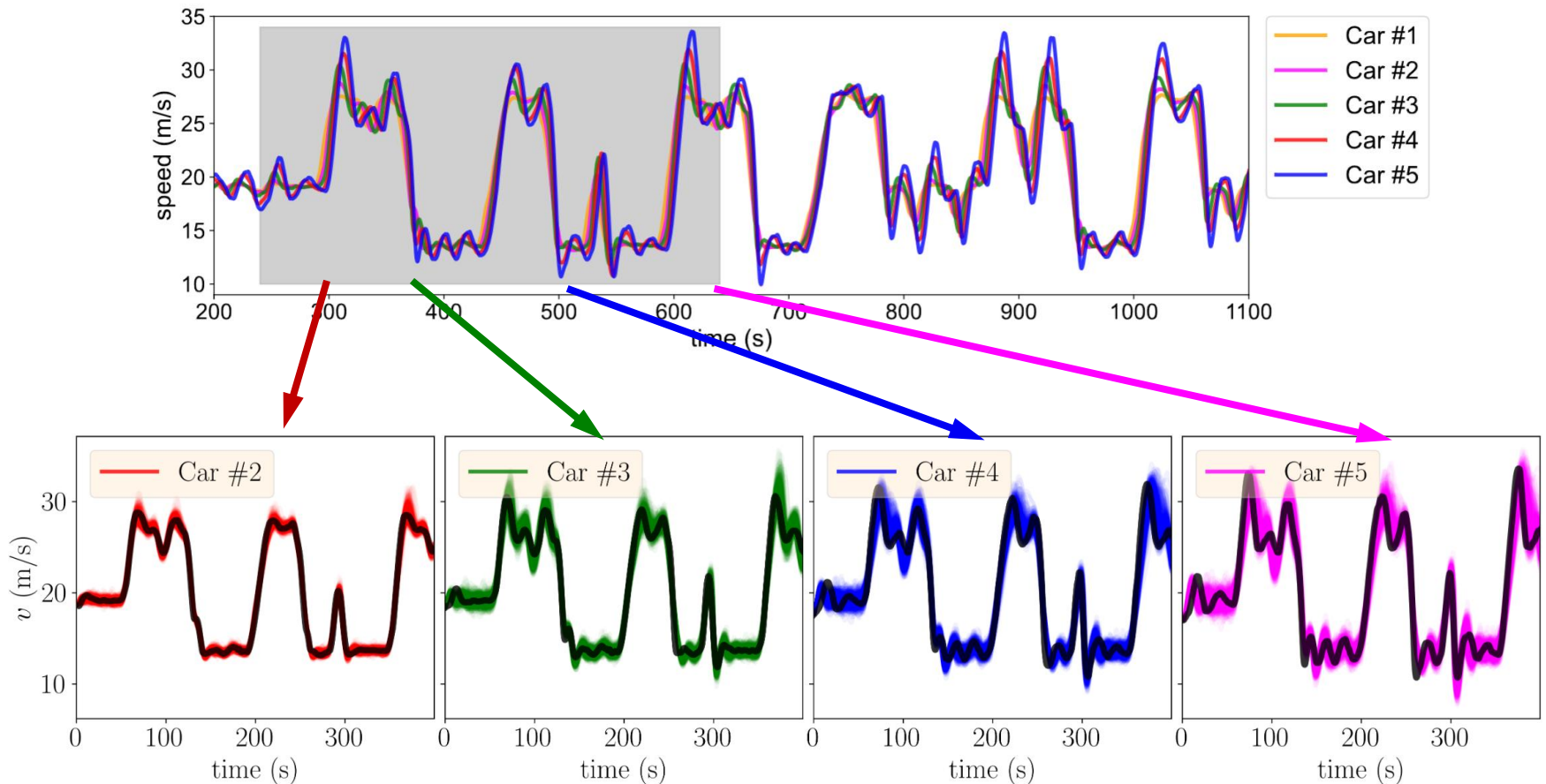
Kernel variance:
 Free/steady \uparrow
 Safety-critical \downarrow

(d) Example \mathbf{K} in Pair #1.(e) Example \mathbf{K} in Pair #2.(f) Example \mathbf{K} in Pair #3.

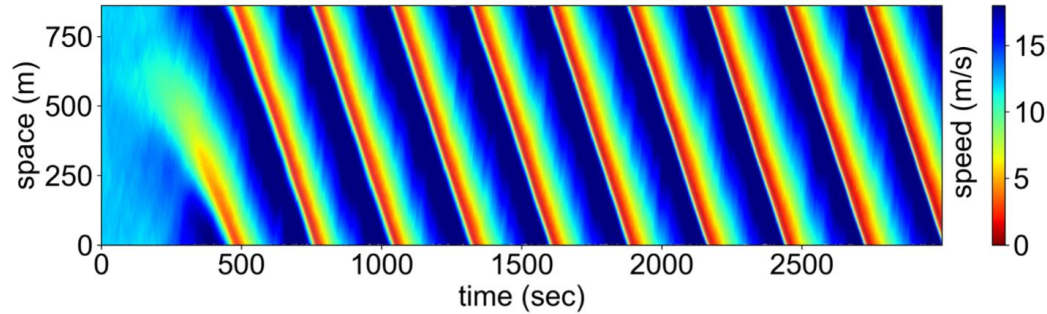
Simulations – Multi-vehicle scenario: Platoon

OpenACC dataset

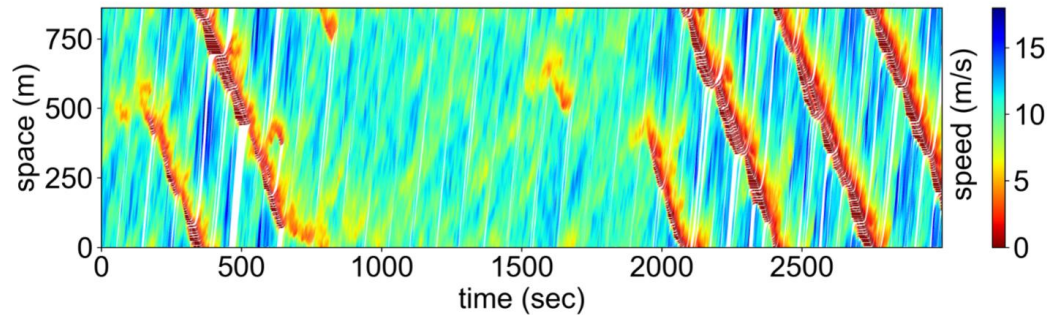
<http://data.europa.eu/89h/9702c950-c80f-4d2f-982f-44d06ea0009f>



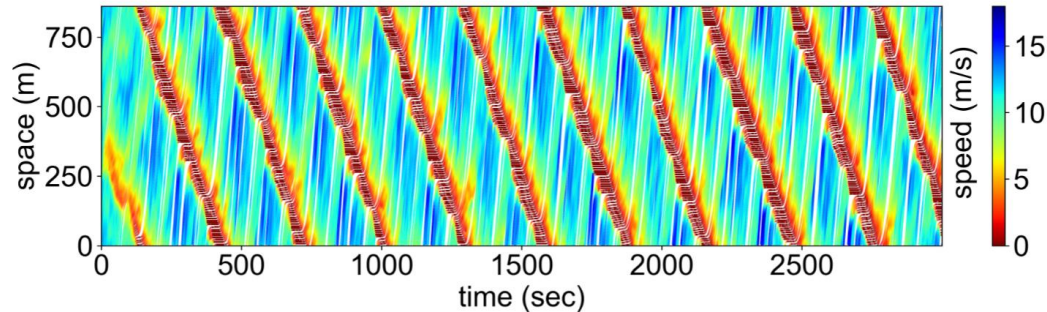
Simulations – Multi-vehicle scenario: Ring road



(a) Simulation with fixed IDM parameters and random white noise.



(b) Light traffic simulation with dynamic IDM ($p = 4$).



(c) Dense traffic simulation with dynamic IDM ($p = 4$).



Sugiyama experiment

General Overview

$$a(x, t) = f_{\text{CFM}}(x; \boldsymbol{\theta}) + \delta(t) + \epsilon_t, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

Table 1: Modeling of temporal correlations in the literature.

Reference	$f_{\text{CFM}}(\mathbf{x})$	$\delta(t)$	Nonstationary?
Treiber et al. (2006)	IDM	Ornstein-Uhlenbeck (OU) processes	✗
Hoogendoorn and Hoogendoorn (2010)	GHR/IDM	Cochrane-Orcutt correction (i.e., AR(1) process)	✗
Zhang and Sun (2024)	IDM	Gaussian processes (GPs)	✗
Zhang et al. (2024a)	IDM	AR processes with higher orders	✗
Zhang et al. (2025)	NN	nonstationary GPs	✓

- The **GP with a Matérn 1/2 kernel** can be seen as the continuous-time counterpart of the discrete-time **AR(1)** process; The **AR(1)** process can be seen as a discrete-time analog of the **Matérn 1/2 kernel**.
- The **OU process** is a continuous-time counterpart to the **AR(1)** process. The **OU process** is equivalent to the **GP with a Matérn 1/2 kernel**.
- The **AR(1)** process can be considered a discrete-time version of the **OU process**. Both processes have exponential autocorrelation functions, but the **AR(1)** process is defined in discrete time, while the **OU process** is defined in continuous time.
- the **Cochrane-Orcutt correction** is a method for addressing autocorrelation in regression models, assuming an **AR(1)** structure

Discussion and takeaway

Human-like car-following behaviors:

$$a(x, t) = \boxed{f_{\text{CFM}}(x; \theta)} + \boxed{\delta(t)} + \boxed{\epsilon_t}, \quad \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

- **Modeling: Significance of appropriate uncertainty quantification!**
- **Simulation: Significance of stochastic simulation!**
- **Inappropriate assumptions and solutions:**
 - Independent and identically distributed (i.i.d.) errors; (GP/AR)
 - Homoscedasticity (constant variance of errors); (nonstationary GP)

References

- Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2), 1805.
- Punzo, V., Zheng, Z., & Montanino, M. (2021). About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes. *Transportation Research Part C: Emerging Technologies*, 128, 103165.
- Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). 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) (pp. 2118-2125). IEEE.
- Anesiadou, A., Makridis, M., Ciuffo, B., & Mattas, K. (2020): Open ACC Database. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/9702c950-c80f-4d2f-982f-44d06ea0009f>
- Treiber, M., Kesting, A., & Helbing, D. (2006). Delays, inaccuracies and anticipation in microscopic traffic models. *Physica A: Statistical Mechanics and its Applications*, 360(1), 71-88.

Read More

➤ Paper:

Zhang, C., & Sun, L. (2024). Bayesian calibration of the intelligent driver model. *IEEE Transactions on Intelligent Transportation Systems*. (IDM with GP)

Zhang, C., Wang, W., & Sun, L. (2024). Calibrating car-following models via Bayesian dynamic regression. *Transportation Research Part C: Emerging Technologies*, 104719. (ISTTT25 Special Issue). (IDM with AR)

Zhang, C., Zheng, H., Wu, C., & Sun, L. (2025). Stochastic Modeling of Car-Following Behaviors with Nonstationary Temporal Correlations. *Preprint (under review)*. (NN with nonstationary GP)

➤ Code:

https://github.com/Chengyuan-Zhang/IDM_Bayesian_Calibration



Thanks! Questions?

Chengyuan Zhang (enzozcy@gmail.com)

Ph.D. Candidate (advised by Prof. Lijun Sun)

Department of Civil Engineering

McGill University

<https://chengyuan-zhang.github.io/>



LinkedIn

Feb. 13, 2025

JTL Research Seminar @ MIT (online)



McGill