Stochastic Modeling and Simulations of Car-Following Behaviors

Chengyuan Zhang
Ph.D. Candidate (advised by Prof. Lijun Sun)
Department of Civil Engineering
McGill University
enzozcy@gmail.com

Feb. 06, 2025
JTL Research Seminar @ MIT (online)

About me

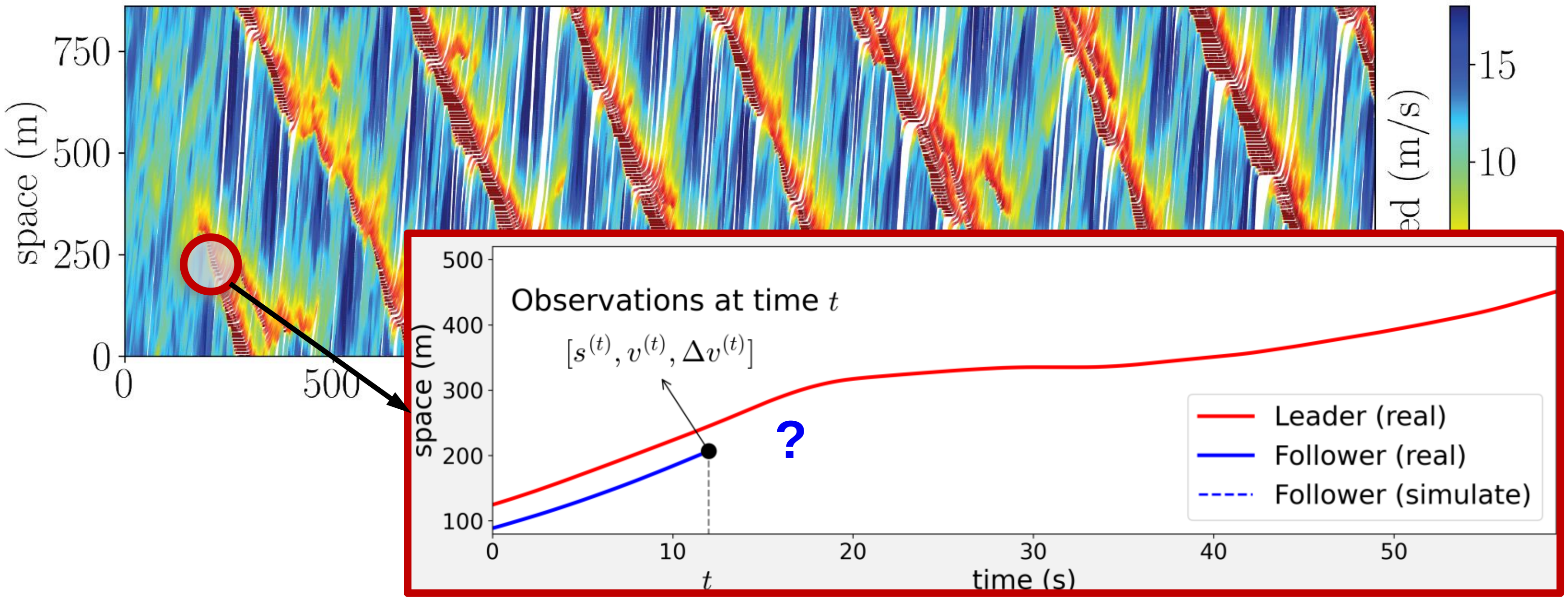
- Smart Transportation Lab at McGill University

- Homepage: https://chengyuan-zhang.github.io/

- My research focuses on **Bayesian learning, spatiotemporal modeling, traffic flow theory,** and **multi-agent interaction modeling** within intelligent transportation systems. 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.

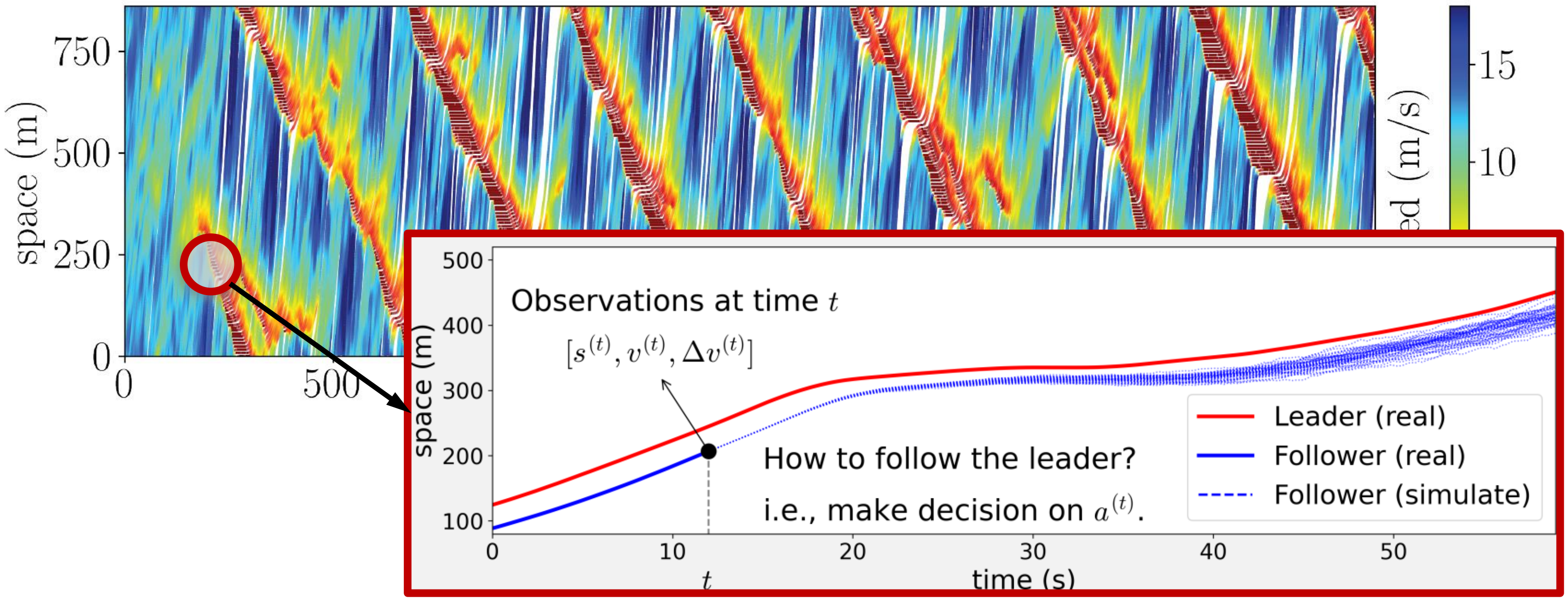
Motivation / background

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



Motivation / background

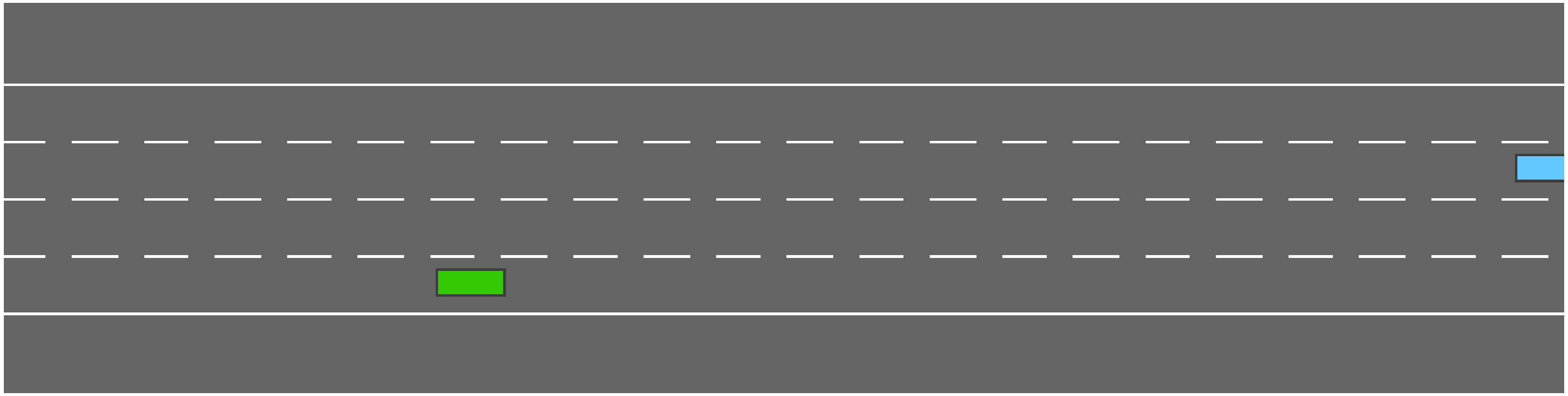
- How would the vehicle react in response to the leading vehicle?



- What do we need for 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 randomness?
  - x Deterministic car-following models
  - ✓ **Probabilistic car-following models with uncertainty quantification**

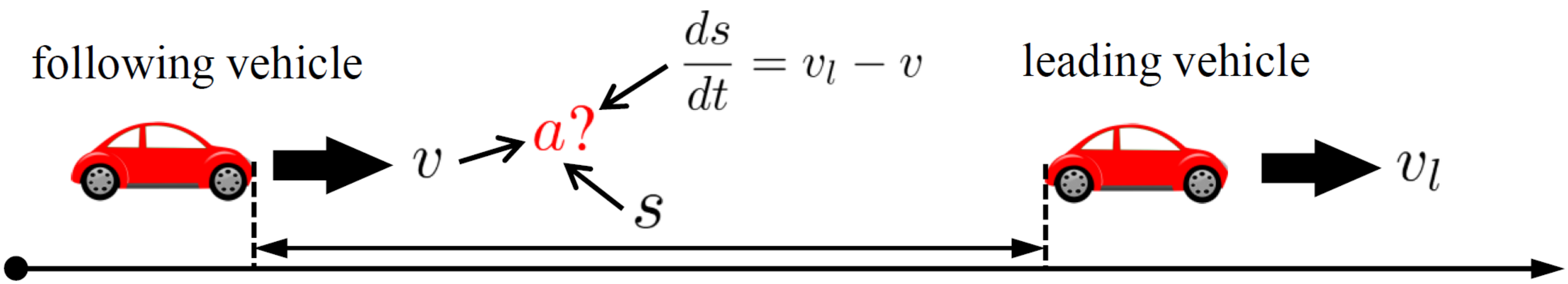
In this work, we are interested in:

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

Outline

- Intelligent Driver Model (IDM, as an example)
- Probabilistic Modeling Framework
    - Mean model (IDM/NN) + stochastic processes (GP/AR)
- Stochastic Simulation
- Discussions

Intelligent driver model



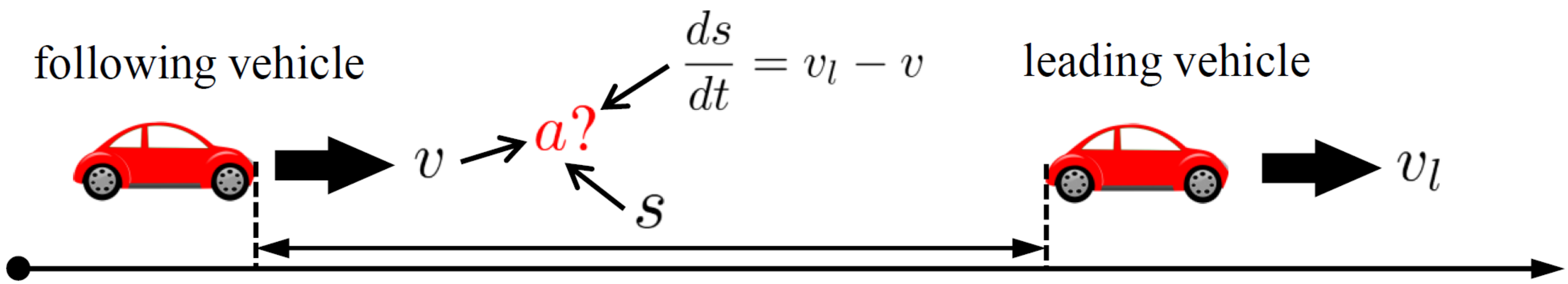
- **Intelligent Driver Model (IDM)** (Treiber et al. 2000)

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

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

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

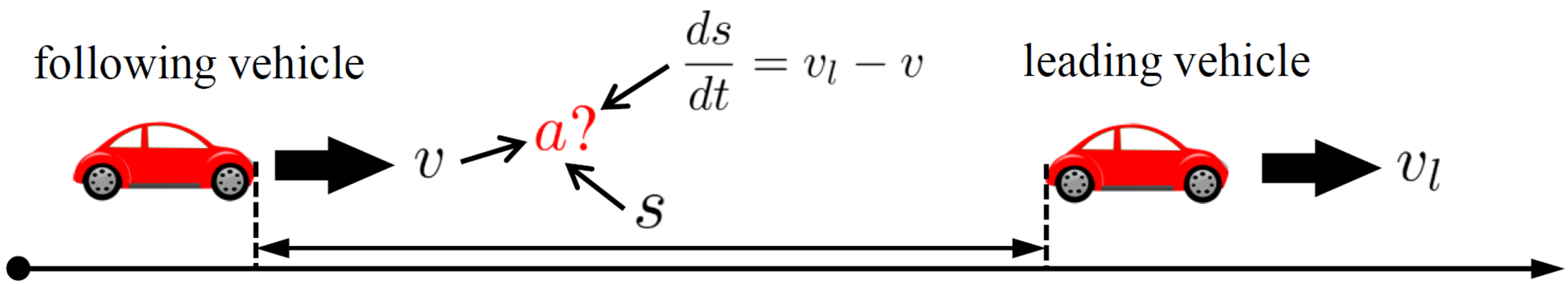
Intelligent driver model



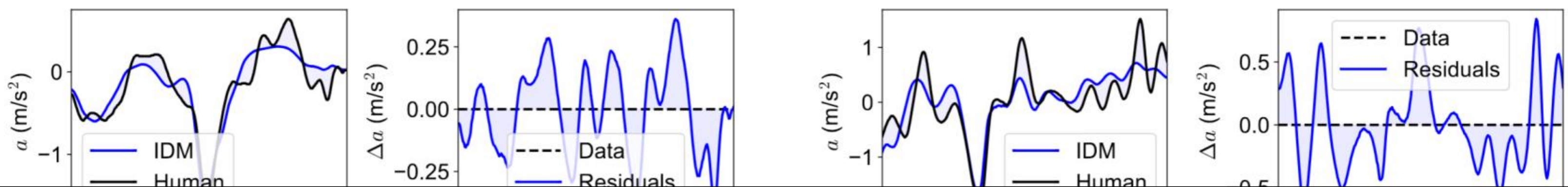
- Gaussian IDM assumes: $a^{(t)} \approx a_{\mathrm{IDM}}^{(t)}$. ➡ Likelihood: $\mathcal{N}(\hat{a}^{(t)} | a_{\mathrm{IDM}}^{(t)}, \sigma_\epsilon^2)$
- Calibration by MLE: $\max_{\boldsymbol{\theta}} \prod_{t=1}^{T} \text{likelihood}$ and $\boldsymbol{\theta} = [v_0, s_0, T, \alpha, \beta]$
- Loss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

$$\min_{\boldsymbol{\theta}} \frac{1}{T} \sum_{t=1}^{T} (a_{\mathrm{IDM}}^{(t)} - \hat{a}^{(t)})^2 + \frac{\alpha}{T} \sum_{t=1}^{T} (v_{\mathrm{IDM}}^{(t)} - \hat{v}^{(t)})^2 + \frac{\beta}{T} \sum_{t=1}^{T} (x_{\mathrm{IDM}}^{(t)} - \hat{x}^{(t)})^2$$

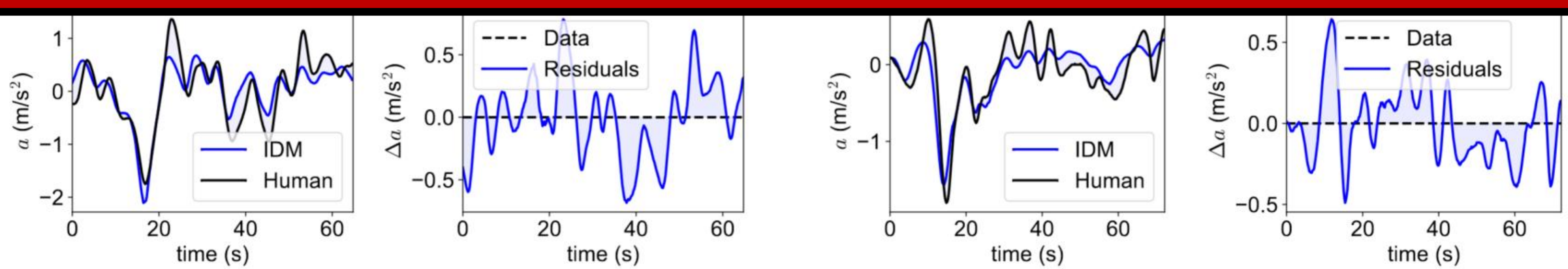
Intelligent driver model



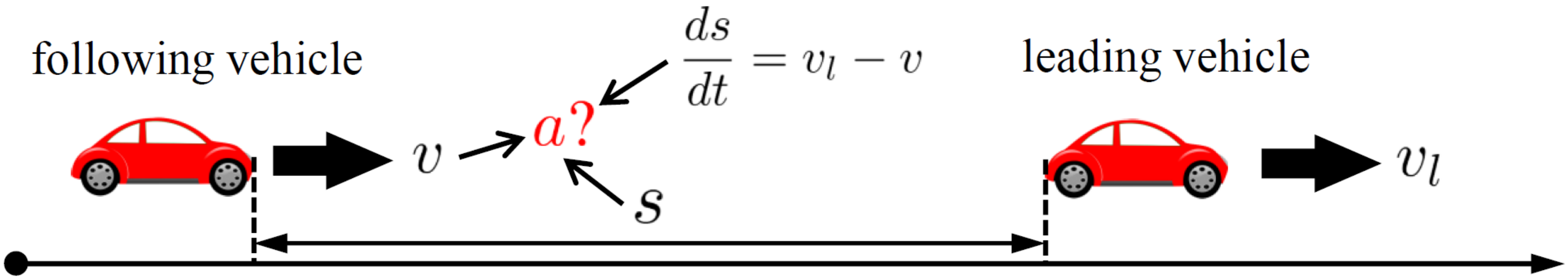
- IDM assumes: $a^{(t)} \approx a^{(t)}_{\text{IDM}}$. Let's visualize the residuals $a^{(t)} - a^{(t)}_{\text{IDM}}$.



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



The general form of car-following models

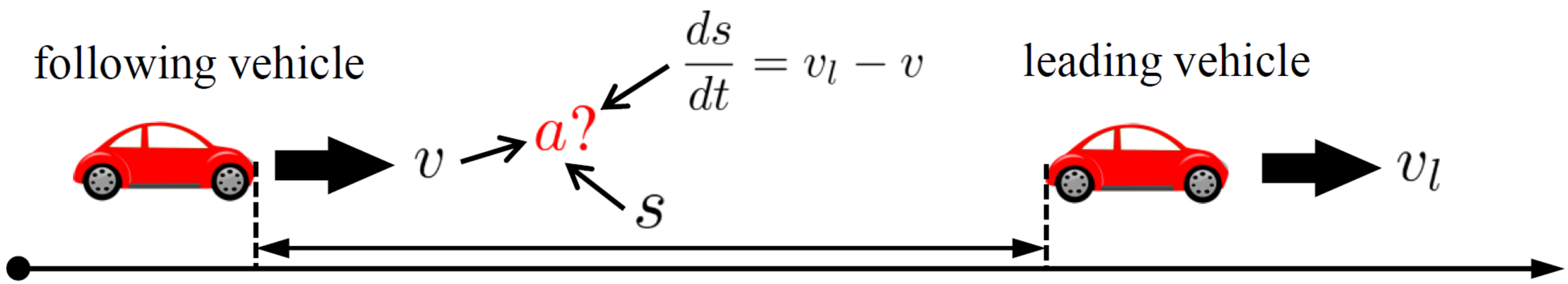


- For a human-driver CF model, what do we miss?
  - **Temporally correlated errors / Time delay** (e.g., Wiener process)
    - IDM as a **parsimonious** model can hardly explain all the variation in the data; as a result, the residual terms are **serially correlated**;
  - …

- Therefore, we assume $a(x,t) \approx \underbrace{\boxed{f_{\mathrm{CFM}}(x;\boldsymbol{\theta})}}_{\text{mean car-following model}} + \overbrace{\boxed{\delta(t)}}^{\text{temporal correlations}}$!!! (Zhang et al. 2025)

Inspired by GLS, see my post From Ordinary Least Squares (OLS) to Generalized Least Squares (GLS)

The general form of car-following models



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

<u>Missed the temporal part</u> $\boxed{\delta(t)}$

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:**

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

- **IDM assumes:**

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

Vector form

- **MA-IDM assumes:**

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

residuals

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:**

$$a^{(t)} = a_{\mathrm{IDM}}^{(t)} + \epsilon_t, \;\; \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\epsilon^2) \qquad \Rightarrow \boldsymbol{a}|\boldsymbol{i}, \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{a}_{\mathrm{IDM}}, \sigma_\epsilon^2 \boldsymbol{I})$$

Vector form

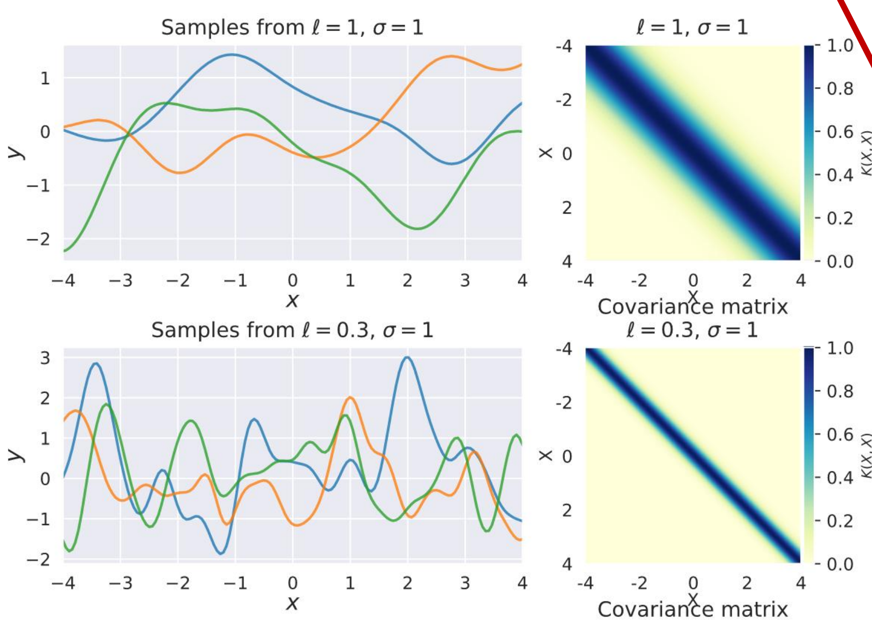
- **MA-IDM assumes:**

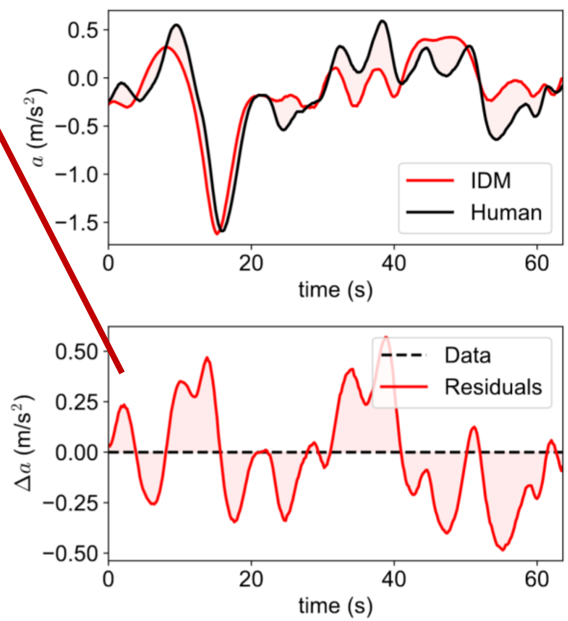
$$a^{(t)} = a_{\mathrm{IDM}}^{(t)} + \boxed{a_{\mathrm{GP}}^{(t)}} \qquad \Rightarrow \boldsymbol{a}|\boldsymbol{i}, \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{a}_{\mathrm{IDM}}, \boxed{\boldsymbol{K}} + \sigma_\epsilon^2 \boldsymbol{I})$$

residuals

where ***K*** is a kernel matrix .

- Gaussian processes



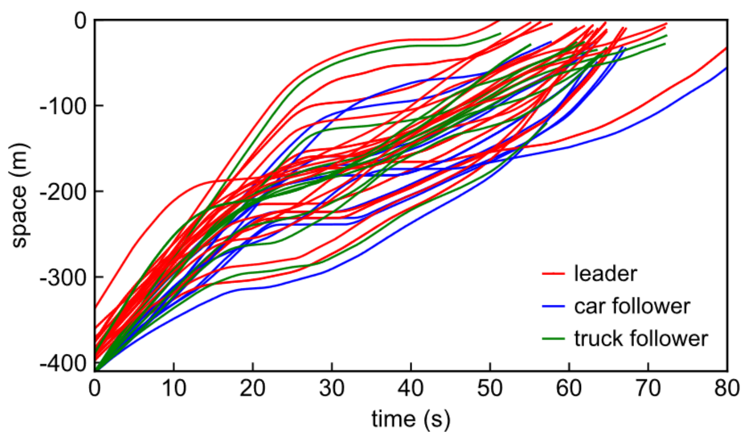


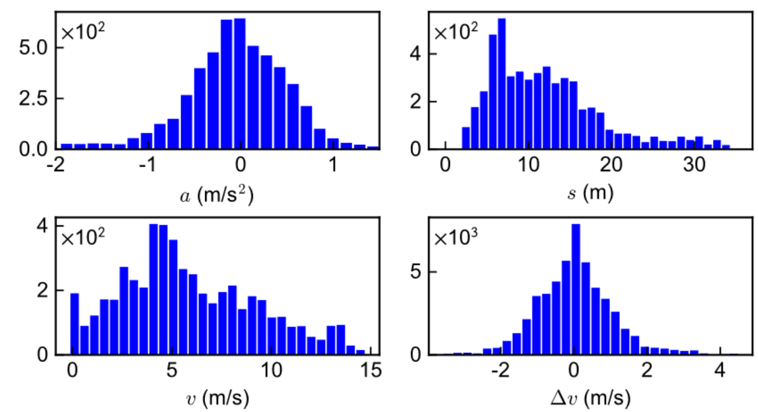
Experiments – Car-Following Data Extraction

- **HighD** dataset: (Krajewski et al. 2018)

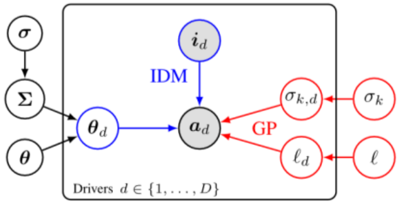
https://levelxdata.com/highd-dataset/

- **20** leader-follower pairs.

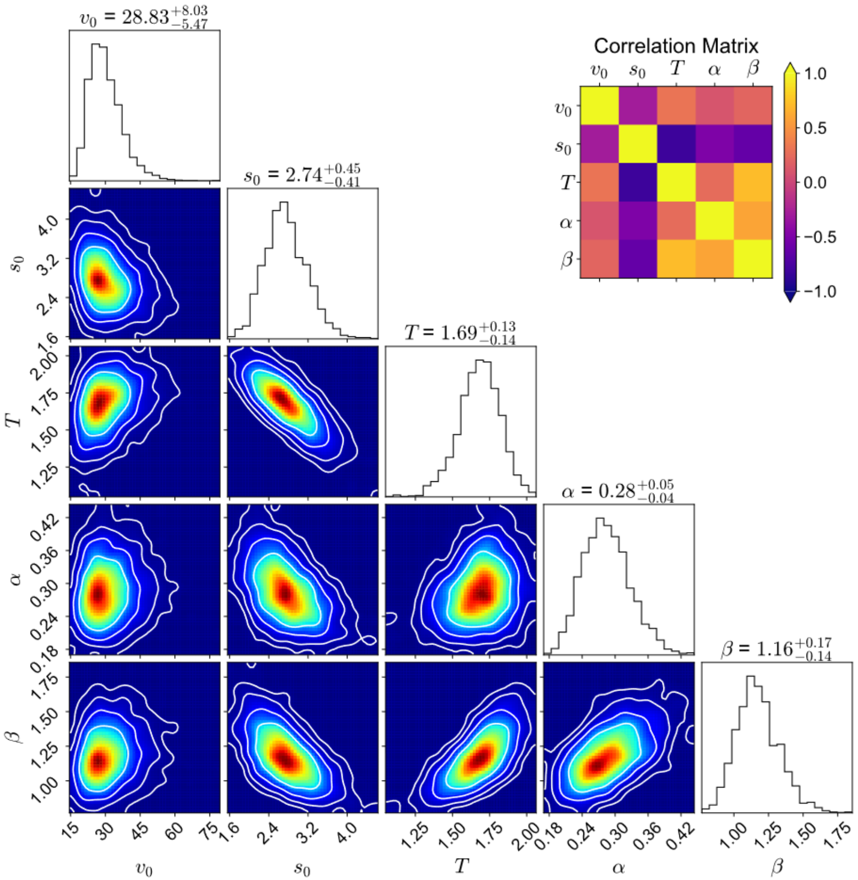




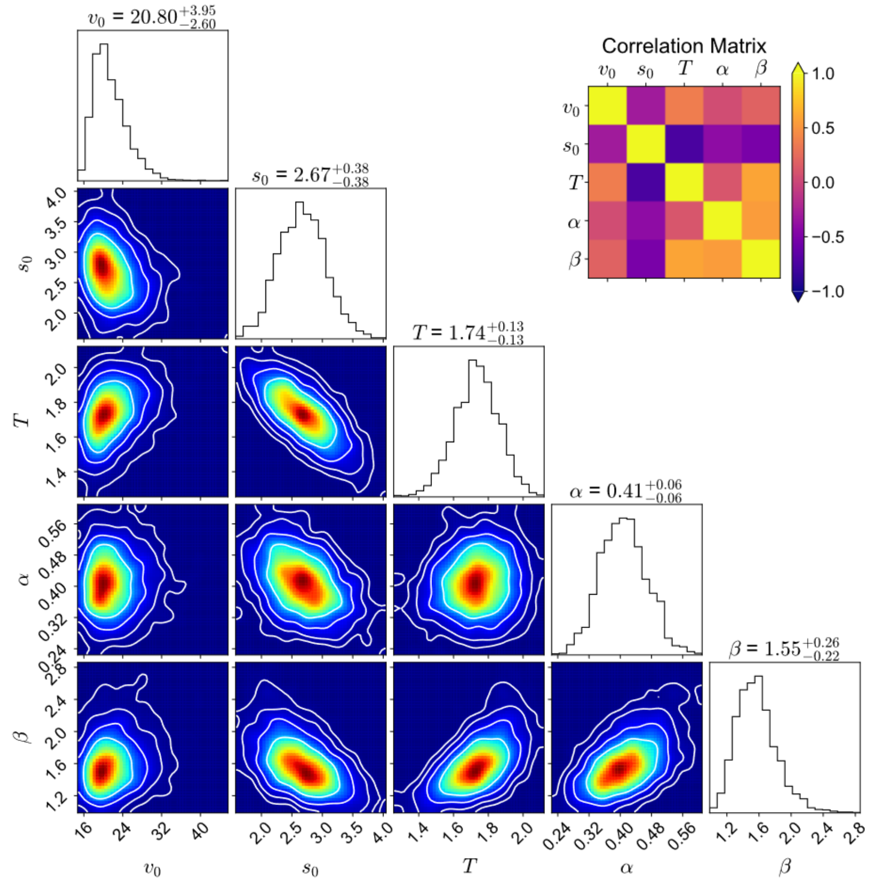
Experiments – Identified IDM Parameters



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!!

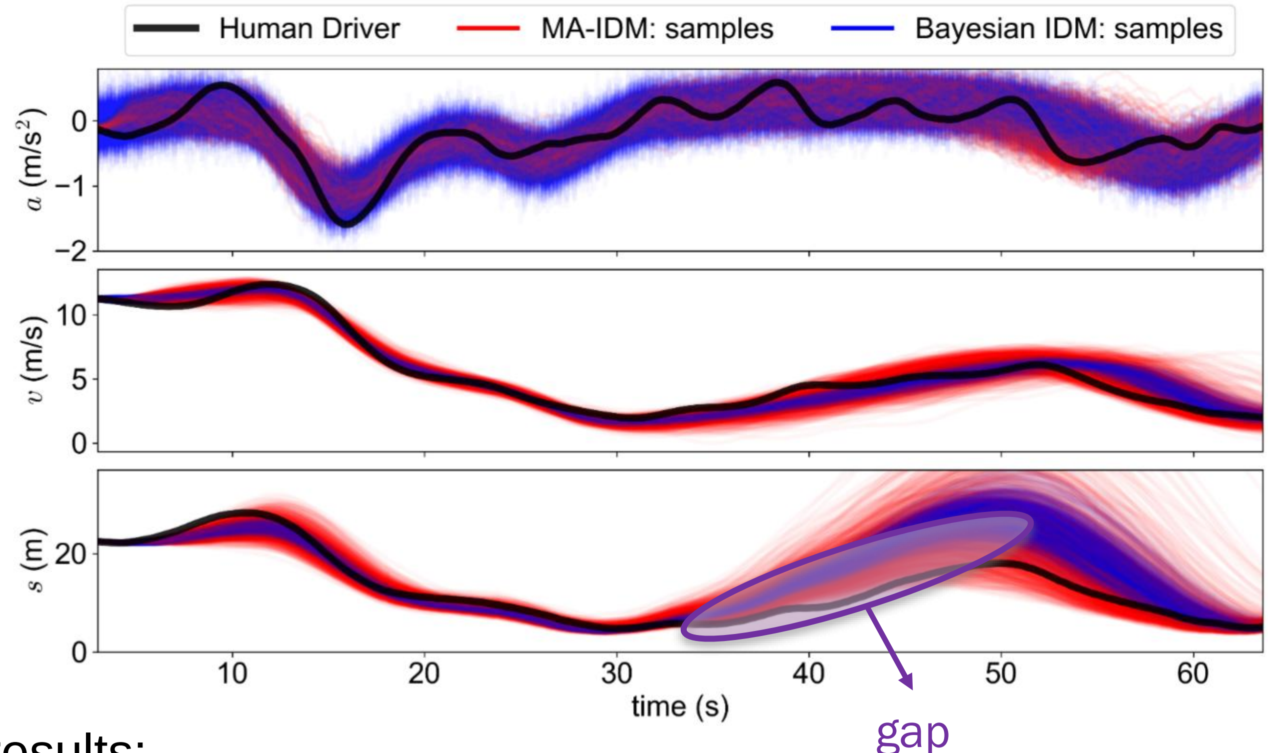
Simulations – Deterministic *v.s.* Stochastic

- **MA-IDM assumes:** $a_d^{(t)} = \boxed{a_{\text{IDM } d}^{(t)}} + \boxed{a_{\text{GP } d}^{(t)}}$

- **Stochastic simulation for step t_0:**

  1) Obtain the first term $\boxed{a_{\text{IDM},d}^{(t)}}$ by feeding $\boldsymbol{\theta}_d$ and inputs into the IDM function;
  2) Draw a sample $\boxed{a_{\text{GP},d}^{(t)} | \boldsymbol{a}_{\text{GP},d}^{(t-T:t-1)}}$ at time t from the GP to obtain the temporally correlated information $a_{\text{GP},d}^{(t)}$;

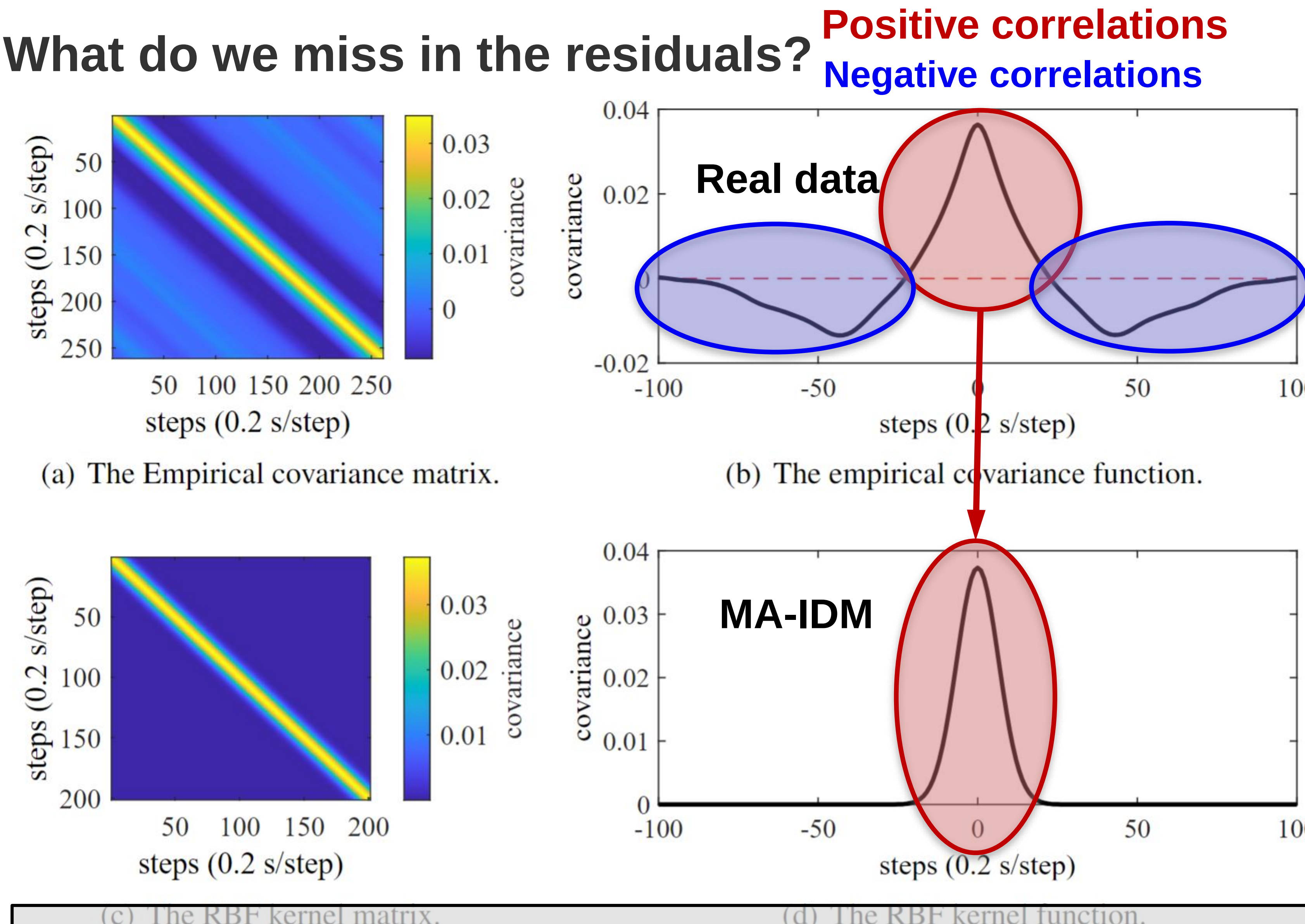
Simulations – Stochastic Simulation (MA-IDM v.s. B-IDM)



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.)
- But MA-IDM has a much larger variance. **Why?**

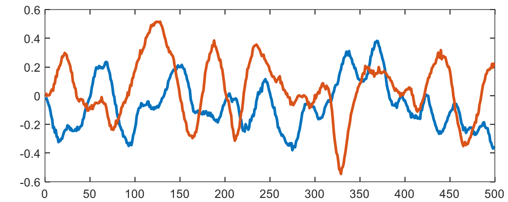
What do we miss in the residuals?



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

Dynamic IDM (Zhang et al. 2024)

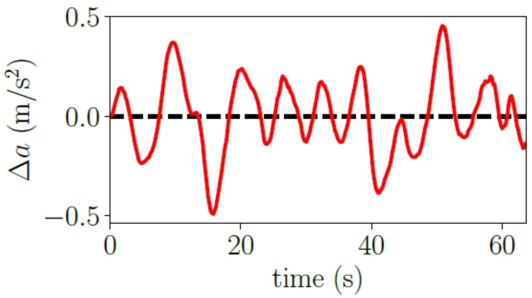
How to model $\delta(t)$ and $\delta(t+1)|\delta(t)$?



Two random series generated by AR(4)

- **We assume:**

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



- **Dynamic IDM assumes:**

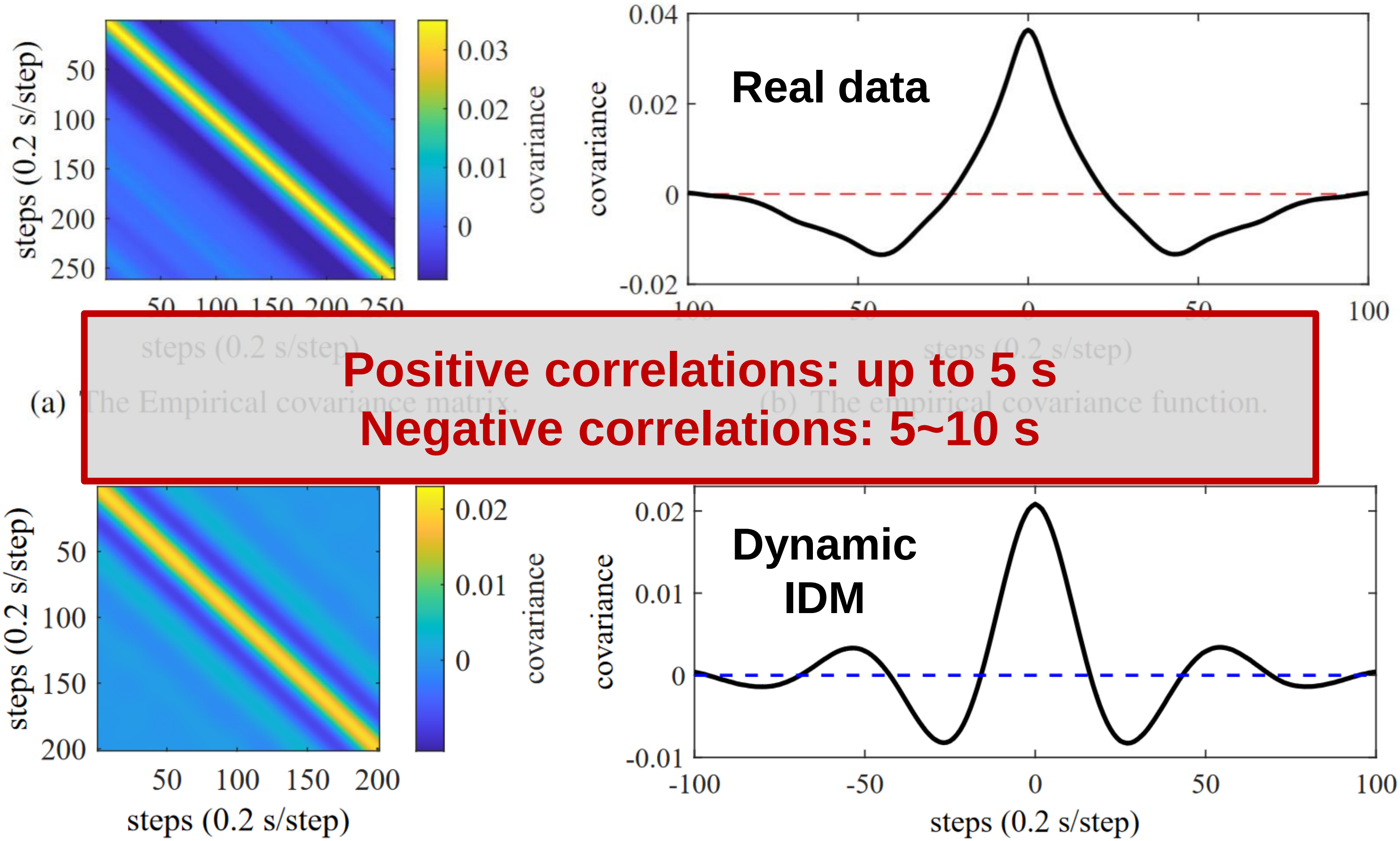
  Autoregressive (AR) processes

$$a_d^{(t)} = \boxed{\mathrm{IDM}_d^{(t)}} + \varepsilon_d^{(t)},$$
$$\varepsilon_d^{(t)} = \boxed{\rho_{d,1}\varepsilon_d^{(t-1)} + \rho_{d,2}\varepsilon_d^{(t-2)} + \cdots + \rho_{d,p}\varepsilon_d^{(t-p)}} + \boxed{\eta_d^{(t)},}$$
$$\eta^{(t)} \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma_\eta^2).$$

ADVANTAGE: It involves rich information from **several historical steps** instead of using only one step.

[Chengyuan Zhang, Wenshuo Wang, and Lijun Sun. (2024). Calibrating car-following models via Bayesian dynamic regression. (ISTTT Special Issue) Transportation Research Part C: Emerging Technologies, 104719.]

Experiments – Identified AR Parameters



Positive correlations: up to 5 s
Negative correlations: 5~10 s

(e) The AR(5) covariance matrix.

(f) The AR(5) covariance functions.

$$\boldsymbol{\rho} = [0.874, 0.580, -0.105, -0.315, -0.071]$$

Simulations – Deterministic *v.s.* Stochastic

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

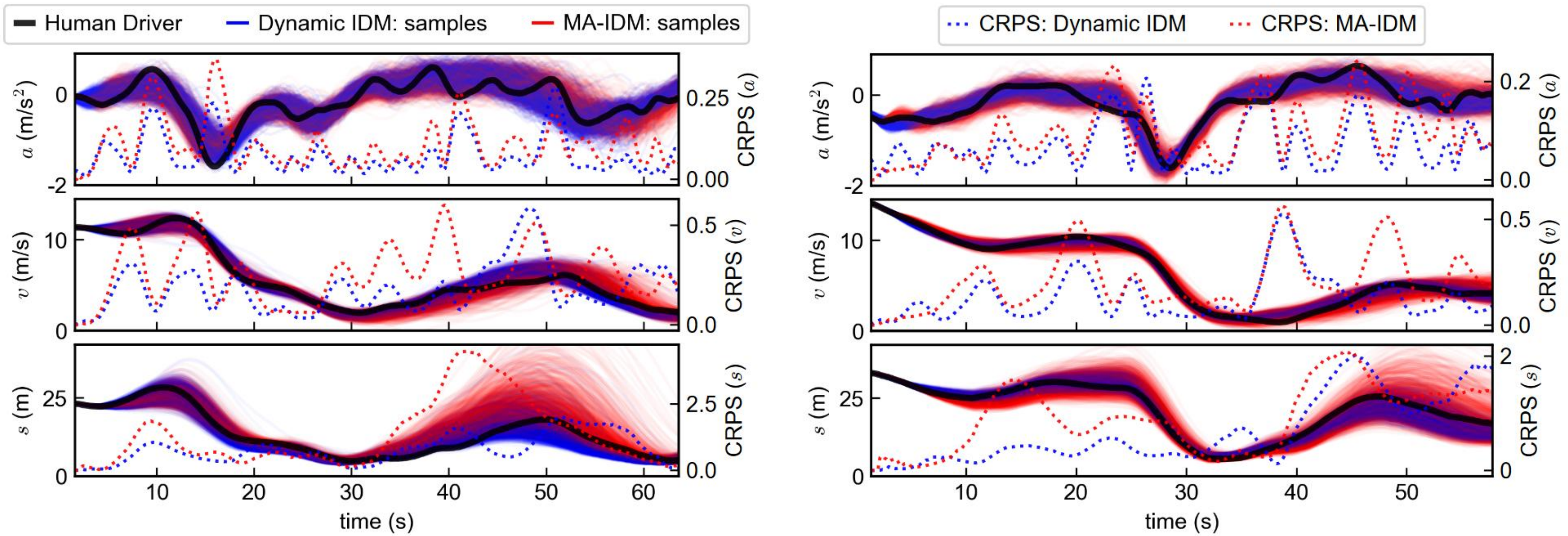
- **Dynamic IDM:**

$$
\begin{aligned}
a_d^{(t)} &= \boxed{\mathrm{IDM}_d^{(t)}} + \varepsilon_d^{(t)}, \\
\varepsilon_d^{(t)} &= \boxed{\rho_{d,1}\varepsilon_d^{(t-1)} + \rho_{d,2}\varepsilon_d^{(t-2)} + \cdots + \rho_{d,p}\varepsilon_d^{(t-p)}} + \boxed{\eta_d^{(t)}}, \\
\eta^{(t)} &\overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\eta^2).
\end{aligned}
$$

- **Stochastic simulation for step t_0:**

  1. **generating the mean model by sampling a set of IDM parameters;**
  2. **computing the serial correlation term according to the historical information;**
  3. **sampling white noise randomly.**

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



(a) Truck driver #211.　　(b) Car driver #273.

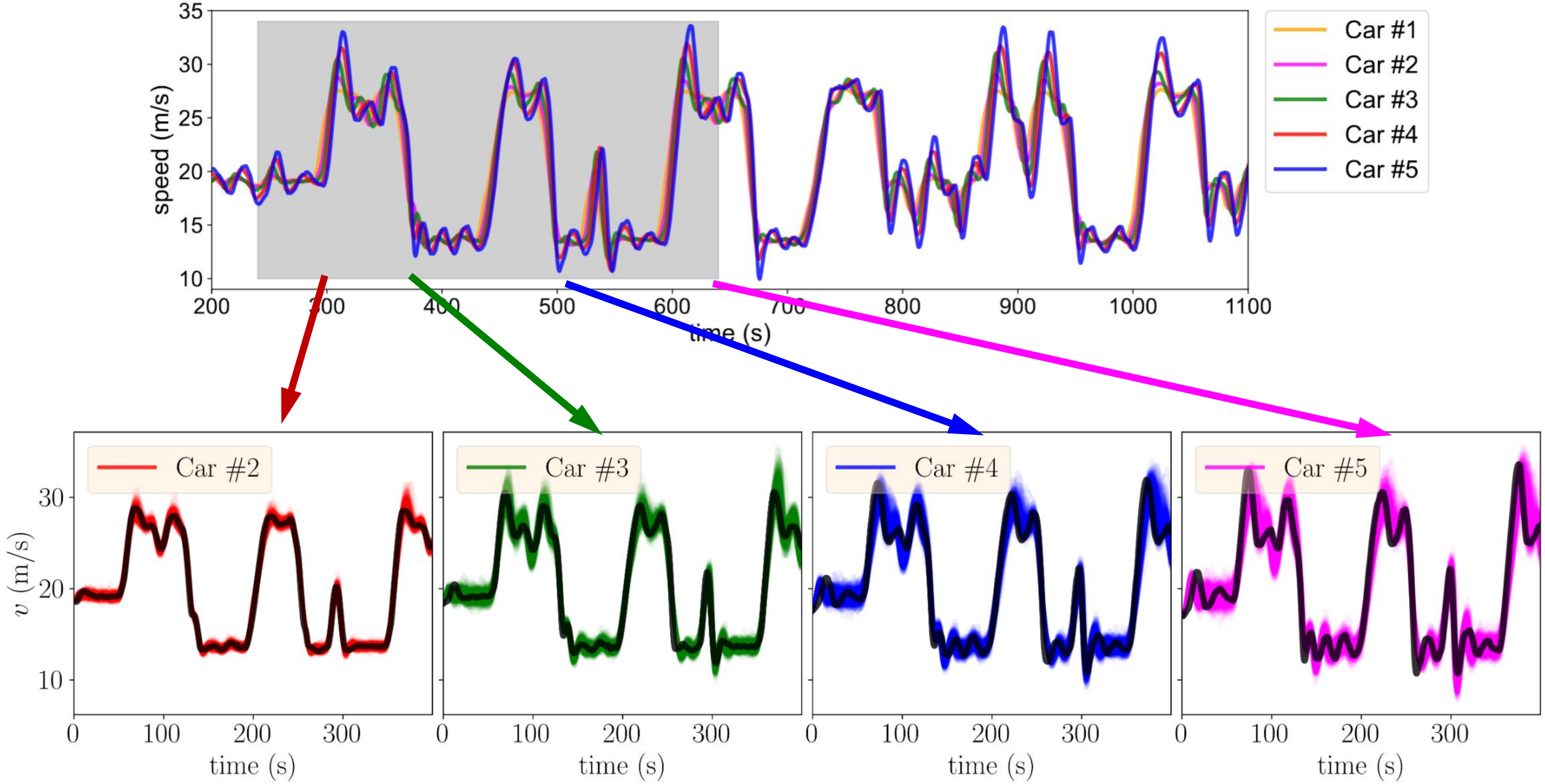
Brief results:

- Dynamic IDM can also obtain accurate parameter estimations;
- Dynamic IDM has much lower variances than MA-IDM;

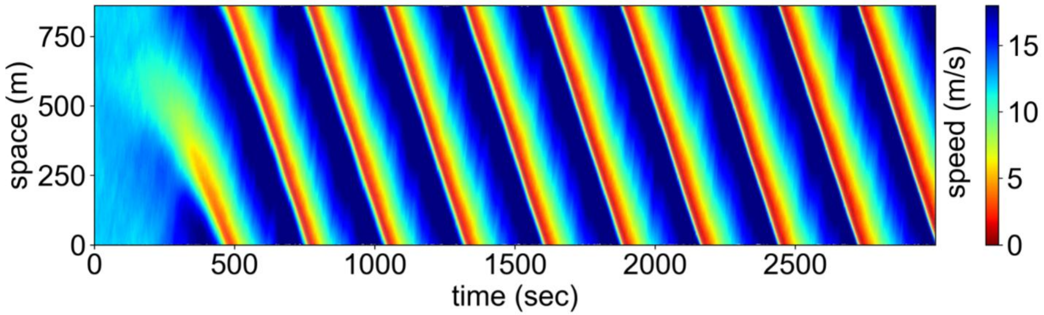
Simulations – Multi-vehicle scenario: Platoon

OpenACC dataset

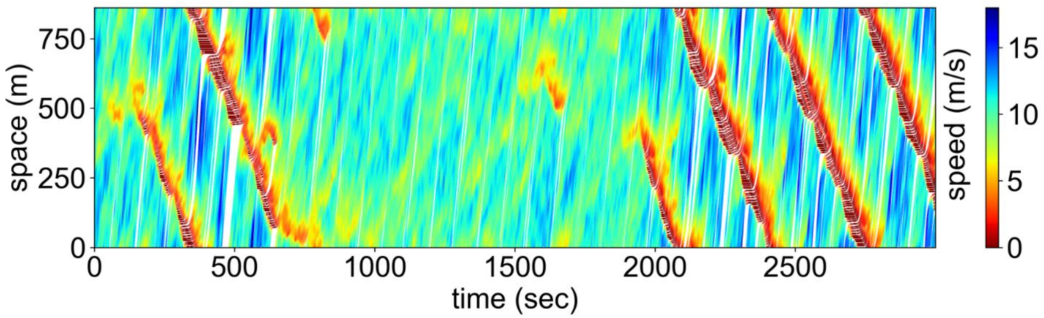
http://data.europa.eu/89h/9702c950-c80f-4d2f-982f-44d06ea0009f



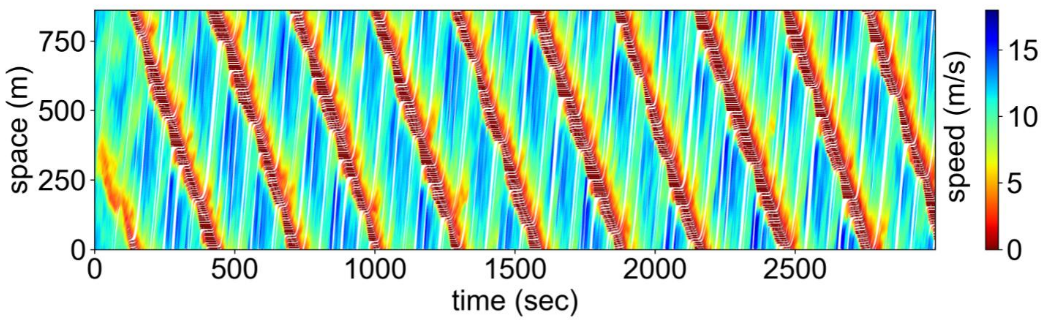
Simulations – Multi-vehicle scenario: Ring road



(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

Discussion and takeaway

- ✓ Generate diverse types of drivers. [Bayesian calibration/Hierarchical structure]
- ✓ Produce good uncertainty for each driver. [GP/AR]
- ✓ Simulate human-like car-following behaviors. [Stochastic Simulation]

- ➢ **Importance of probabilistic simulation!**
- ➢ positive correlations (0~5 sec) & negative correlations (5~10 sec)

  → at least 10 sec historical information as input.
- ➢ Provide enough information to calibrate car-following models.
- ➢ IDM is very powerful.

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*.

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)*.

- **Code:**

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