Harnessing Mean-Field Game & Data Science for Mixed Autonomy

Sharon Di, Ph.D.

Civil Engineering & Engineering Mechanics

Data Science Institute

Columbia University in the City of New York

Data & innovative-technology driven Transportation Lab

Ditech

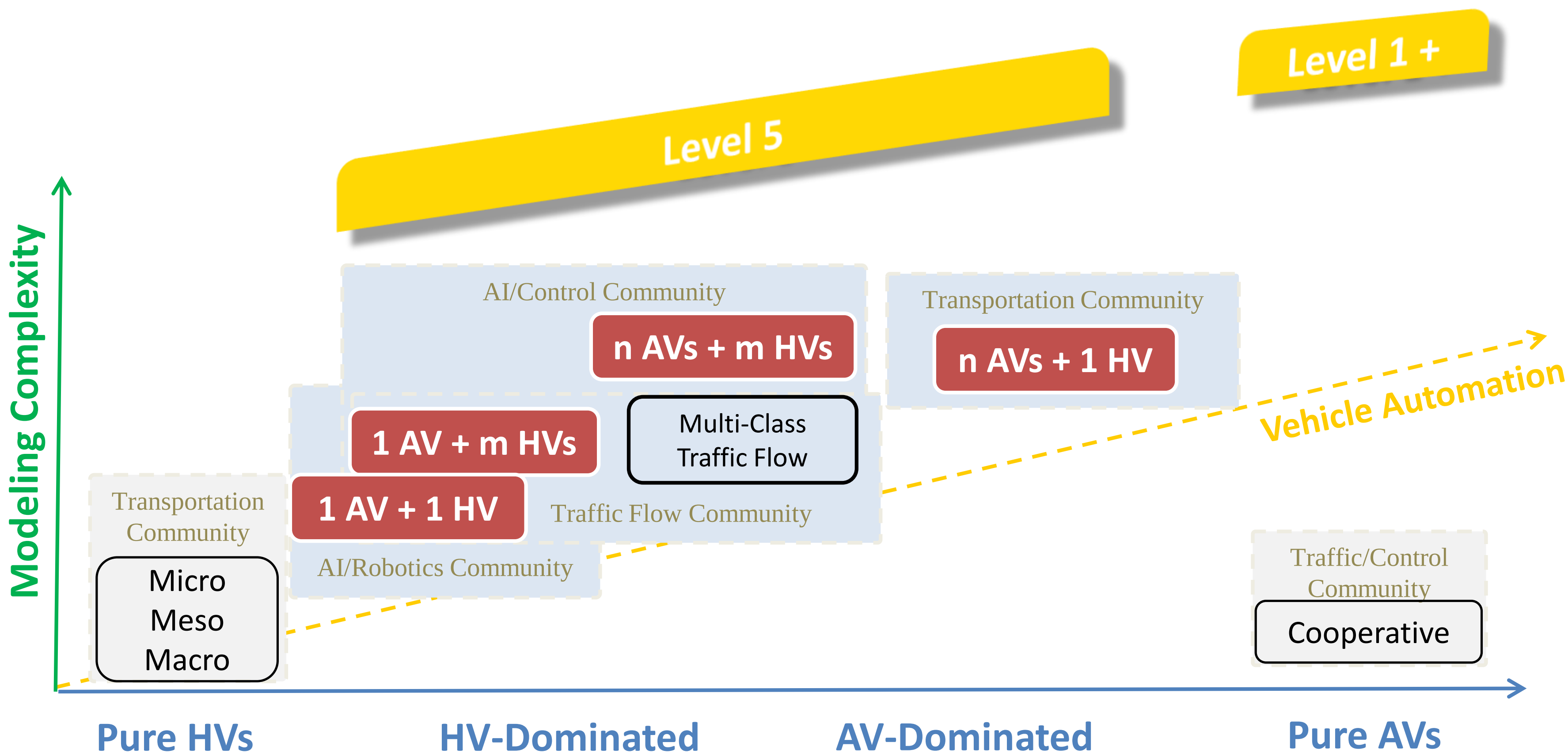
Kuang Huang, Qiang Du, Xi Chen (Columbia University)

Columbia University Data Science Institute

NSF

Mixed Autonomy: the Fundamental

1. Scalable AV Controller Design
2. Human Behavioral Estimation & Adaption
3. HV-AV Interaction Characterization



(**AV**: Autonomous Vehicle **HV**: Human-Driven Vehicle)

Game Theory

Empowering Driving Intelligence

AI

Control Agents

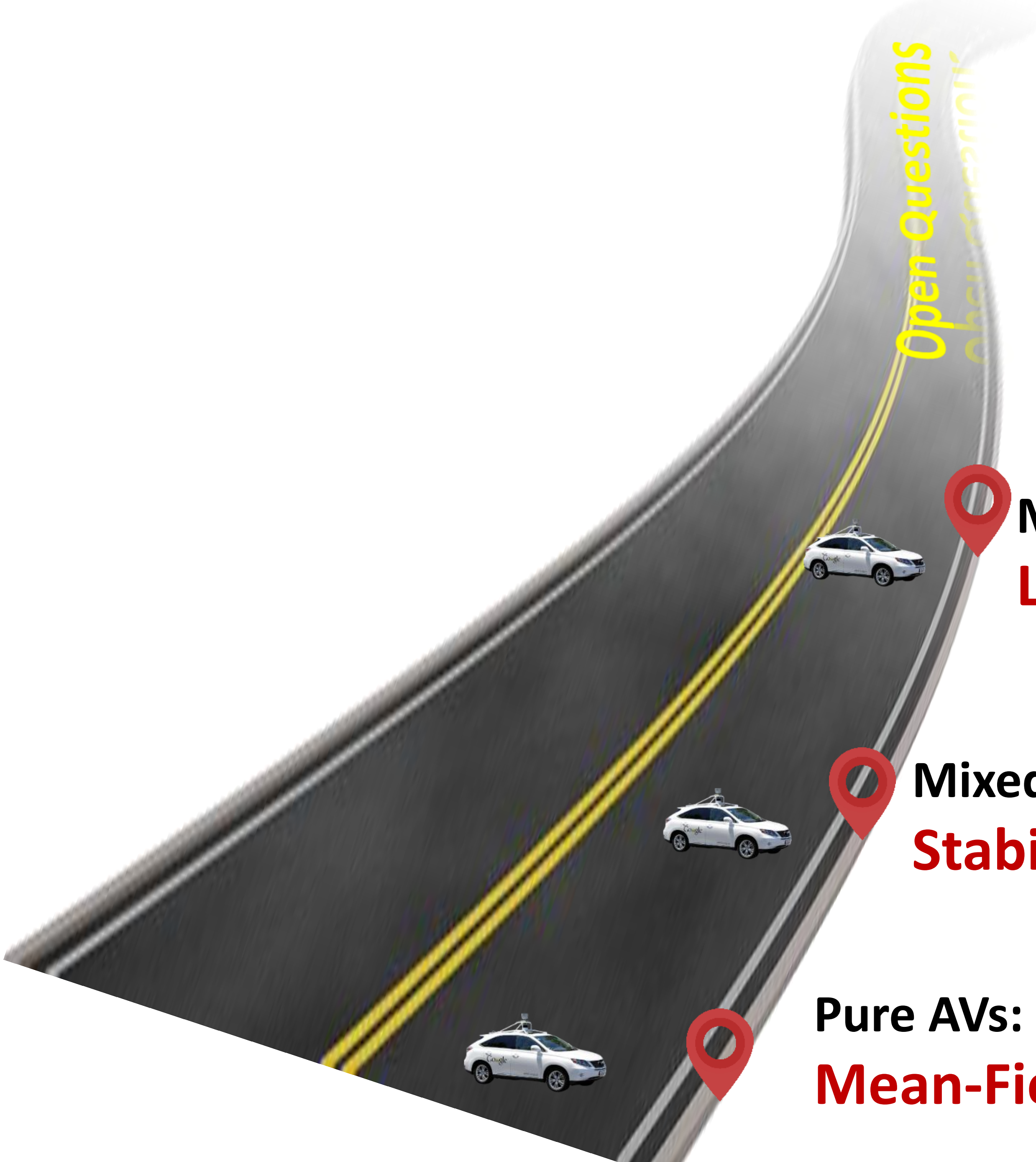
- ✓ Rational, utility-optimizing
- ✓ (Non)cooperative
- ✓ Learning & adaptation

Non-cooperative

Nobody slows down …

Proactive, Anticipating

Approaching the end, shall I force in?



Mixed AV-HV:
Learning Based Game

Mixed AV-HV:
Stability Analysis

Pure AVs:
Mean-Field Game

Multi-Autonomous Vehicle Control in Mixed Traffic

Game-Based Control

- Mean Field Game
- Stability

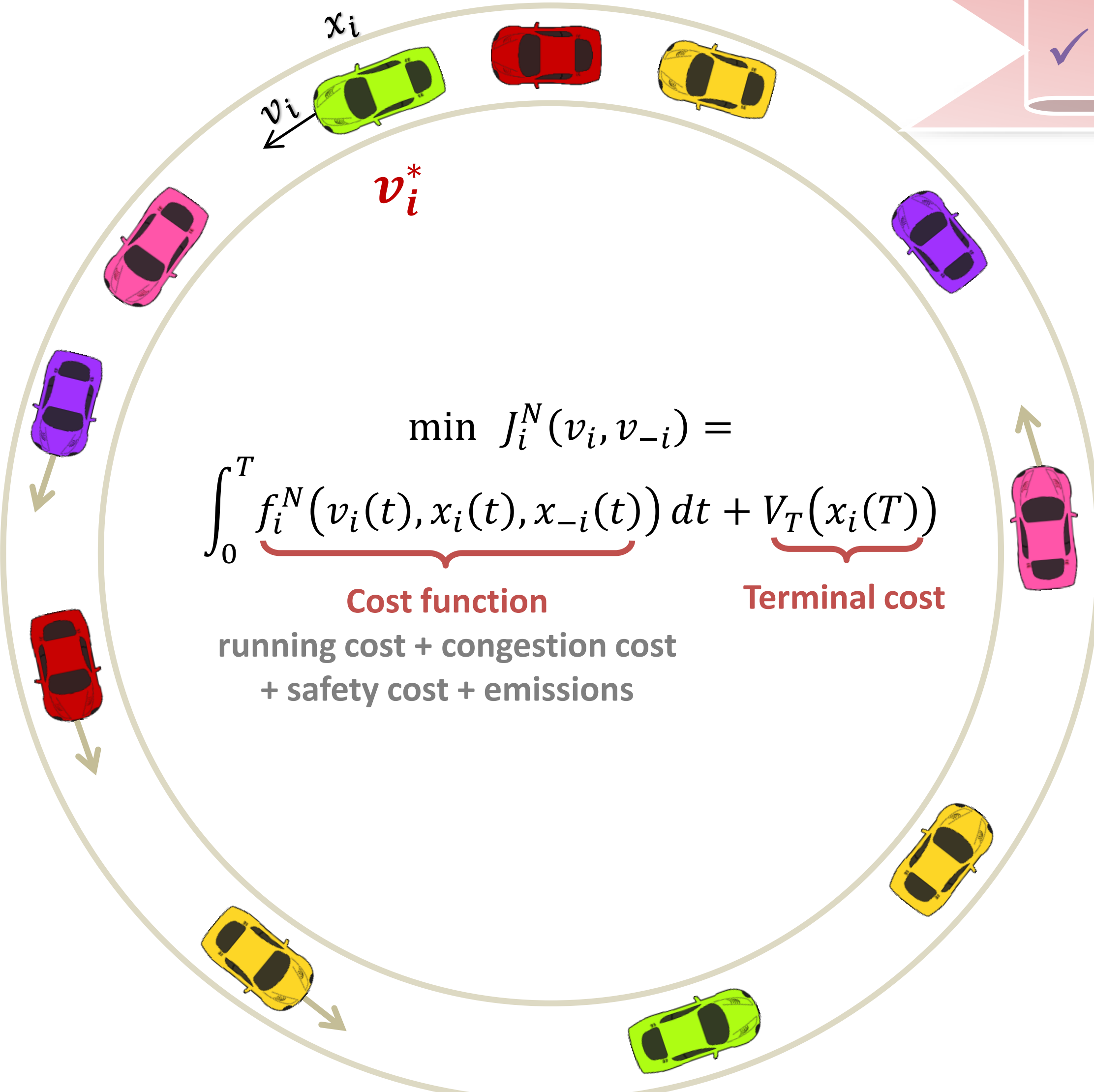
Learning-Based Control

Deterministic environment	Evolutionary game	Equilibrium
Stochastic, dynamic environment	Markov game	Learning

[Huang-Di-Du-Chen, *DCDS-B*, 2019; Huang-Di-Du-Chen, *TR part C*, 2020]

N-Car Differential Game

- ✓ Controls of many cars
- ✓ "New" traffic flow theory



$$\min \; J_i^N(v_i, v_{-i}) = \int_0^T \underbrace{f_i^N\big(v_i(t), x_i(t), x_{-i}(t)\big)}_{\text{Cost function}} dt + \underbrace{V_T\big(x_i(T)\big)}_{\text{Terminal cost}}$$

running cost + congestion cost + safety cost + emissions

Nash Equilibrium

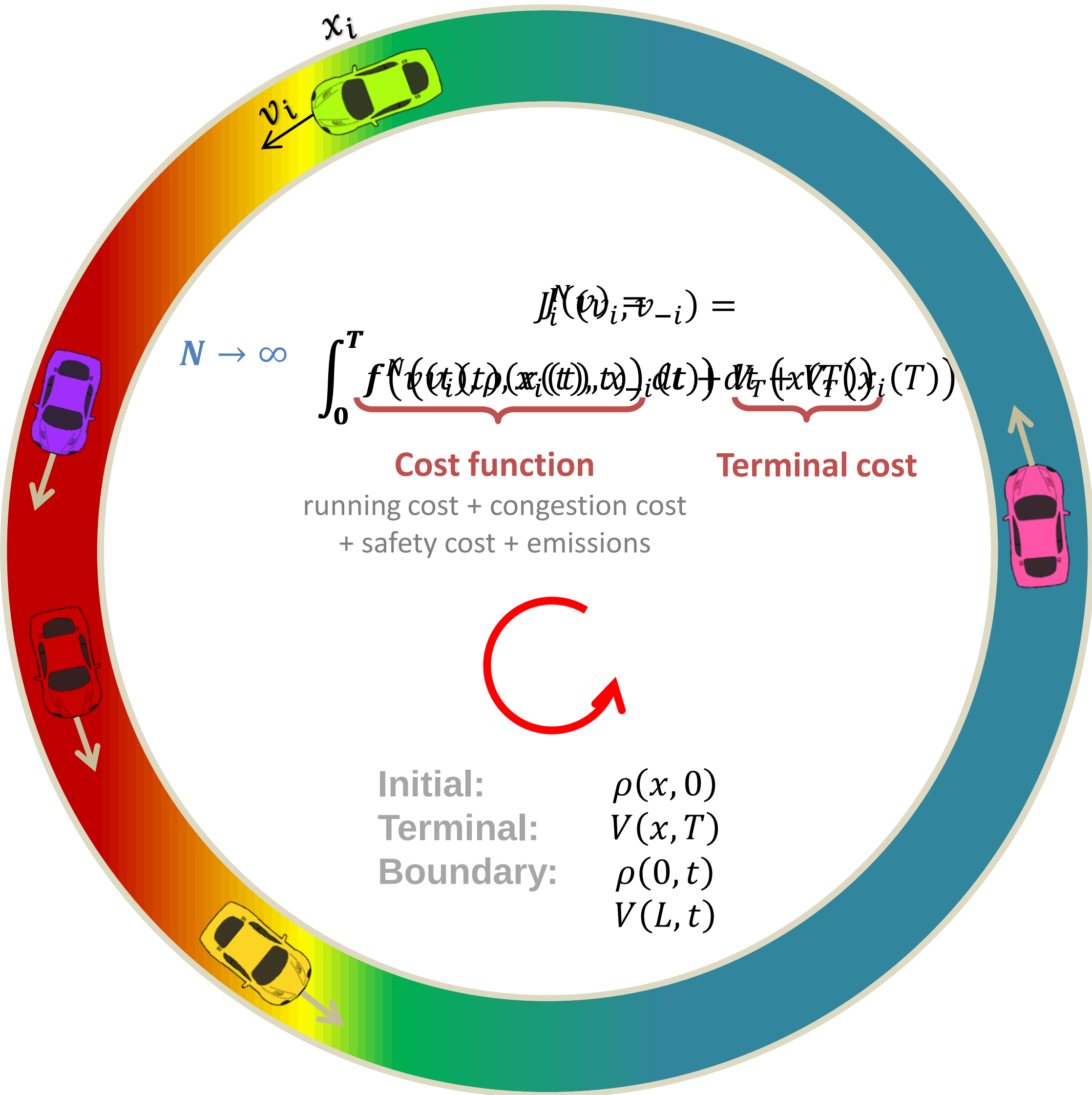
$$J_i^N(v_i^*, v_{-i}^*) \leq J_i^N(v_i, v_{-i}^*), \forall v_i, i$$

Mean Field Game: Decision-Making for Multi-Agent Systems

- Complex **multi-agent dynamic** systems if **N→∞**
- N-player limit of dynamic non-cooperative Nash game
- Micro (Agent) → Macro (Mass)

[Huang-Malhame-Caines, *Com. Info. Sys.*, 2006; Lions-Lasry, ***Paris-Princeton lectures,*** 2007]

Mean Field Approximation



$N \to \infty$

$$J_i^N(v) = J_i(v_i, v_{-i}) = \int_0^T \underbrace{f^N(v(t), x(t))}_{\text{Cost function}} dt + \underbrace{V_T^N(x(T))}_{\text{Terminal cost}}$$

$$\int_0^T f(v_i(t), \rho(x_i(t), t))\, dt + V_T(x_i(T))$$

Cost function

running cost + congestion cost + safety cost + emissions

Terminal cost

Initial:	$\rho(x,0)$
Terminal:	$V(x,T)$
Boundary:	$\rho(0,t)$
	$V(L,t)$

Mean Field Equilibrium

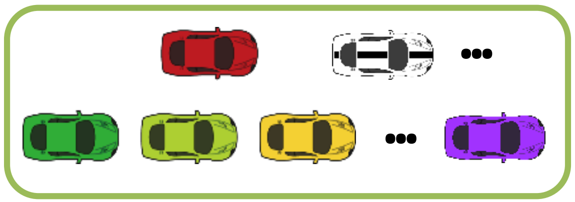
$$\boldsymbol{u^*(x,t)}$$

$$\boldsymbol{\rho^*(x,t)}$$

Forward-Backward Equations

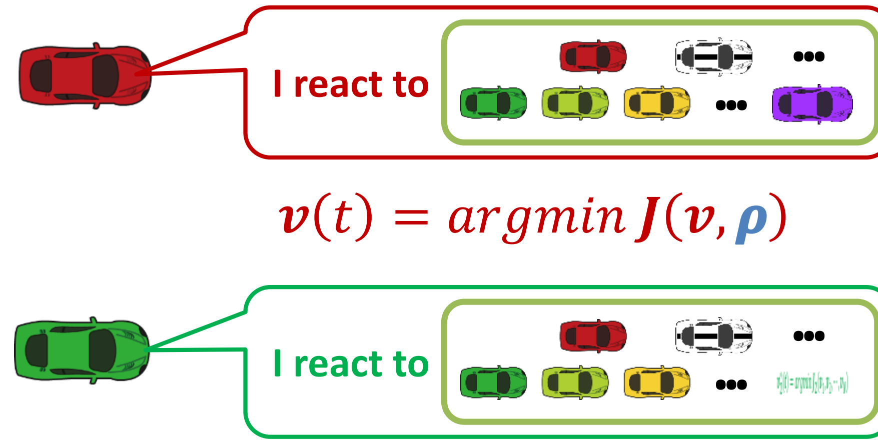
Mass Dynamic

Agent Dynamic



$$\boldsymbol{u(x,t)}$$

$$\boldsymbol{\rho(x,t)}$$



$$\rho_t + (\rho u)_x = 0$$

$$\boldsymbol{v}(t) = argmin\, \boldsymbol{J}(\boldsymbol{v}, \boldsymbol{\rho})$$

Fokker-Planck-Kolmogorov (FPK)

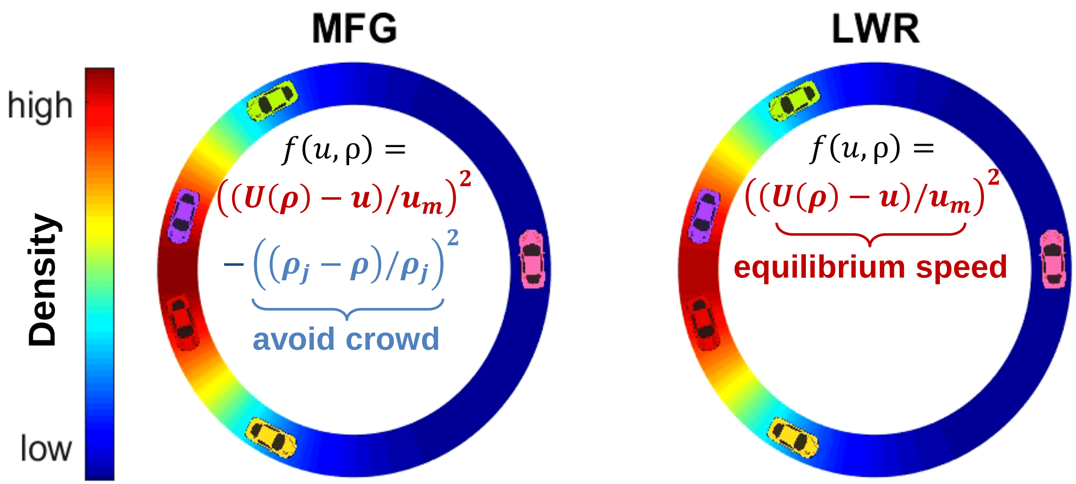
Hamilton-Jacobi-Bellman (HJB)

$$v_2^*(t) = argmin\, J_2(v_1, v_2, \cdots, v_N)$$

Forward

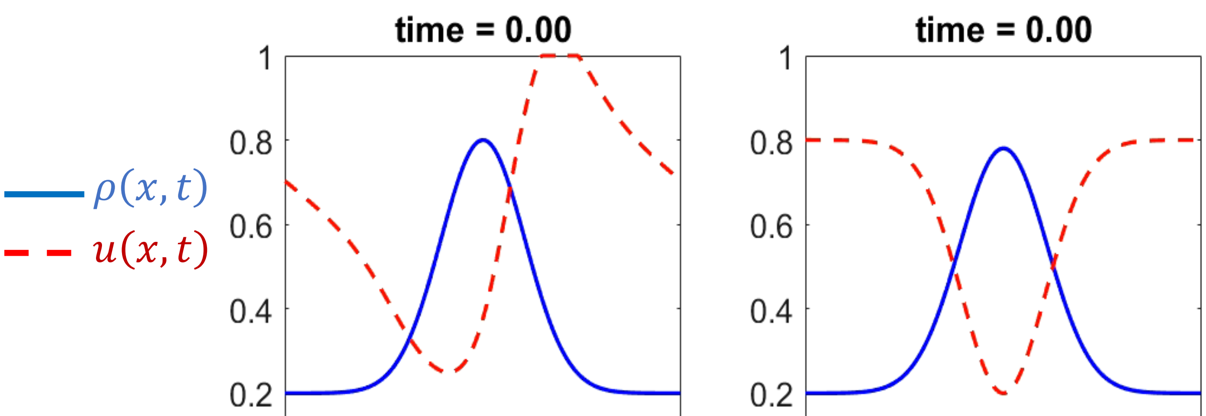
Backward

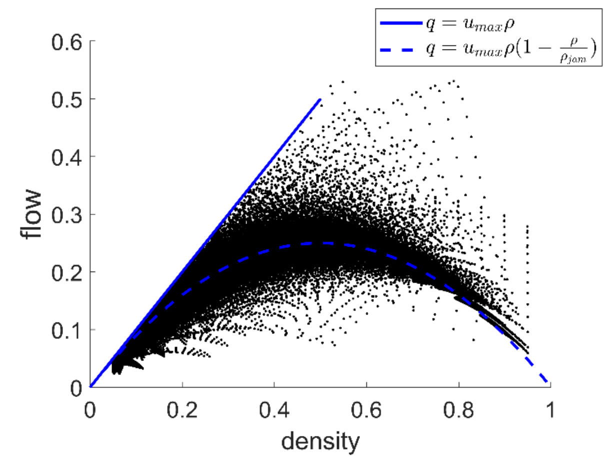
Game-Theoretic Interpretation of LWR



$$J(\boldsymbol{u}) = \int_0^T \underbrace{f(u(t),\rho(x(t),t))}_{\text{Cost function}}\,dt + V_T(x(T))$$

running cost + congestion cost
+ safety cost + emissions





LWR is a special type of MFG

1. LWR is the MFG with a cost function that aims to move toward an equilibrium speed
2. LWR is the *myopic* MFG with generic cost functions under regularity conditions

Linear Stability of Mixed Traffic

- ✓ Critical AV penetration
- ✓ Stability by design

Uniform Flow:

$$\rho^{AV}(x,t) \equiv \bar{\rho}^{AV}, \qquad \rho^{HV}(x,t) \equiv \bar{\rho}^{HV}$$

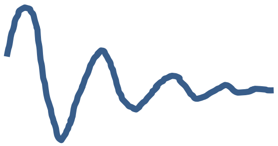
$(\bar{\rho}^{AV}, \bar{\rho}^{HV}, \bar{u})$

$$u^{AV}(x,t) \equiv u^{HV}(x,t) \equiv \bar{u}$$

Stability Conditions:

$$\sum_{i=AV,HV} \left\|\rho^i(\cdot,0) - \bar{\rho}^i\right\| + \|u^{HV}(\cdot,0) - \bar{u}\,\| \leq \delta$$

$$\sup_{0\leq t\leq T} \left\{ \sum_{i=AV,HV} \left\|\rho^i(\cdot,t) - \bar{\rho}^i\,\right\| + \left\|u^i(\cdot,t) - \bar{u}\,\right\| \right\} \leq \epsilon$$



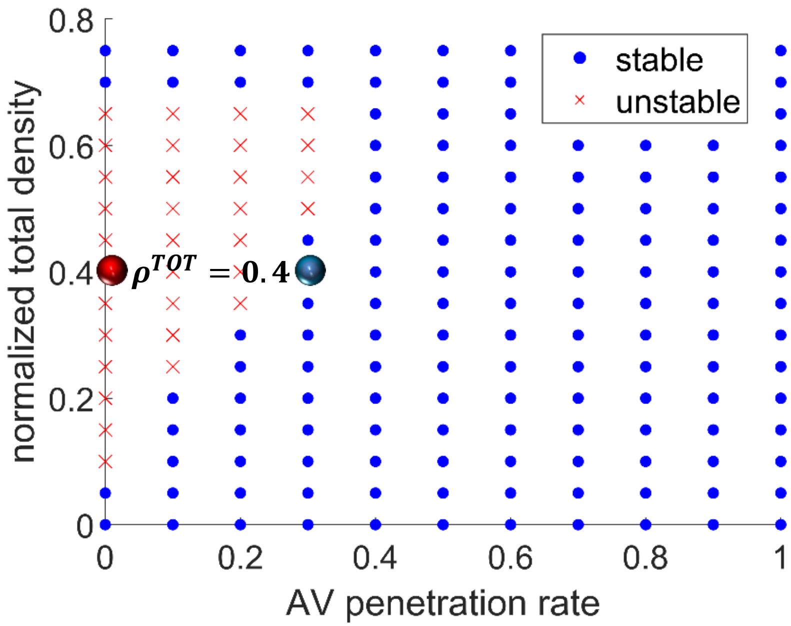
Mixed Traffic: Mean Field Game (MFG) + ARZ

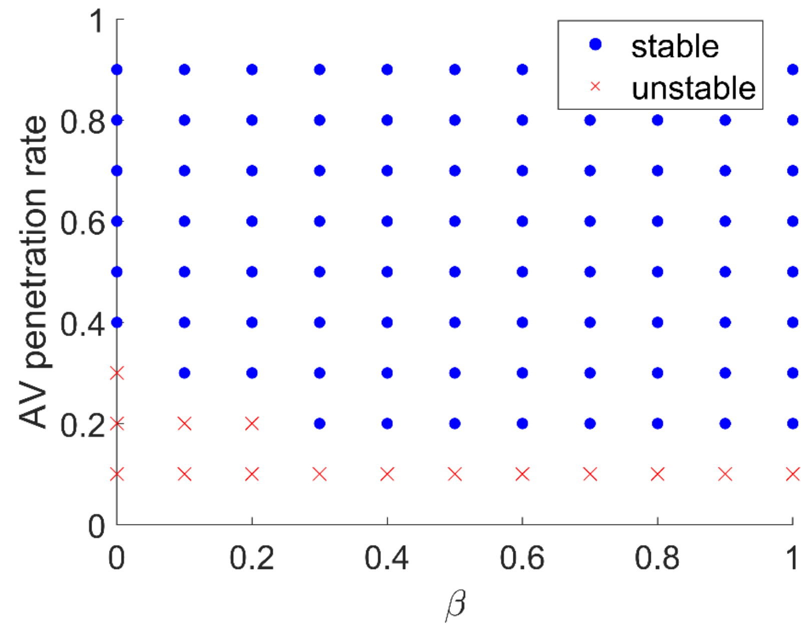
$$(\rho^{AV}, \rho^{HV}, u^{AV}, u^{HV})$$

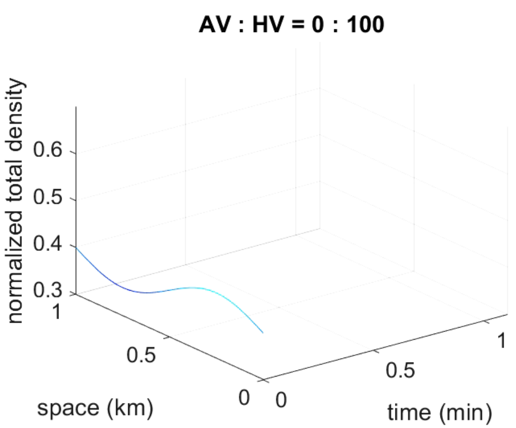
	MFG		ARZ
Flow	$\rho_t^{AV} + (\rho^{AV} u^{AV})_x = 0$	$+$ ρ^{TOT}	$\rho_t^{HV} + (\rho^{HV} u^{HV})_x = 0$
Dynamic	$\begin{cases} V_t + uV_x + \frac{1}{2}\left(\frac{u^{AV}}{u_m}\right)^2 - \frac{u^{AV}}{u_m} + \frac{u^{AV}\rho^{TOT}}{u_m \rho_j} + \boldsymbol{\beta}\frac{\rho^{HV}}{\rho_j} = 0 \\ \boldsymbol{u} = u_{max}\left(1 - \frac{\rho^{TOT}}{\rho_j} - u_m V_x\right) \end{cases}$		$[u^{HV} + h(\rho^{TOT})]_t + u^{HV}[u^{HV} + h(\rho^{TOT})]_x = \frac{1}{\tau}(U(\rho^{TOT}) - u^{HV})$
	(HJB Equation)		(Momentum Equation)

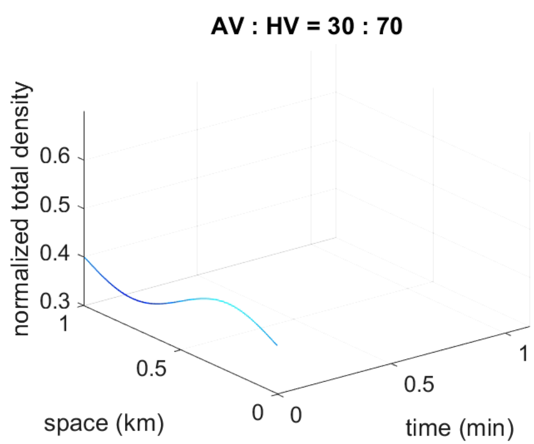
(**AV**: Autonomous Vehicle **HV**: Human-Driven Vehicle)

Stabilizing Traffic via MFG Control









$$f(u^{AV}, \rho^{AV}, \rho^{HV})$$

$$= \frac{1}{2}\left(\frac{u^{AV}}{u_{max}}\right)^2 - \frac{u^{AV}}{u_{max}} + \frac{u^{AV}\rho^{TOT}}{u_{max}\rho_{jam}} + \boldsymbol{\beta}\,\frac{\rho^{HV}}{\rho_{jam}}$$

Autonomous Controller Design

(**AV**: Autonomous Vehicle **HV**: Human-Driven Vehicle)

Multi-Autonomous Vehicle Control in Mixed Traffic

Game-Based Control
- Mean Field Game
- Stability

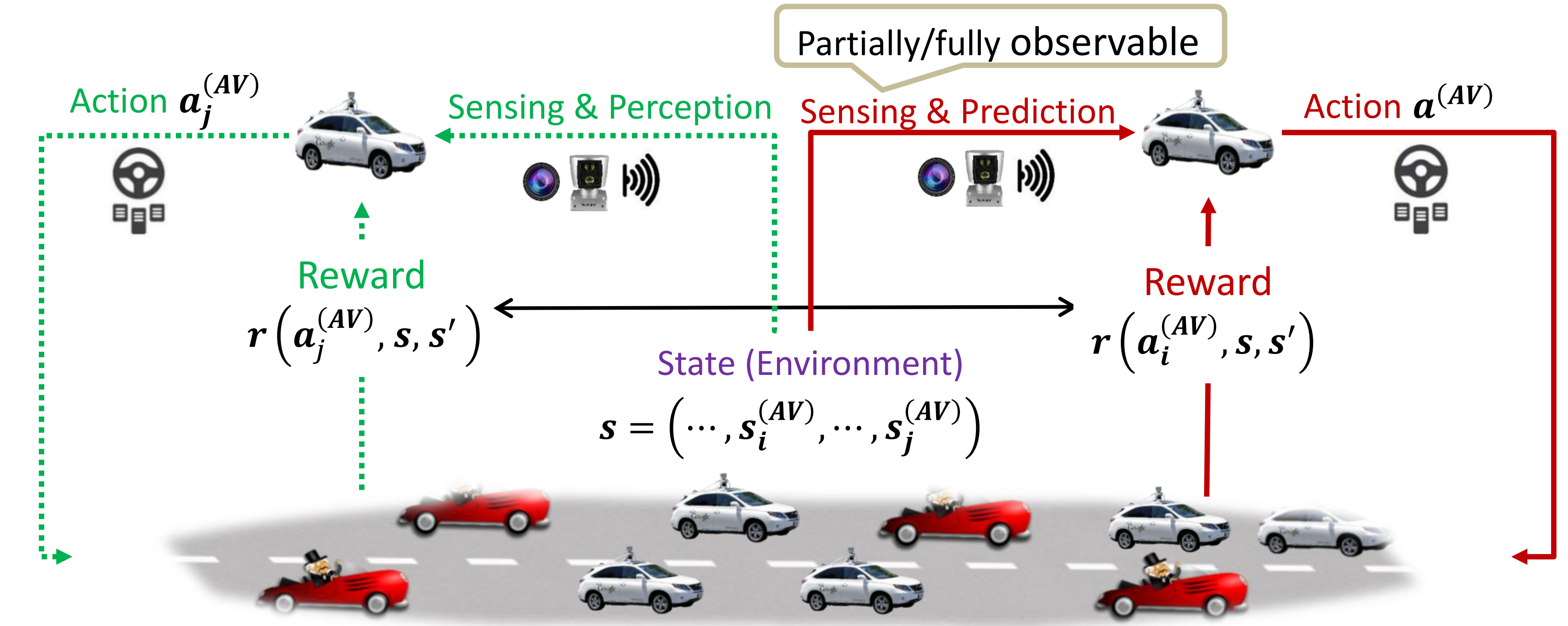
Learning-Based Control
- Reinforcement Learning
- Imitation Learning

Deterministic environment	Evolutionary game	Equilibrium
Stochastic, dynamic environment	Markov game	Learning

“It is not the strongest that survives, nor the most intelligent. It is the one that is most adaptable to change.”

Multi-Agent

Reinforcement Learning



HVs

- ✓ Heterogeneity
- ✓ Stochasticity

Imitation Learning of Mixed Traffic

Uncontrolled AVs

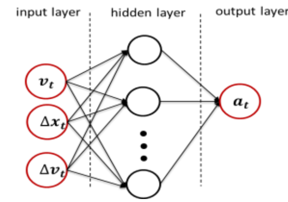
- ✓ Unconventional data

Imitation Learning of Mixed Environment

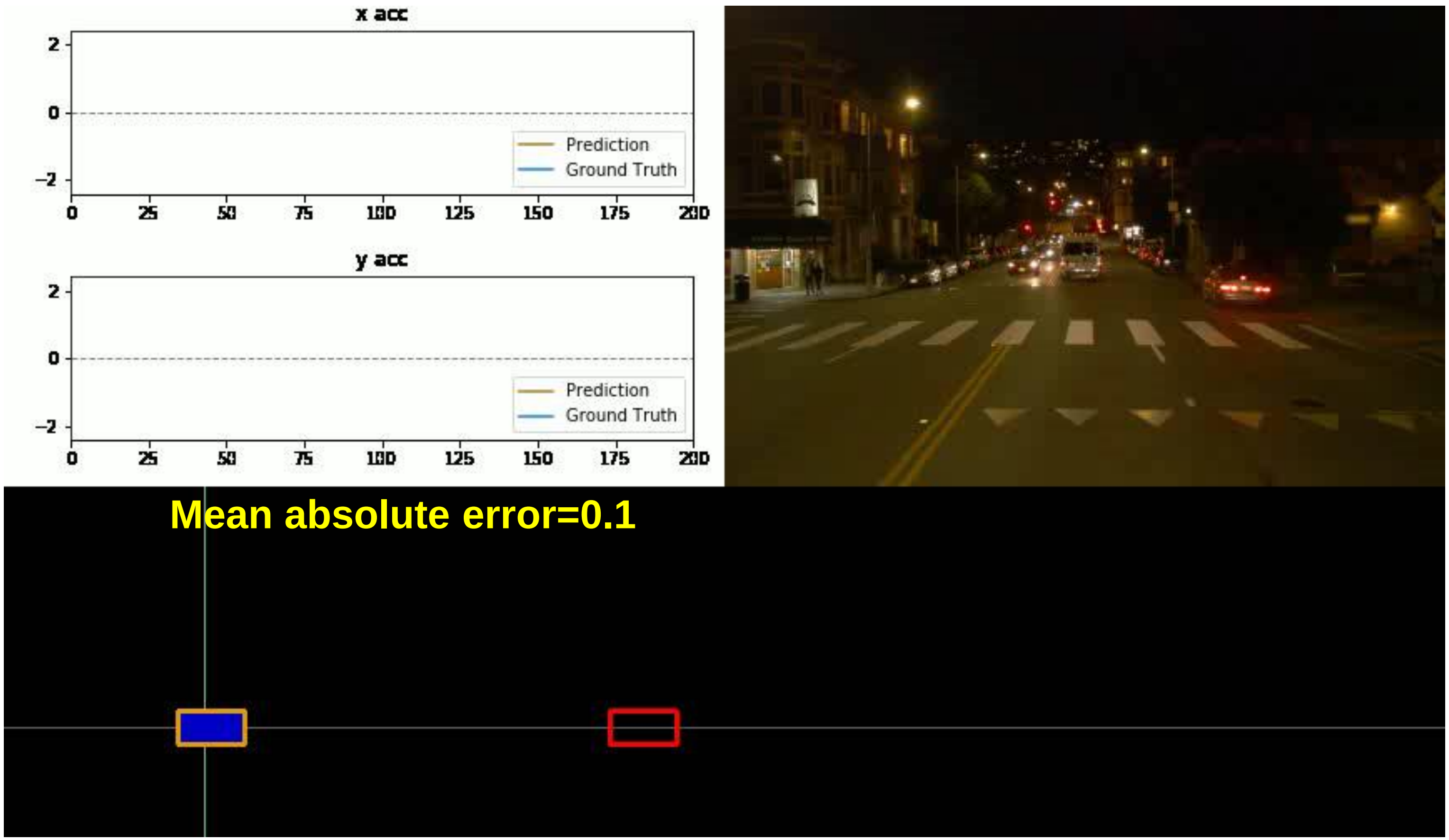
$\Delta \boldsymbol{x}_i(t), \Delta \boldsymbol{v}_i(t), \boldsymbol{v}_i(t)$

images

Long Short-Term Memory (LSTM)



$a_i(t)$



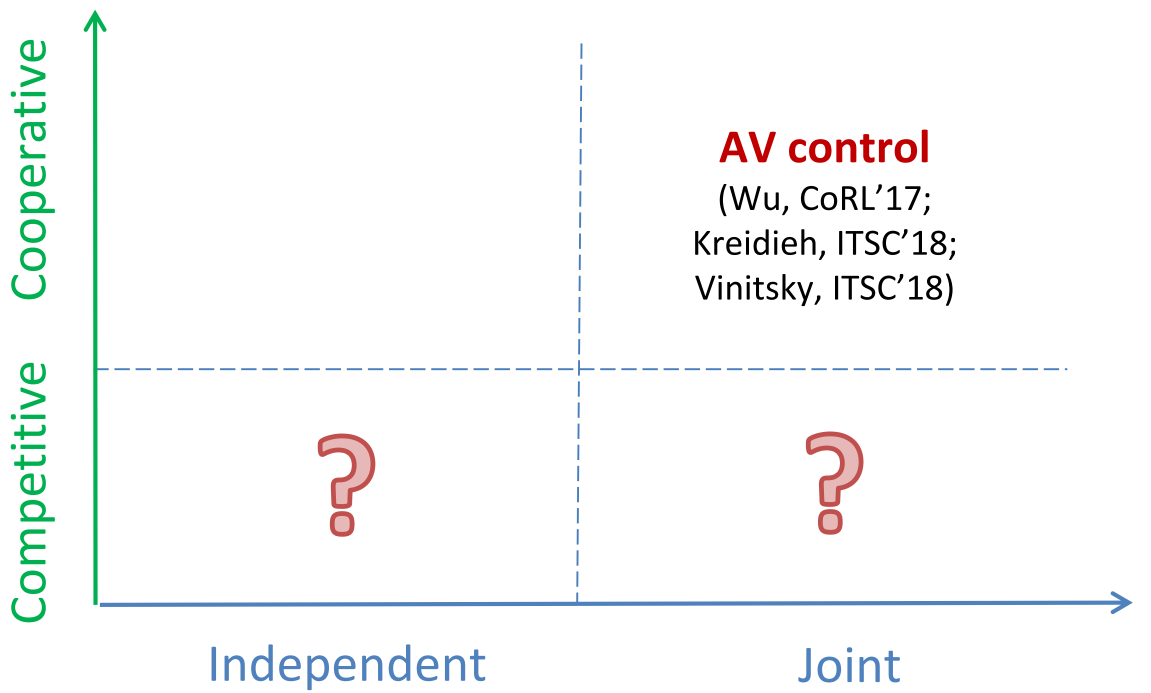
[Li, Z.H., Gu, Z.C., **Di, X.**, Shi, R.Y., 2020. An LSTM-Based Autonomous Driving Model Using Waymo Open Dataset, ***Applied Sciences*** - Intelligent Transportation Systems: Beyond Intelligent Vehicles, 10(6), 2046.]

Multi-Agent Reinforcement Learning (MARL)

Multi-

Single- Agent Reinforcement learning Markovian

Game Theory



Recap…

Multi-Autonomous Vehicle Control in Mixed Traffic

Game-Based Control
- **Mean Field Game**
- **Stability**

Learning-Based Control
- **Reinforcement Learning**
- **Imitation Learning**

Open Questions

- ✓ **Autonomous Driving Model**
- ✓ **Multi-Scale:** micro – macro
- ✓ **Mixed Traffic Simulator**
- ✓ **Social Implications**

1. Scalable AV Controller Design
2. Human Behavioral Estimation & Adaption
3. HV-AV Interaction Characterization

[Rahwan et al, "Machine behavior", *Nature,* 2019]

THANK YOU!

Questions?

Data Science Institute
COLUMBIA UNIVERSITY

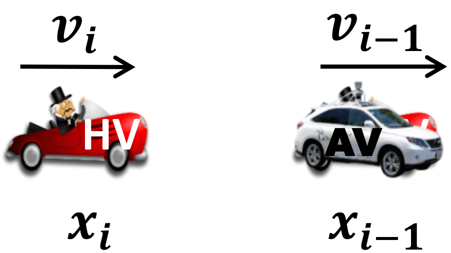
COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Ditech

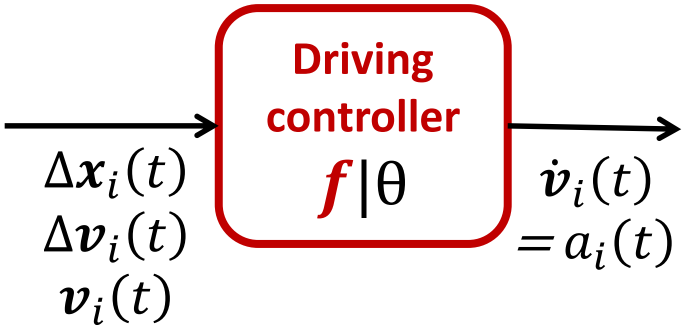
Reference

1. Achdou, Y. and Perez, V., 2012. Iterative strategies for solving linearized discrete mean field games systems. Networks & Heterogeneous Media, 7(2).
2. Couillet, R., Perlaza, S.M., Tembine, H. and Debbah, M., 2012. Electrical vehicles in the smart grid: A mean field game analysis. IEEE Journal on Selected Areas in Communications, 30(6), pp.1086-1096.
3. Cui, S., Seibold, B., Stern, R. and Work, D.B., 2017, June. Stabilizing traffic flow via a single autonomous vehicle: Possibilities and limitations. In 2017 IEEE Intelligent Vehicles. Symposium (IV) (pp. 1336-1341). IEEE.
4. 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: Methodological 94, 314-334.
5. Lachapelle, A. and Wolfram, M.T., 2011. On a mean field game approach modeling congestion and aversion in pedestrian crowds. Transportation research part B: methodological, 45(10), pp.1572-1589.
6. Li, N., Oyler, D. W., Zhang, M., Yildiz, Y., Kolmanovsky, I., Girard, A. R., 2018. Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems. IEEE Transactions on control systems technology, 26 (5), 1782-1797.
7. 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.
8. Sadigh, D., Sastry, S., Seshia, S. A., Dragan, A. D., 2016. Planning for autonomous cars that leverage effects on human actions. In: Robotics: Science and Systems.
9. Sadigh, D., Sastry, S.S. and Seshia, S.A., 2019. Verifying Robustness of Human-Aware Autonomous Cars. IFAC-PapersOnLine, 51(34), pp.131-138.
10. Stern, R.E., Cui, S., Delle Monache, M.L., Bhadani, R., Bunting, M., Churchill, M., Hamilton, N., Pohlmann, H., Wu, F., Piccoli, B. and Seibold, B., 2018. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. Transportation Research Part C: Emerging Technologies, 89, pp.205-221.
11. Wu, C., Bayen, A. M., Mehta, A., 2018. Stabilizing traffic with autonomous vehicles. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 1-7.
12. Wang, M., Hoogendoorn, S.P., Daamen, W., van Arem, B. and Happee, R., 2015. Game theoretic approach for predictive lane-changing and car-following control. Transportation Research Part C: Emerging Technologies, 58, pp.73-92.
13. Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.F., Breazeal, C., Crandall, J.W., Christakis, N.A., Couzin, I.D., Jackson, M.O. and Jennings, N.R., 2019. Machine behaviour. *Nature*, *568*(7753), pp.477-486.

Control Agents



- ✓ (Collaborative) adaptive cruise control
- ✓ Nonlinear car following model
- ✓ Optimal control problem
- ✓ Model predictive control



		
Reaction time	Long	**Short**
Information	Local	**Global**
Driving control	Heterogeneous	**Homogeneous**

Reinforcement & Imitation Learning

Reinforcement
$(\boldsymbol{S}, \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{s}_0, \boldsymbol{R})$

Find π_θ
s.t. $\pi_\theta = \arg\max E[V^\pi(s_0)]$

Optimal sequence of actions $\boldsymbol{a}_0 \boldsymbol{a}_1 \cdots \boldsymbol{a}_t$

- Model-based: dynamic programming
- Model-free: Q-learning

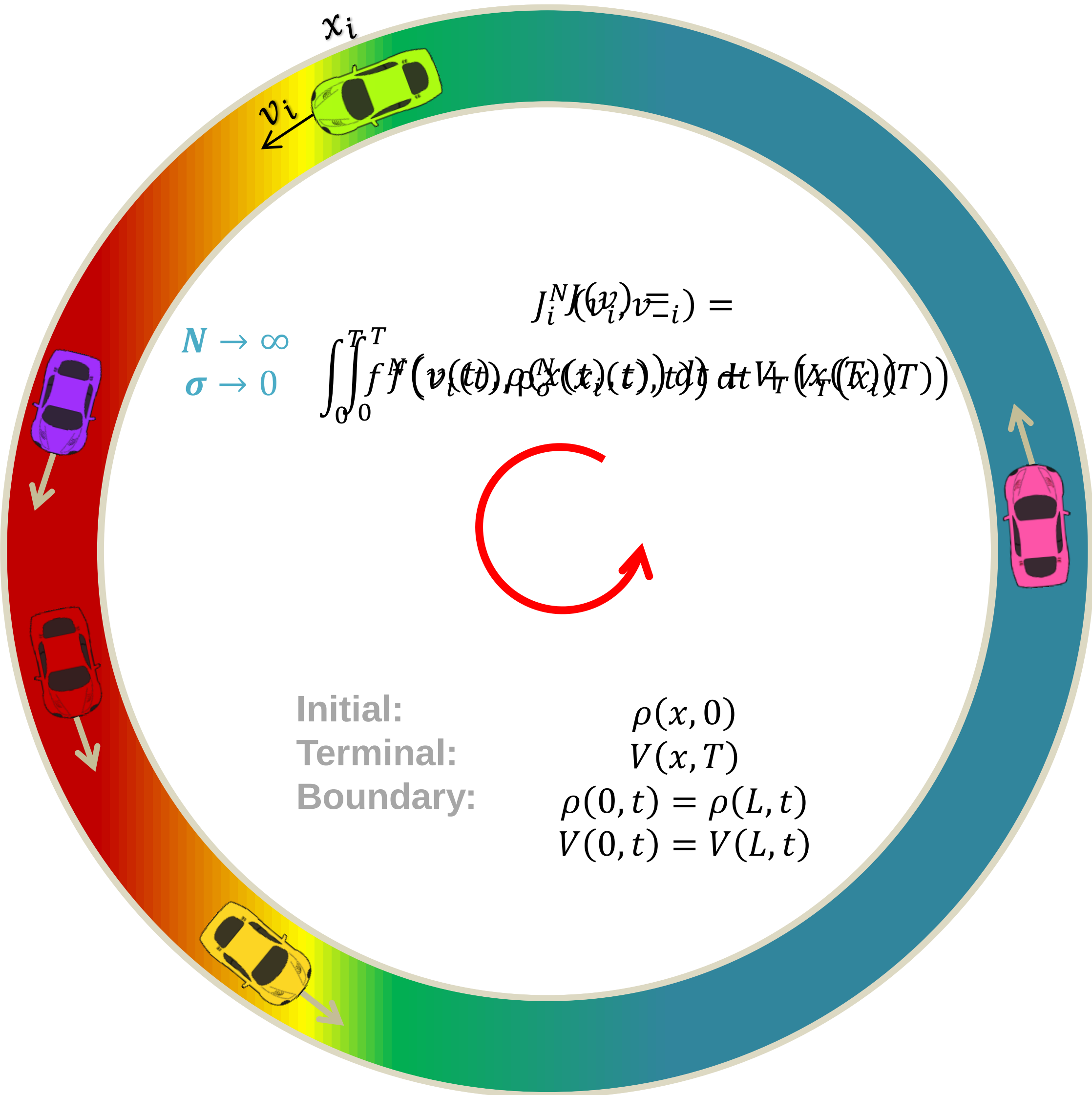
Imitation
$(\boldsymbol{S}, \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{s}_0\ \boldsymbol{R})$

Sampled trajectories $\tau = \boldsymbol{s}_0 \boldsymbol{a}_0 \boldsymbol{s}_1 \boldsymbol{a}_1 \cdots \boldsymbol{s}_t \boldsymbol{a}_t \cdots$

Find π_θ
s.t. $\min E_\tau \|\pi(s_0) - \pi_\theta(s_0)\|$

- Supervised learning: behavioral cloning
- Imitation learning

Supplemental Materials



$N \to \infty$
$\sigma \to 0$

$J_i^N(v_i, v_{-i}) =$ $J(v) =$

$\int_0^T f^N\left(v_i(t), \rho_\sigma^N(x_i(t), t)\right) dt + V_T(x_i(T))$ $\int_0^T f\left(v(t), \rho(x(t), t)\right) dt + V_T(x(T))$

Initial: $\rho(x, 0)$
Terminal: $V(x, T)$
Boundary: $\rho(0, t) = \rho(L, t)$
$V(0, t) = V(L, t)$

Mean Field Equilibrium:

$\rho^*(x, t)$
$u^*(x, t)$

$$(FPK)\ \rho_t + (\rho u)_x = 0$$

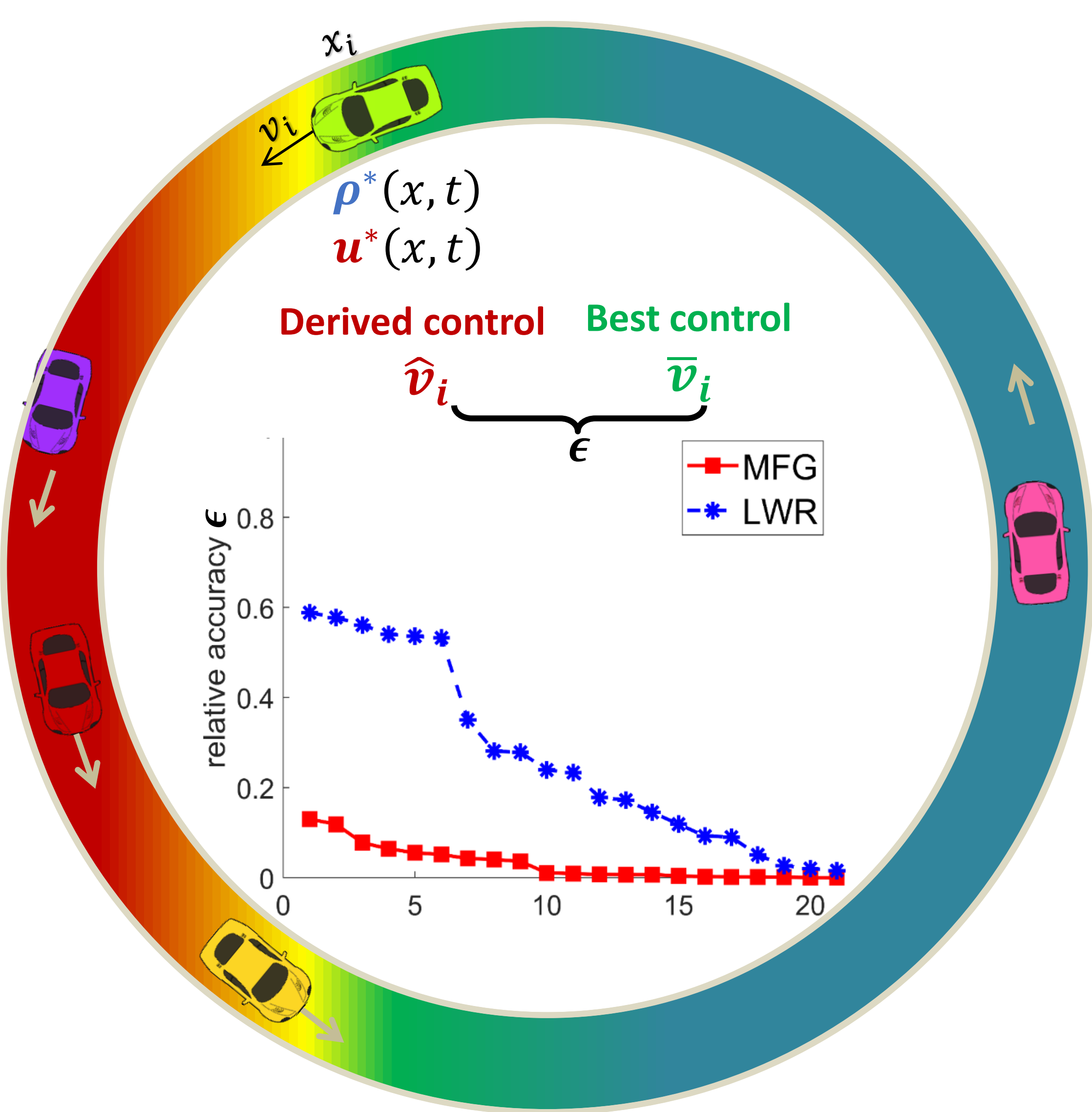
$$(HJB)\begin{cases} V_t + f^*(V_x, \rho) = 0 \\ u = f_p^*(V_x, \rho) \end{cases}$$

$$\parallel$$

$$argmin_\alpha \{f(\alpha, \rho) + \alpha V_x\}$$

$$f^*(p, \rho) = argmin_\alpha \{f(\alpha, \rho) + \alpha p\}$$

ϵ-Nash Equilibrium of Differential Game



$$u^*(x,t)$$

$$\begin{cases} \hat{v}_i(t) = u^*(x_i(t),t) \\ \dot{x}_i(t) = \hat{v}_i(t), x_i(0) = x_{i,0} \end{cases}$$

ϵ-Nash Equilibrium:

$$J_i^N(\hat{v}_i, \hat{v}_{-i}) \le J_i^N(v_i, \hat{v}_{-i}) + \epsilon,$$

$$\forall v_i \in A, i = 1, \cdots, N$$

Challenge: forward-backward structure

Fixed-point Iteration

- Alternatingly solve the forward and backward equations
- Simple to implement
- Fail to converge when T is relatively large

Variational Form

- Only valid for separable cost function (potential game)
- Transform into an optimization

Finite-Difference Newton’s Method

- View forward and backward equations as a whole nonlinear system
- Use Newton’s method to solve the nonlinear system
- Not so efficient, no guarantee on convergence

Couillet, R., Perlaza, S.M., Tembine, H. and Debbah, M., 2012. Electrical vehicles in the smart grid: A mean field game analysis. IEEE Journal on Selected Areas in Communications, 30(6), pp.1086-1096.

Lachapelle, A. and Wolfram, M.T., 2011. On a mean field game approach modeling congestion and aversion in pedestrian crowds. Transportation research part B: methodological, 45(10), pp.1572-1589.

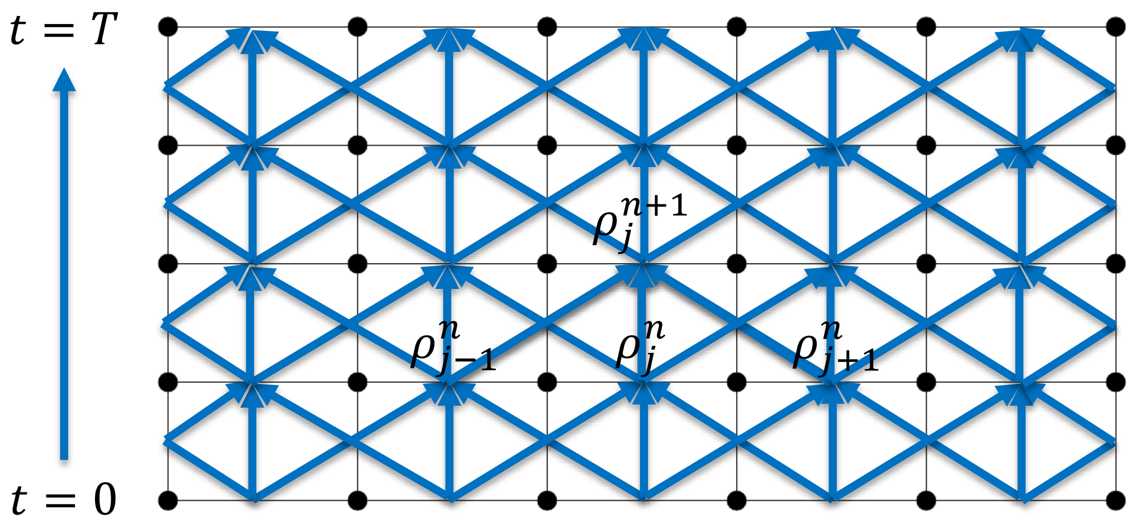
Achdou, Y. and Perez, V., 2012. Iterative strategies for solving linearized discrete mean field games systems. Networks & Heterogeneous Media, 7(2).

Newton's Method

Forward

continuity equ.:

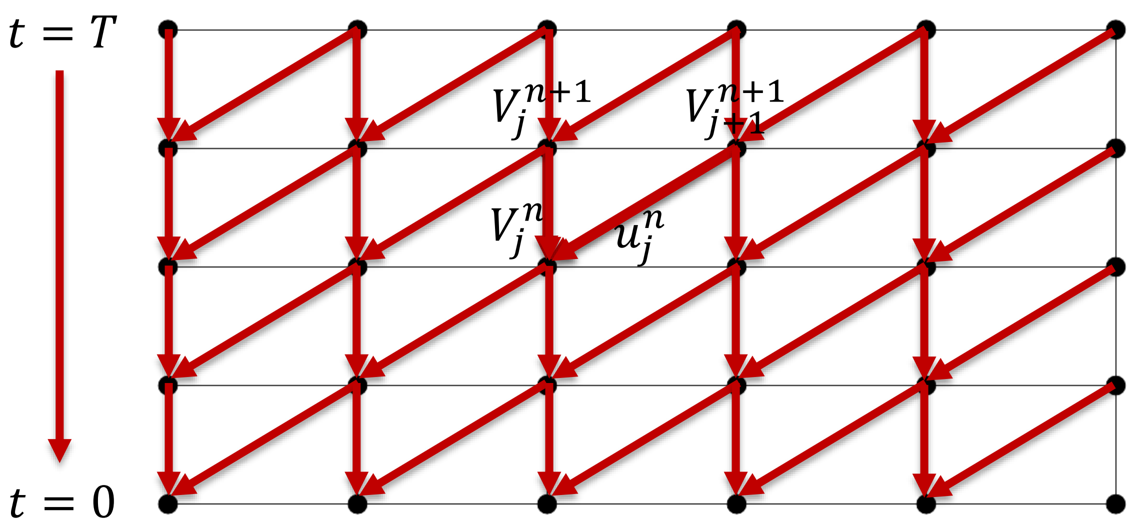
Lax-Friedrichs scheme



Backward

HJB equ.:

upwind scheme



Step 1. Space-time grids

$$0 = x_0 < x_1 < \cdots < x_{N_x} = L, \qquad x_j = j\Delta x, \qquad \Delta x = \frac{L}{N_x}$$

$$0 = t^0 < t^1 < \cdots < t^{N_t} = T, \qquad t^n = n\Delta t, \qquad \Delta t = \frac{T}{N_t}$$

Discretize density, velocity and cost

$$\rho_j^n = \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} \rho(x, t^n)\, dx \qquad u_j^n = u(x_{j+1/2}, t^n) \qquad V_j^n = V(x_j, t^n)$$

Proposed Algorithm

Step 2. Discretization of equations

- Lax-Friedrichs scheme for continuity equation
- Upwind scheme for HJB equations (dynamic programming)

$$\rho_j^{n+1} = \frac{1}{2}\left(\rho_{j-1}^n + \rho_{j+1}^n\right) - \frac{\Delta t}{2\Delta x}\left(\rho_{j+1}^n u_{j+1}^n - \rho_{j-1}^n u_{j-1}^n\right)$$

$$\frac{V_j^{n+1} - V_j^n}{\Delta t} = \frac{1}{2}\left(\frac{u_j^n}{u_f}\right)^2$$

$$\frac{\rho_j^n}{\rho_j} + \frac{u_j^n}{u_f} + \frac{V_{j+1}^{n+1} - V_j^{n+1}}{\Delta x} = 1$$

Plus initial/terminal conditions

Proposed Algorithm

Step 3. Solve the discretized system as a large system of nonlinear equations

$\rho_j^n, u_j^n, V_j^n \rightarrow$ Vector w

$$\rho_j^{n+1} = \frac{1}{2}(\rho_{j-1}^n + \rho_{j+1}^n) - \frac{\Delta t}{2\Delta x}(\rho_{j+1}^n u_{j+1}^n - \rho_{j-1}^n u_{j-1}^n)$$

$$\frac{V_j^{n+1} - V_j^n}{\Delta t} = \frac{1}{2}\left(\frac{u_j^n}{u_f}\right)^2 \rightarrow F(w) = 0$$

$$\frac{\rho_j^n}{\rho_j} + \frac{u_j^n}{u_f} + \frac{V_{j+1}^{n+1} - V_j^{n+1}}{\Delta x} = 1$$

Step 3. Solve the discretized system as a large system of nonlinear equations

- Newton's method

$$w^{n+1} = w^n - (JF(w^n))^{-1}F(w^n)$$

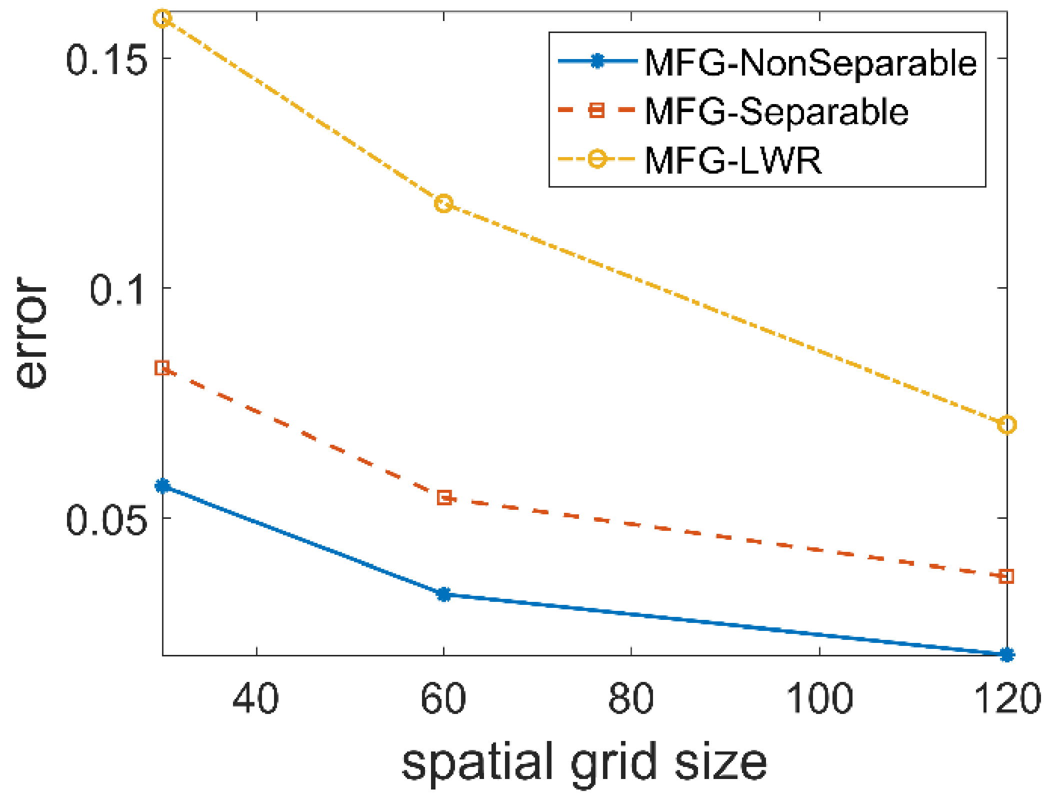
- Multigrid

$$32\text{x}100 \xrightarrow{\text{initial guess}} 64\text{x}200 \xrightarrow{\text{initial guess}} 128\text{x}400$$

- Preconditioning

At each Newton's step, use the uncoupled system as the preconditioner, apply iterative linear solvers

Algorithm Convergence



Theorem 1.

The solution of the LWR model

$$\rho_t + (\rho U(\rho))_x = 0$$

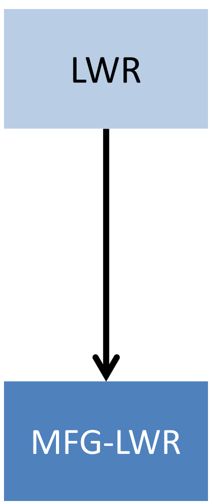
is a solution of the MFG system with cost function

$$f(u,\rho) = \frac{1}{2}(U(\rho) - u)^2$$

under the conditions that:

(i) *They have the same initial density $\rho_0(x)$ and boundary conditions;*

(ii) *The terminal cost $V_T(x) = 0$.*



$$\boldsymbol{u} = U(\boldsymbol{\rho})$$

$$\Big\downarrow \quad \boldsymbol{u} = argmin\, J(\boldsymbol{u},\boldsymbol{\rho}) = \int_0^T f\big(v(t), \rho(x(t),t)\big)\, dt$$

$$f(u,\rho) = \frac{1}{2}(U(\rho) - u)^2$$

Theorem 2.

Under the conditions that:

(i) *The cost function* $f(u,\rho)$ *is continuously differentiable, strictly convex w.r.t* u*;*

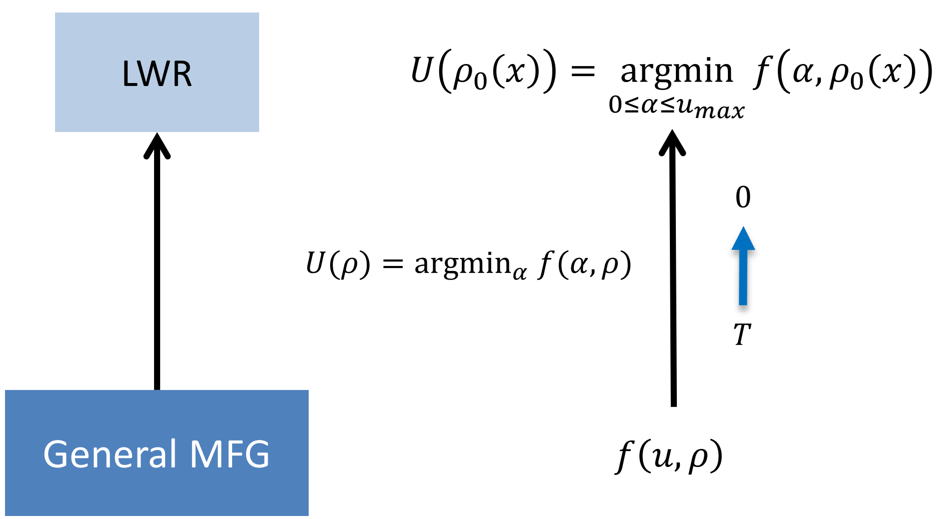
(ii) *The terminal cost* $V_T(x) = 0$*;*

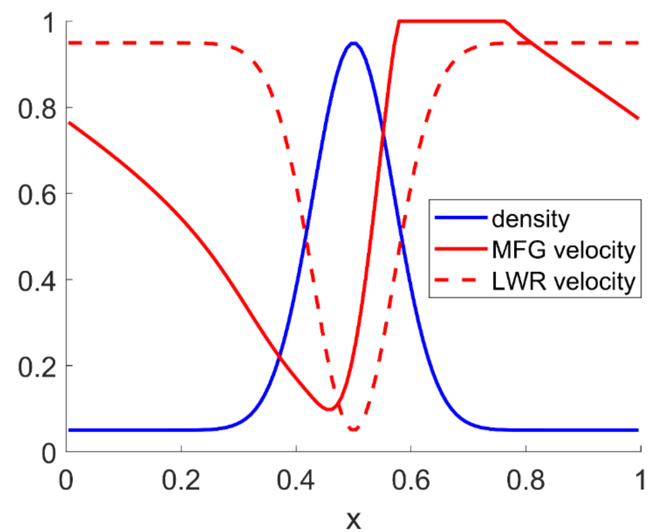
(iii) $\exists T_0 > 0$ *s.t. whenever* $0 < T \le T_0$ *, the MFG system has a unique solution* $\rho^{(T)}(x,t), u^{(T)}(x,t)$ *and* $V^{(T)}(x,t)$ *that are uniformly bounded up to second order derivatives on* $0 \le x \le L, 0 \le t \le T \le T_0$.

When $T \to 0$ *we have:*

$$\lim_{T\to 0} u^{(T)}(x,0) = U(\rho_0(x)) = \underset{0\le\alpha\le u_{max}}{\operatorname{argmin}} f(\alpha,\rho_0(x)),$$

which is LWR with fundamental diagram $U(\rho)$ *and initial density* $\rho_0(x)$.





Stability Analysis of MFG

Uniform Flow:

$(\bar{\rho}, \bar{u})$

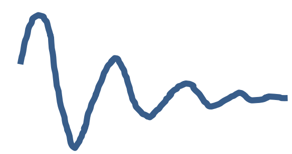
$$\bar{\rho} \equiv \rho(x,t)$$

$$\bar{u} = U(\bar{\rho})$$

Stability Conditions:

$$\|\rho(\cdot,0) - \bar{\rho}\,\| \le \delta$$

$$\sup_{0\le t\le T} \left\{\left\|\rho^{(T)}(\cdot,t) - \bar{\rho}\,\right\| + \left\|u^{(T)}(\cdot,t) - \bar{u}\,\right\|\right\} \le \epsilon$$



Theorem 3.

The MFG system is linearly stable around the uniform flow $(\bar{\rho}, \bar{u})$ *where*

$$f(u,\rho) = \frac{1}{2}\left(\frac{u}{u_{max}}\right)^2 - \frac{u}{u_{max}} + \frac{u\rho}{u_{max}\rho_{jam}}$$

for all $0 \le \bar{\rho} \le \rho_{jam}$.