LEARNING TO CO-DRIVE... BRAIN ARCHITECTURES AND MENTAL IMAGERY MECHANISMS THAT HELP IMPROVING AGENTS FOR AUTOMATED DRIVING AND ENABLE NATURAL HUMAN-ROBOT INTERACTIONS

MAURO DA LIO - UNIVERSITÀ DEGLI STUDI DI TRENTO, ITALY

H2020 DREAMS4CARS RIA (GRANT 731593)

HOGSKOLAN I SKOVDE (SE), MIDDLESEX UNIVERSITY HIGHER EDUCATION CORPORATION (UK), THE UNIVERSITY OF SHEFFIELD (UK), DEUTSCHES FORSCHUNGSZENTRUM FUR KUNSTLICHE INTELLIGENZ (D), HEICH CONSULT (D), CENTRO RICERCHE FIAT (I)

TALK OUTLINE

1. The “dream” of Autonomous Driving
2. Engineering approaches, issues and open challenges
3. Artificial Driving Agent cognitive architectures

    Explainable safe and scalable AI

    Learning by self-instantiated simulations (“dreaming”)

DREAMS4CARS

THE “DREAM” OF AUTONOMOUS DRIVING

- **Motivations**
  - Safety. **Assumptions:** 1) humans are bad drivers, let us 2) ”automate” driving
  - Service people that cannot drive
  - Technological and market leadership

- **44 (at least) corporations were listed working on Automated Driving**
  - https://www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list/

CHALLENGES WITH AUTOMATED DRIVING

- **1) Human beings are actually very good drivers**
  - 1 fatal accident every **100 million of driven miles**; severe accident every 12 millions miles.
  - Average of all conditions and all types of driver (senior/attentive drivers are much better).
  - Target figure should be **~100 times better** (which means **400 deaths/year in the EU**).
  - **Benchmarks**
  - Tesla (level 2). 1 fatal accident after 130 million miles (but in restricted scenarios, level 2 supervised).
  - Google (level 4). Reports 69 safety-critical takeover per year (13 would be crashes) in 2015.

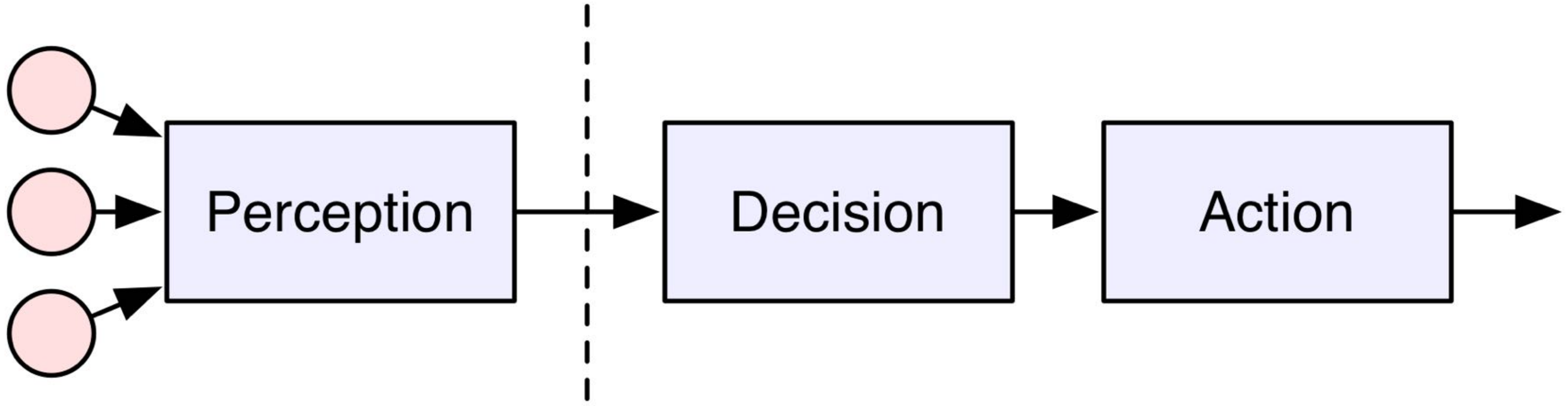
- **2) Autonomous driving needs cognition abilities**
  - Driving a car is **not** a matter of "automation"

EXAMPLE 1


sky
pole
vegetation
building
car
traffic sign
truck
road
terrain
sidewalk
wall
bicycle
fence

BEHAVIORS ARE A MATTER OF MENTAL PREDICTION

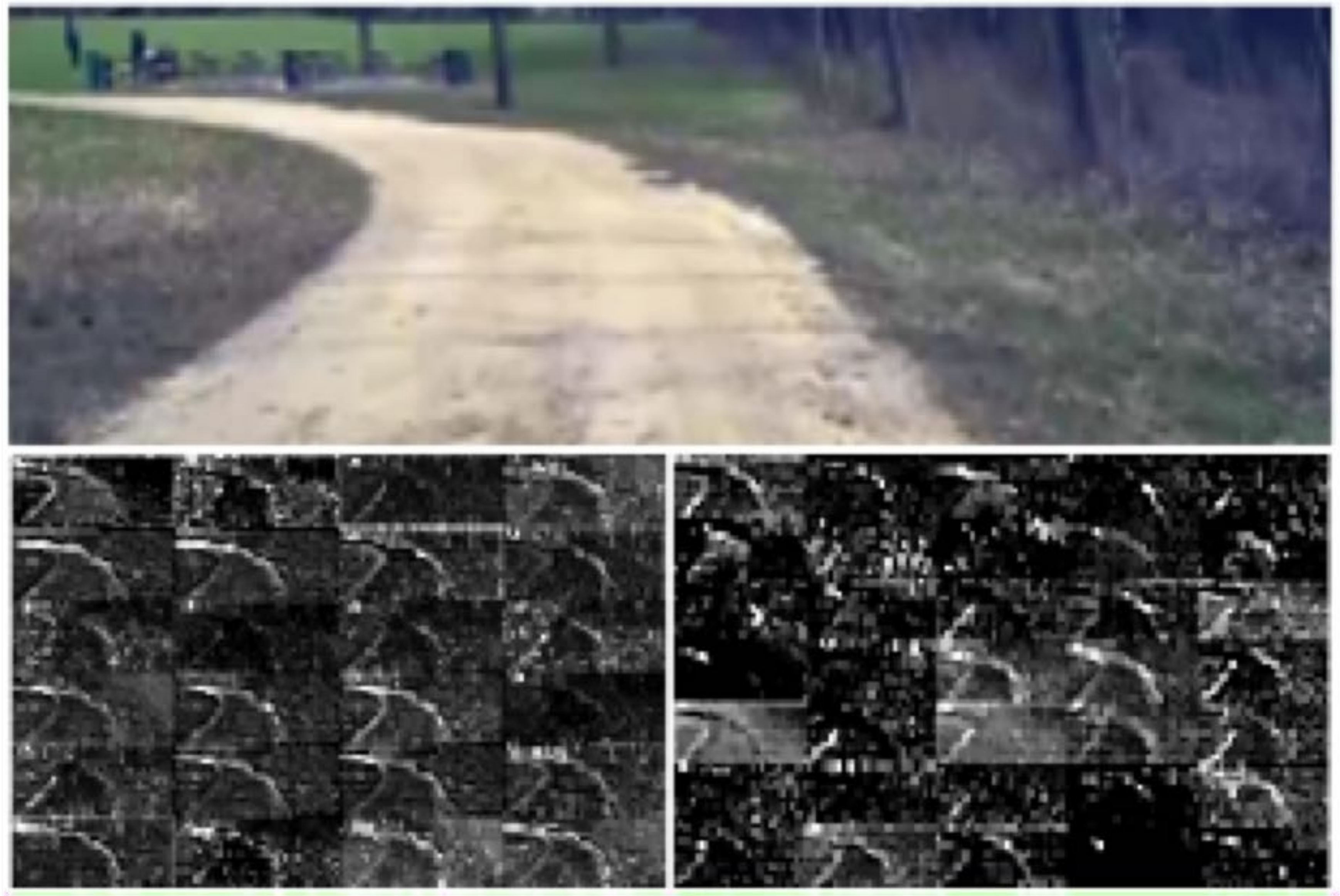
TRADITIONAL APPROACHES TO AUTOMATED DRIVING

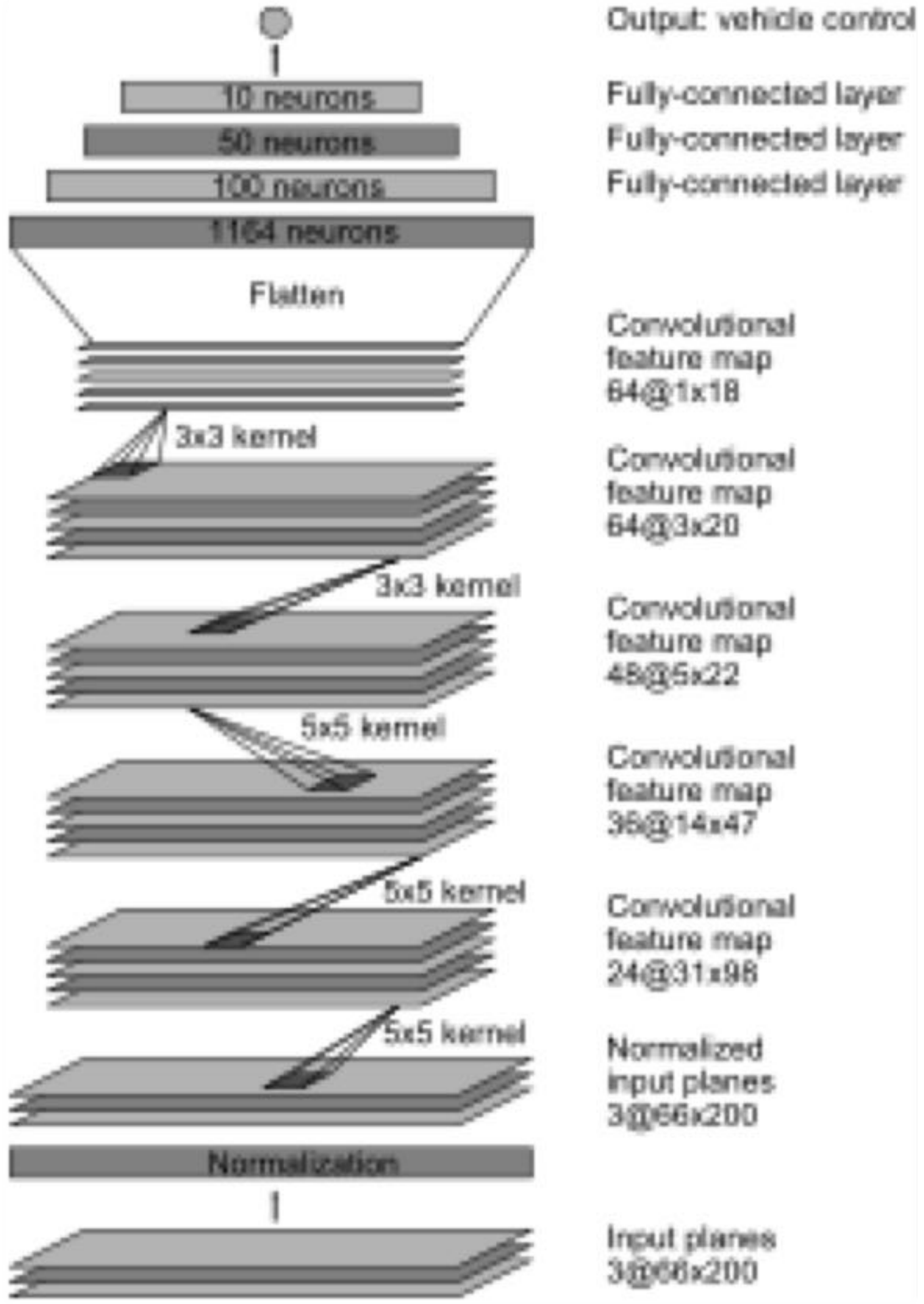


- **Hardcoded Design** (sense-think-act, maybe DL for perception only) is attractive
  - Apparently everything is under control of the designer.
    - However predicting how the system will operate in every possible complex situations is hard.
  - Lack of **autonomy**
    - In the real world a system should be able to correctly operate in situations that were **not predicted**, nor even known to the designer.

END-TO-END?

- **NVIDIA (with Autonomous Stuff vehicle).**





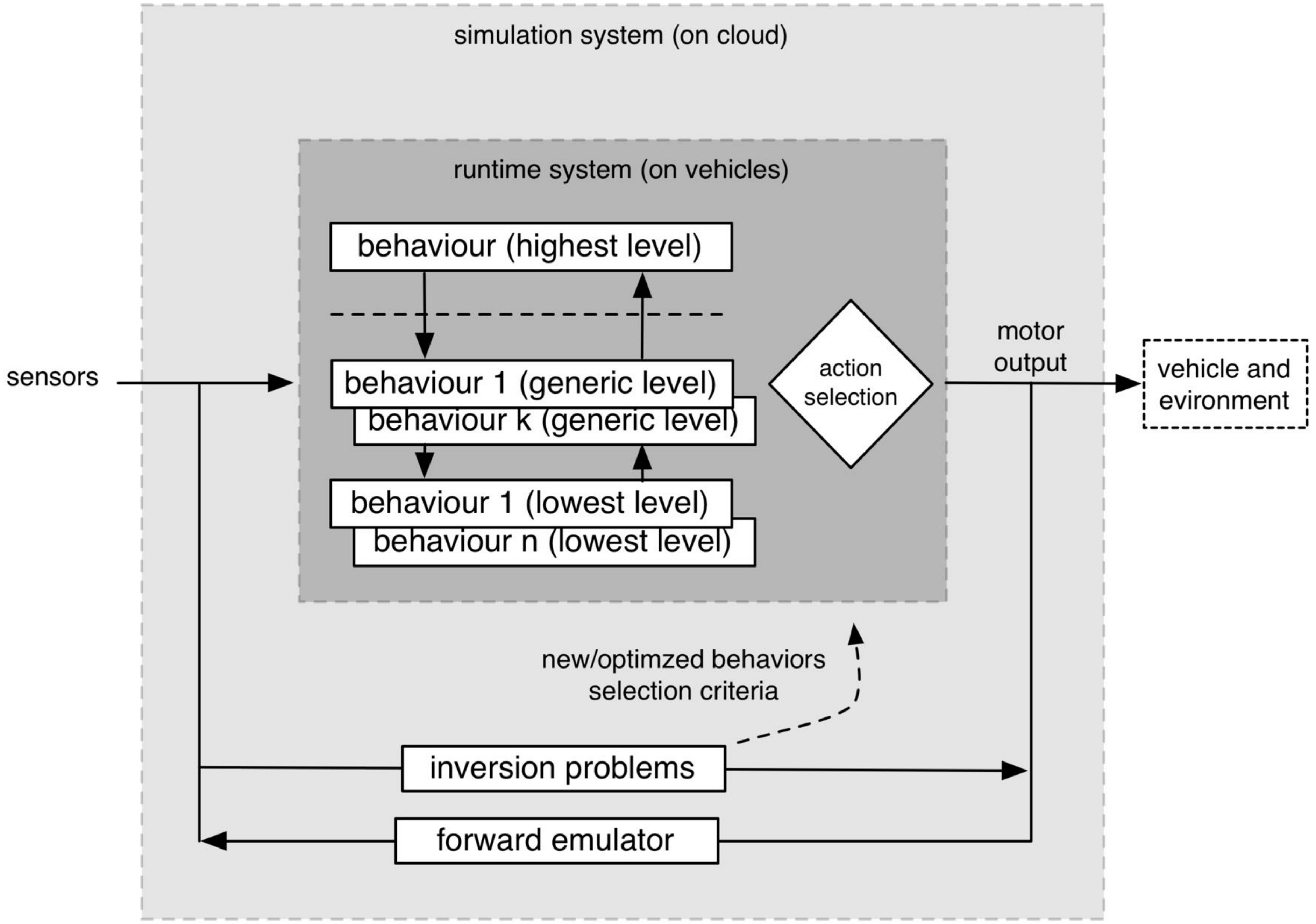
- M. Bojarski et. Al., "End to End Learning for Self-Driving Cars", arXiv: 1604.07316v1 (Apr. 16)

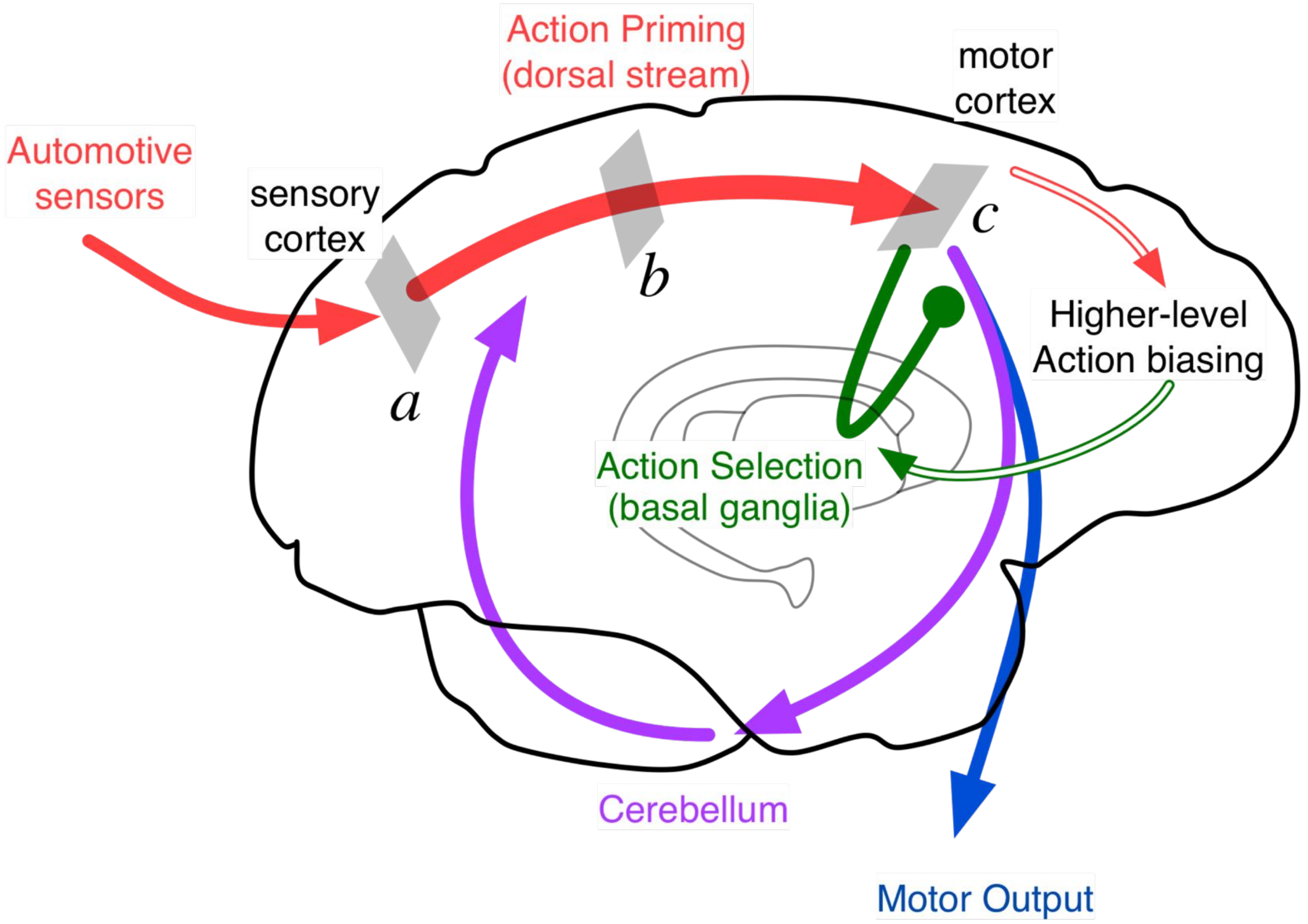
BEYOND PURE IMITATION…

- **Waymo (2018)**
  - “simple imitation of a large number of expert demonstrations is not enough to create a capable and reliable self-driving technology”
  - “Following the success of neural networks for perception, we naturally asked ourselves the question: given that we had millions of miles of driving data (i.e., expert driving demonstrations), can we train a skilled driver using a purely supervised deep learning approach?”
  - **ChauffeurNet**: learns by synthesizing suitable training data via perturbations of the expert driven trajectory
- https://medium.com/waymo/learning-to-drive-beyond-pure-imitation-465499f8bcb2

TOWARDS A “DIFFERENT” PARADIGM

- Dreams4Cars is an H2020 RIA in robotics. The main idea is developing an agent that:
  1. Learns models of the world to enable prediction abilities (both procedural and declarative predictions)
  2. Use predictions to synthetize (mostly offline) improved sensorimotor control/behaviors

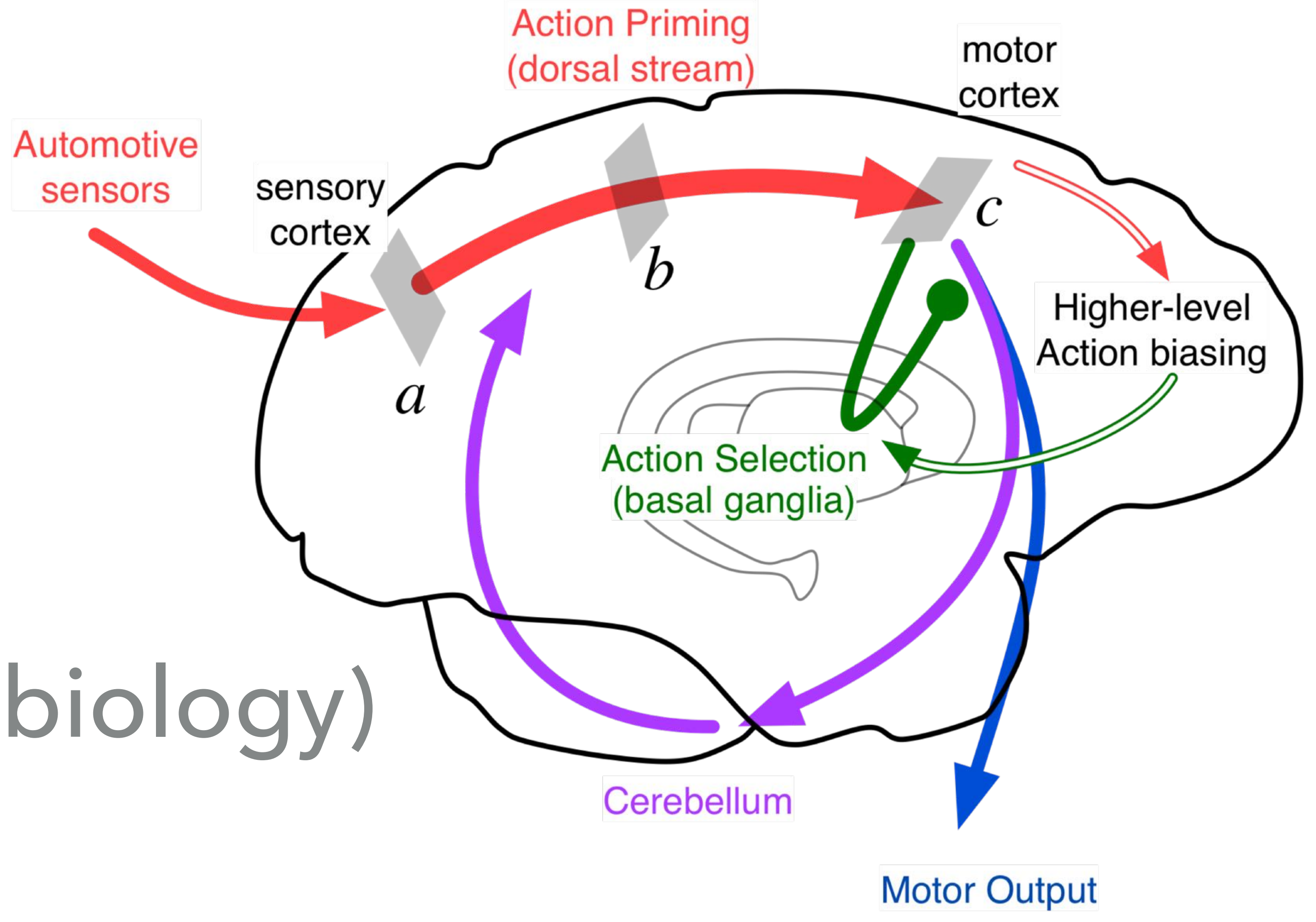


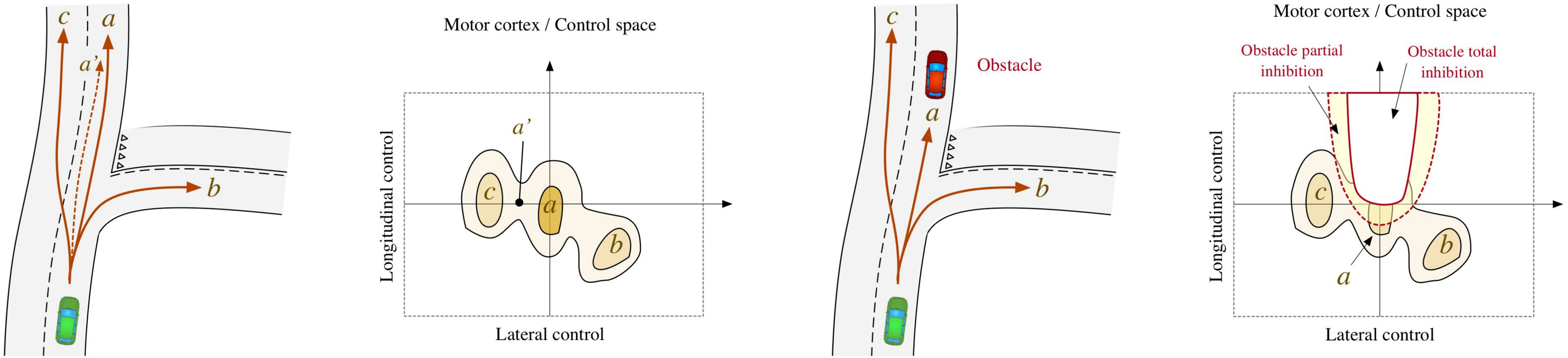


DREAMS4CARS

ACTION PRIMING

- The Dorsal Stream is the direct sensorimotor loop
  - **Parallel action instantiation**
  - **Action value is "salience"** (activation of neural patterns in biology)

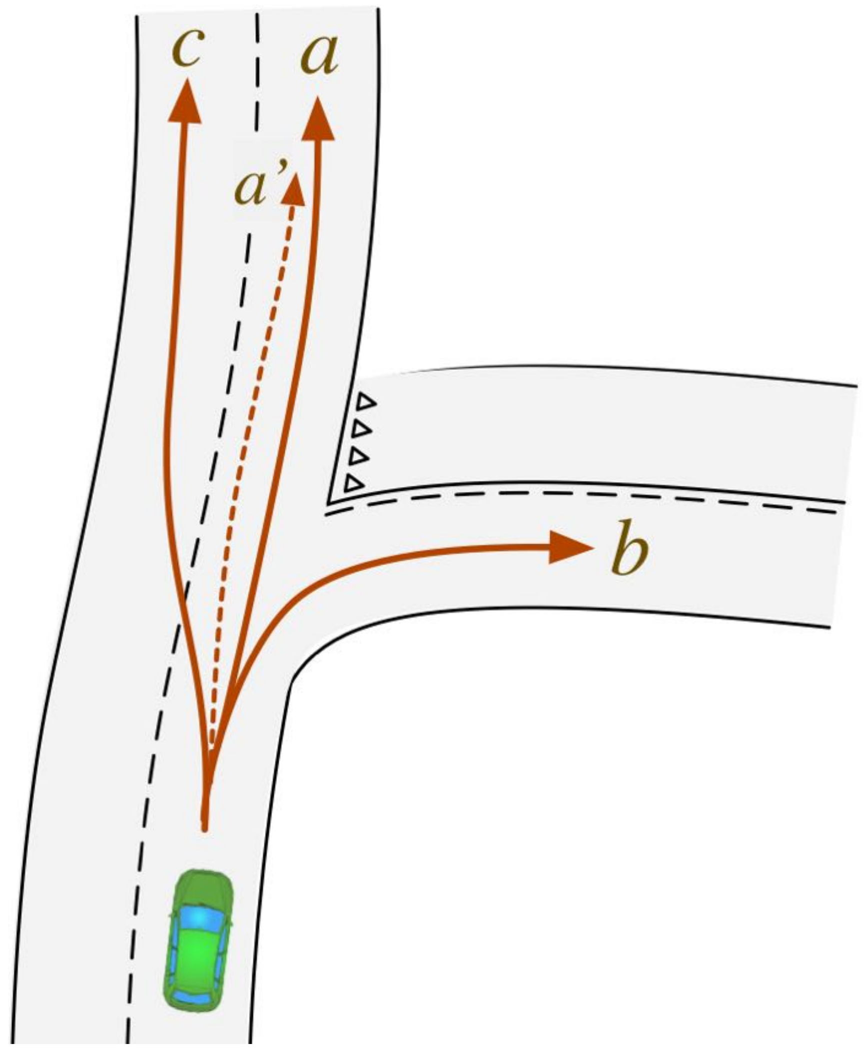


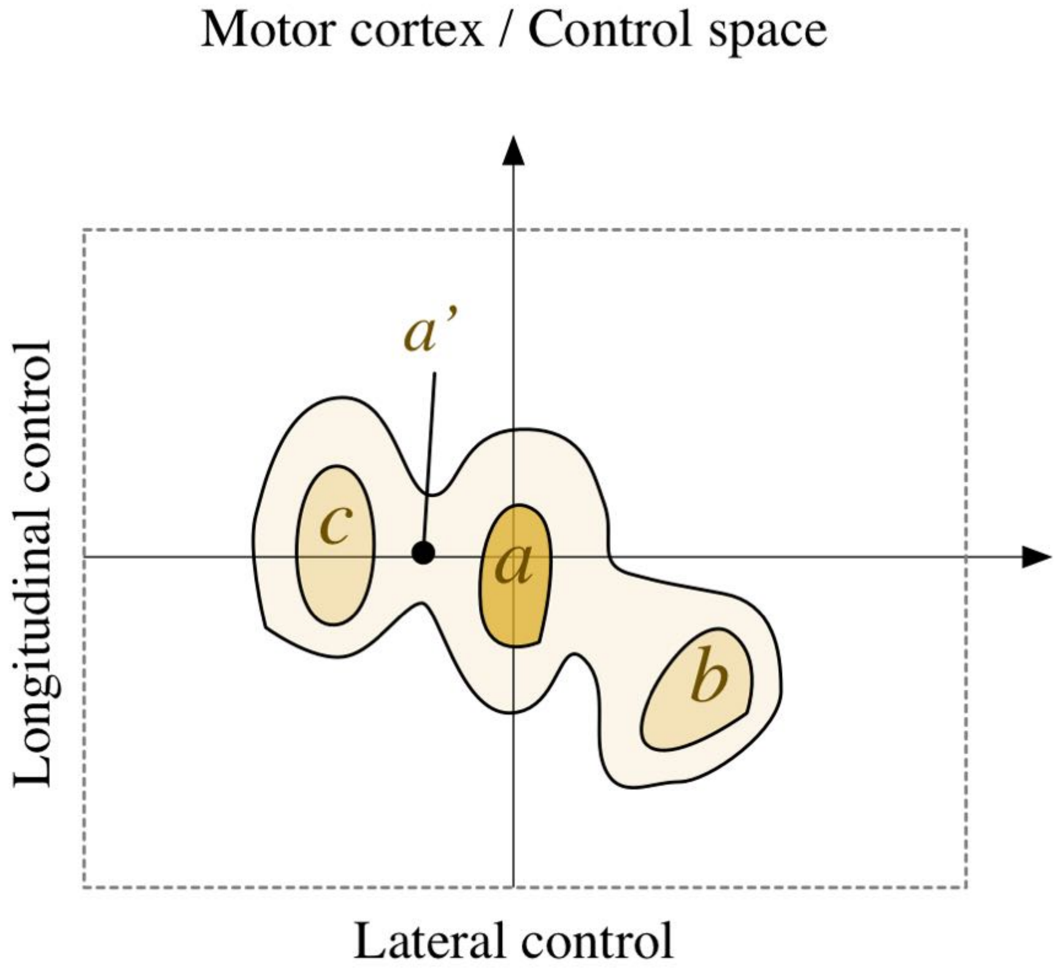


DREAMS4CARS

MODULAR STRUCTURE (EXCITATORY CIRCUITS)

- Salience may be computed **independently** for each space-time affordable location.
- Motor cortex is obtained by **overlapping** the salience of each individual goal.
- Dorsal stream **excitatory module** is a (deep) neural network that computes activation (the output tensor) for a **generic road strip** with given **vehicle forward model** and given **environmental conditions**.

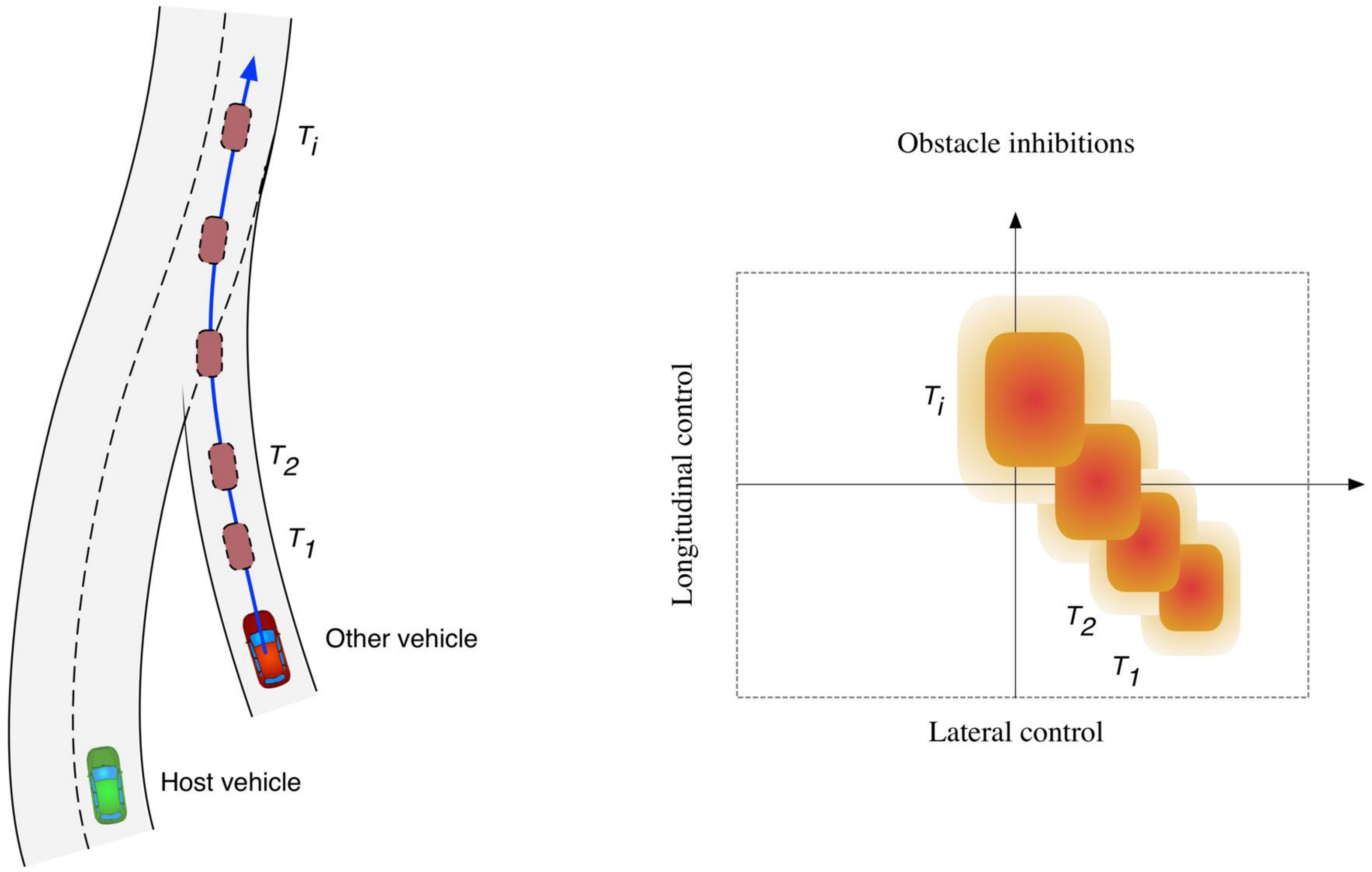


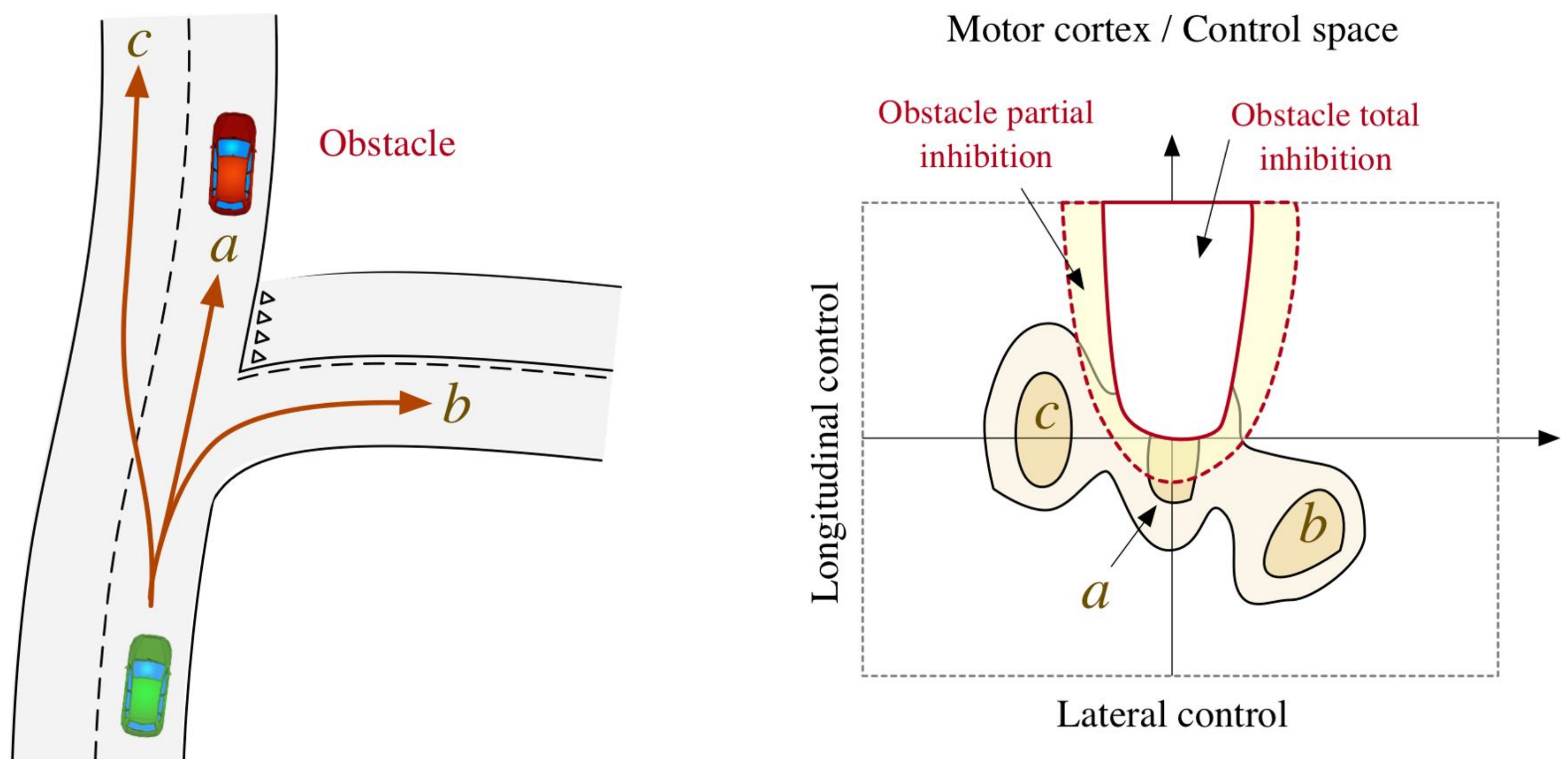


DREAMS4CARS

MODULAR STRUCTURE (INHIBITORY CIRCUITS)

- Dorsal stream basic **inhibitory module** is a (deep) neural network that computes inhibition (the output tensor) for a generic **space-time location** with given **vehicle forward model** and given **environmental conditions**.
- Prediction of obstacle trajectory from a different module.

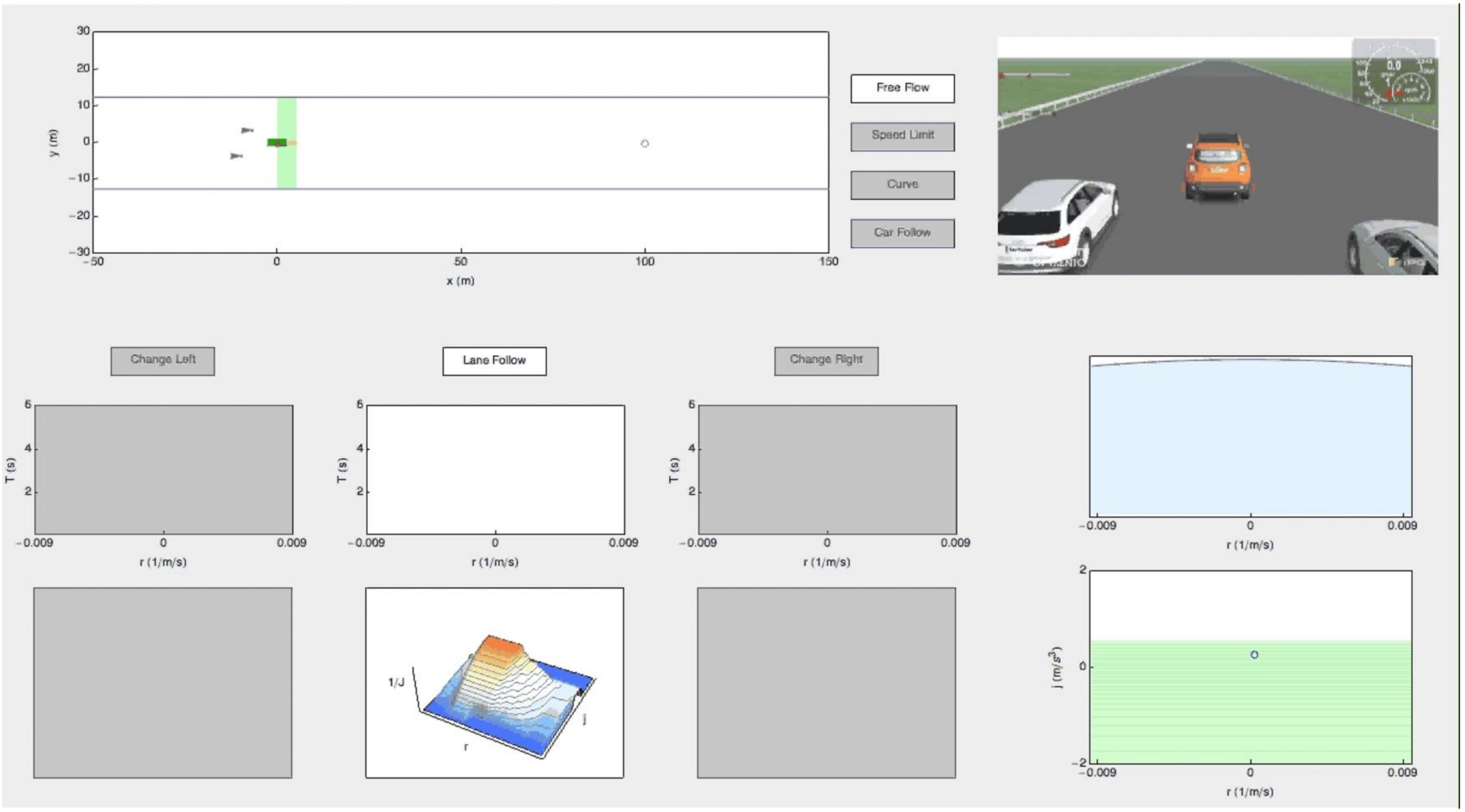




- A. Plebe, M. Da Lio, D. Bortoluzzi, "**On Reliable Neural Network Sensorimotor Control in Autonomous Vehicles**," IEEE Trans. Int. Transportation Sys, in press, 2019.

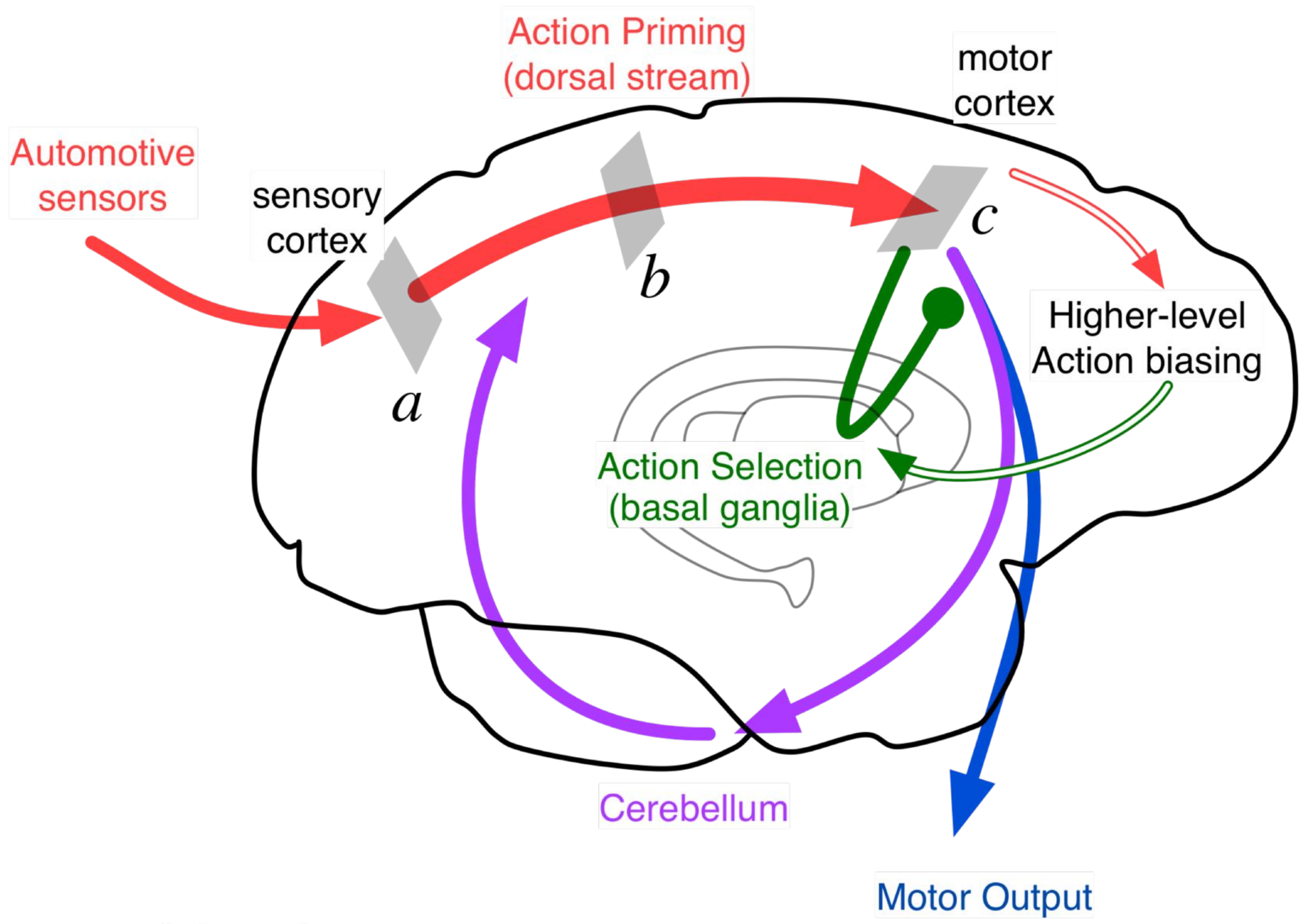
EXPLAINABLE AI (AT SYSTEM LEVEL)

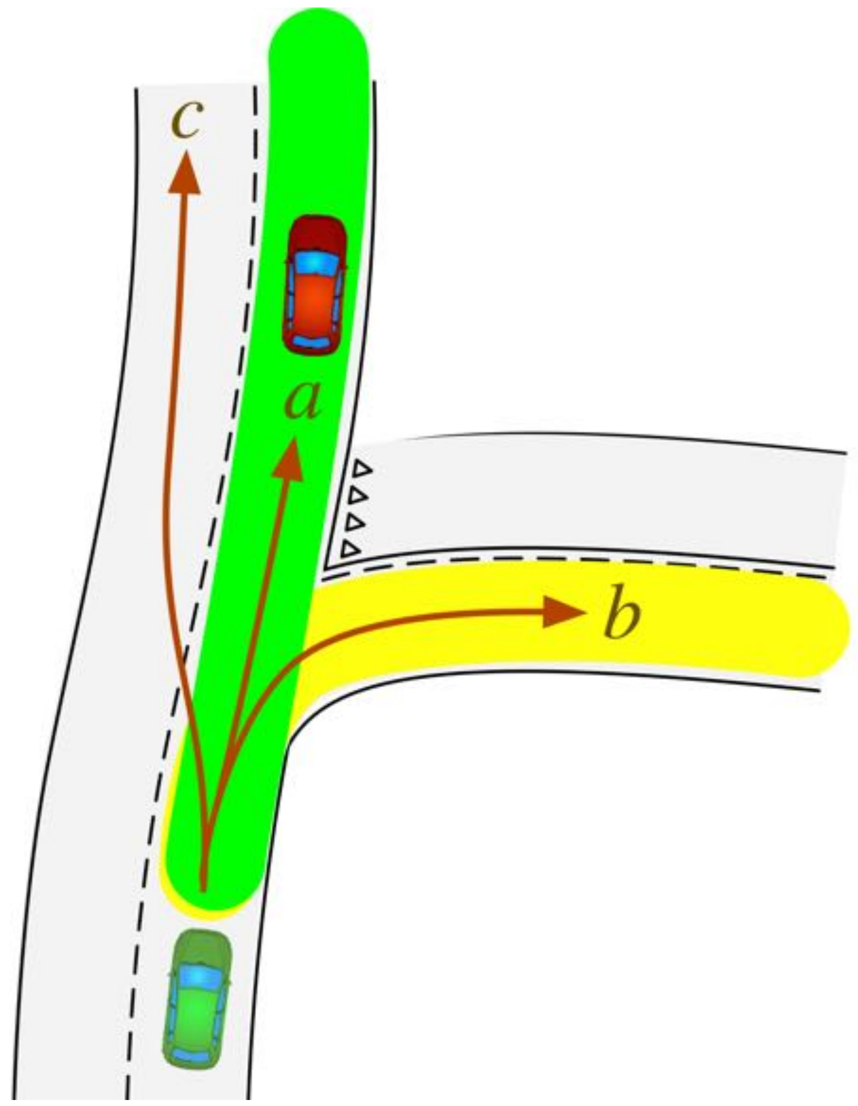
- **The sensorimotor system is interpretable** by inspection of the motor cortex.
  - One can always say which actions were instantiated and why one particular action was selected.
  - As a matter of fact this is the way the agent is debugged.

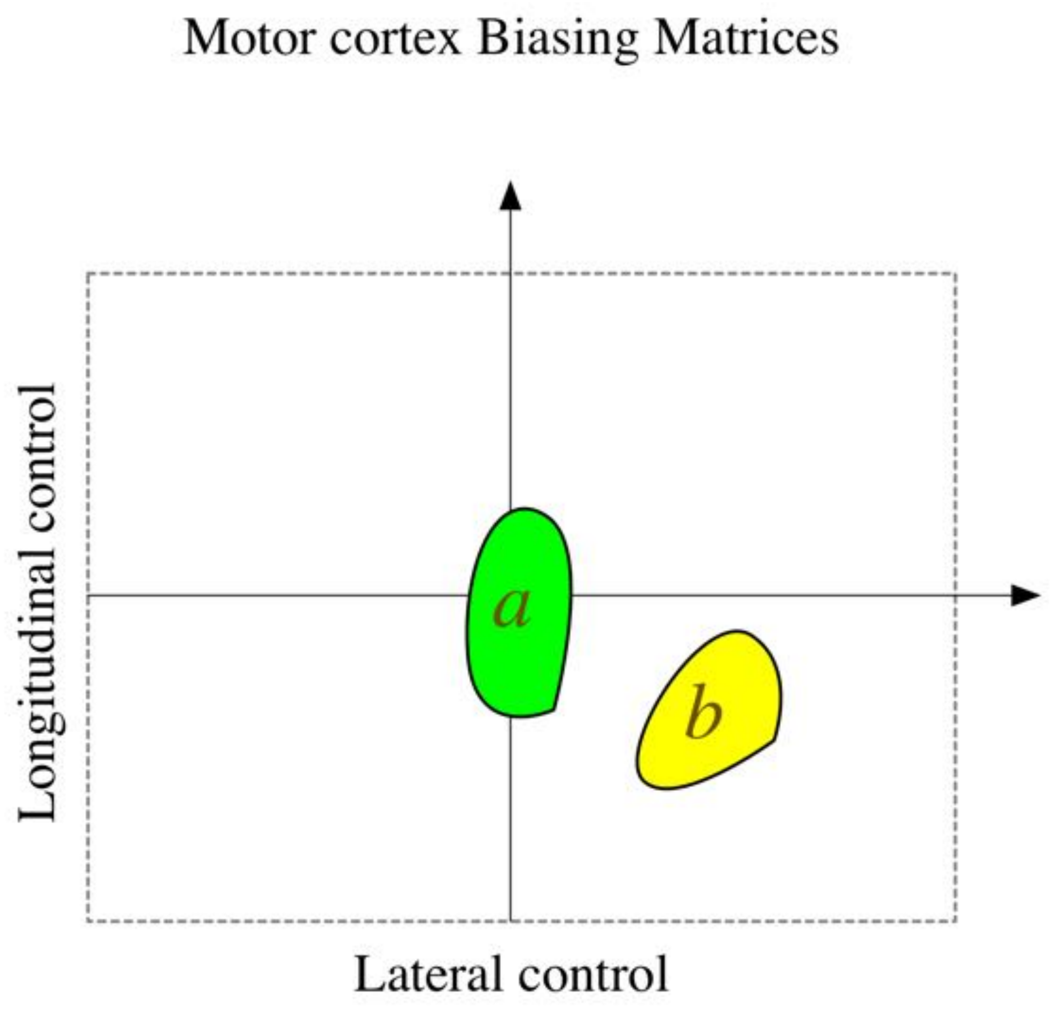


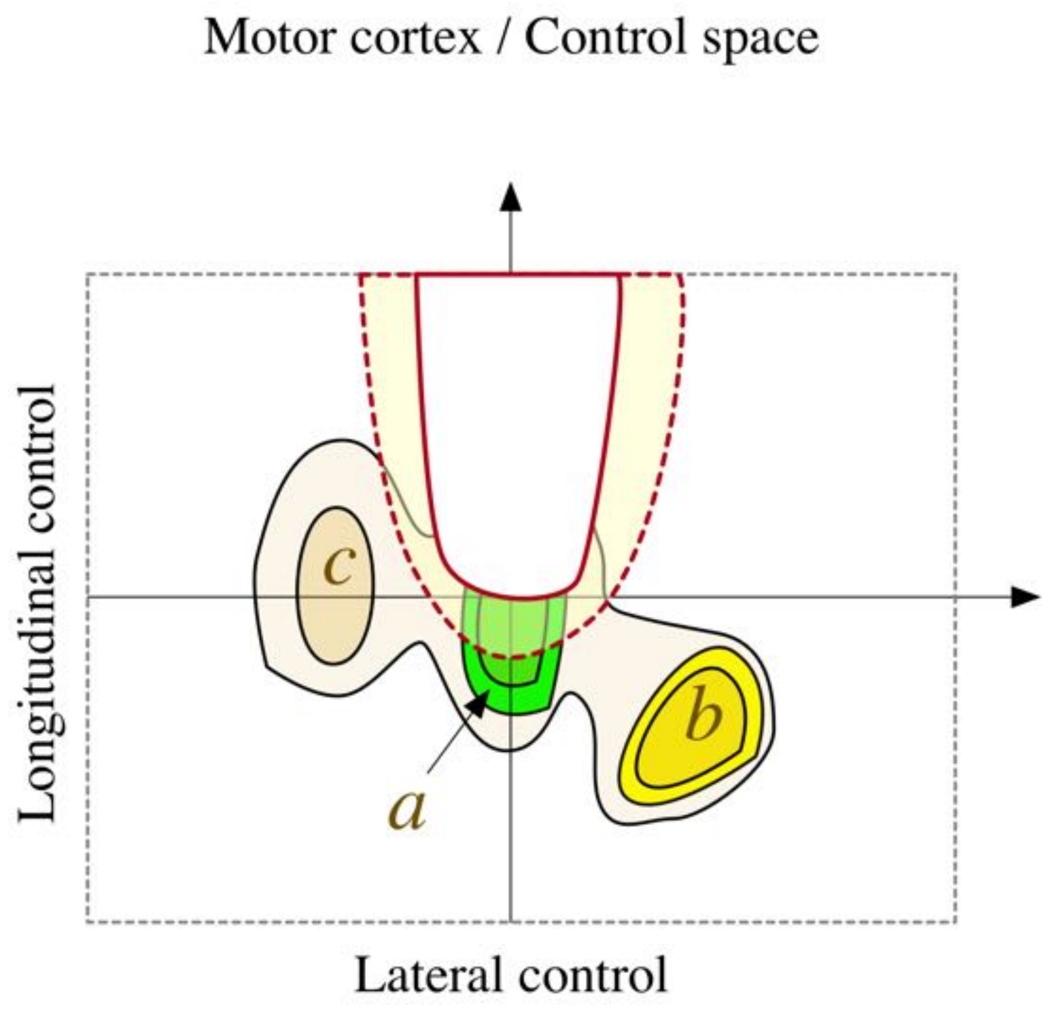
ARTIFICIAL IMPLEMENTATION OF BIASING MECHANICS

- Long-term goals **bias** proximal action by **artificially increasing/decreasing the salience** at lower level competition.
- The low-level may always **veto** higher-level directives (post bias salience may be insufficient).



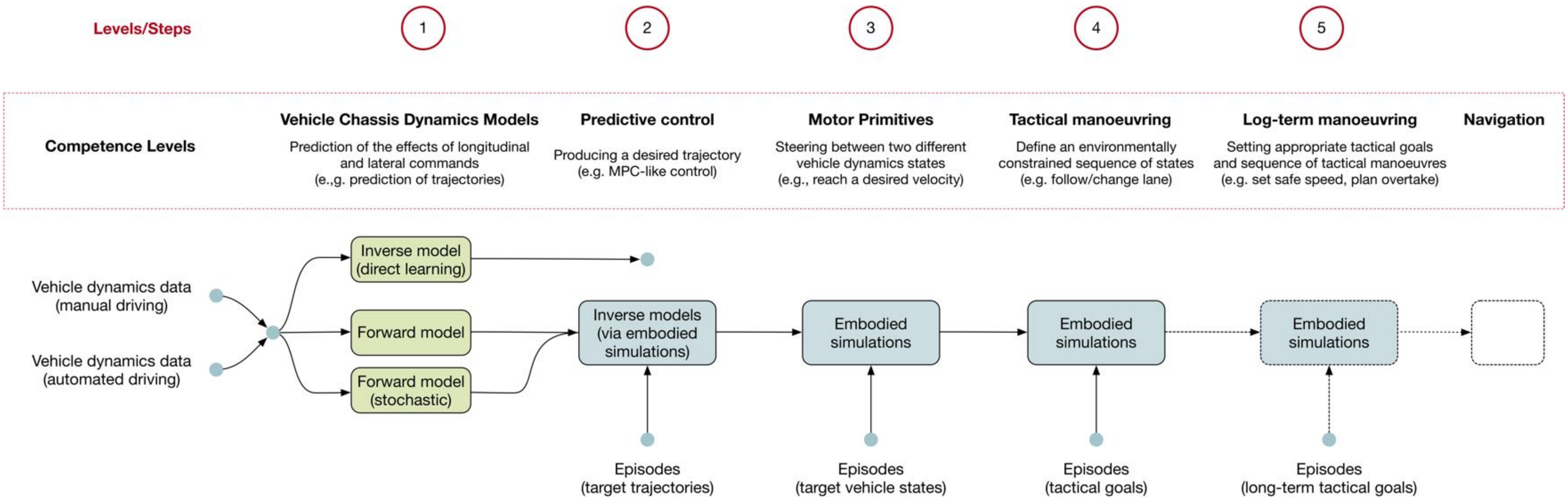






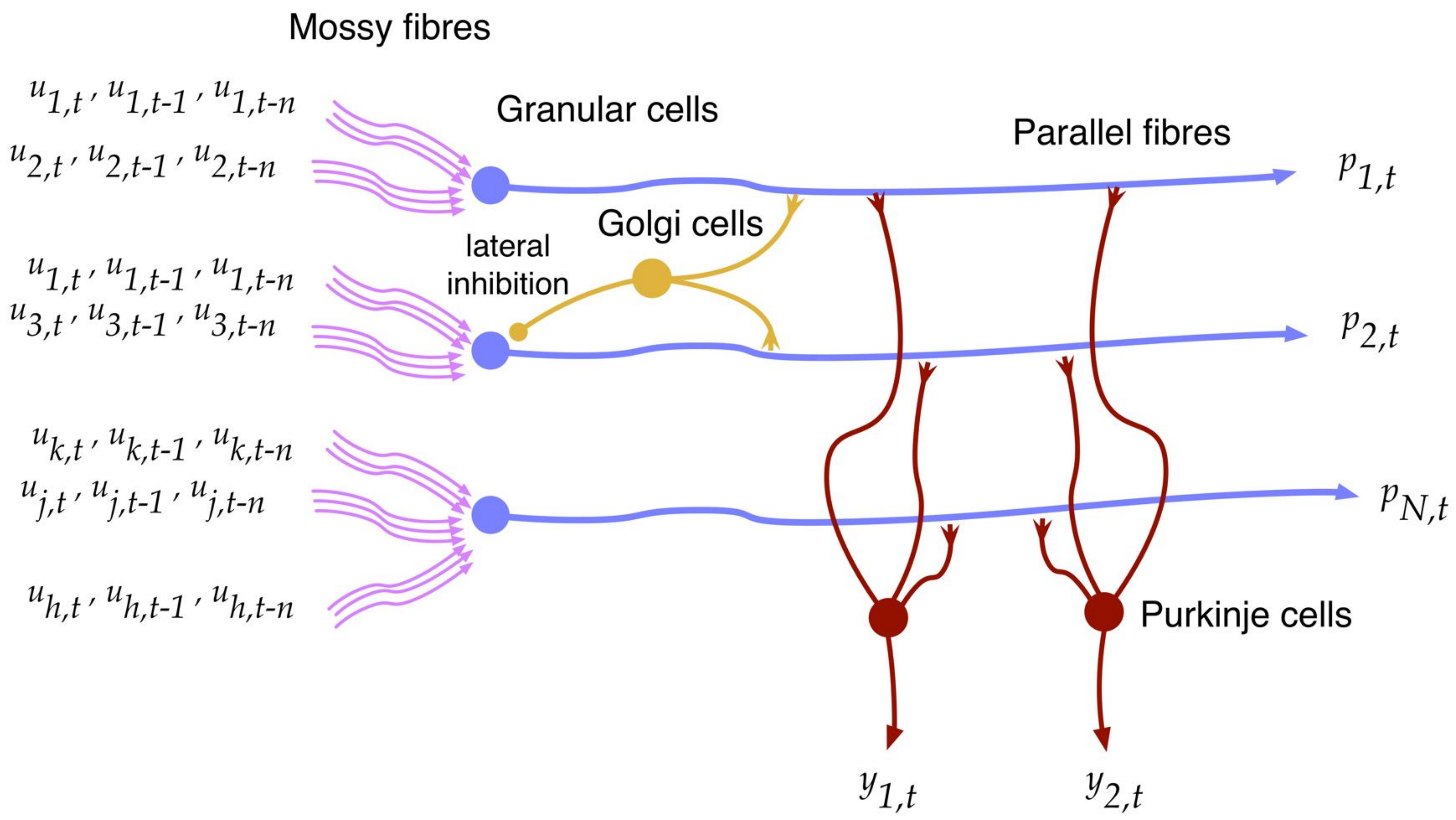
DREAMS4CARS

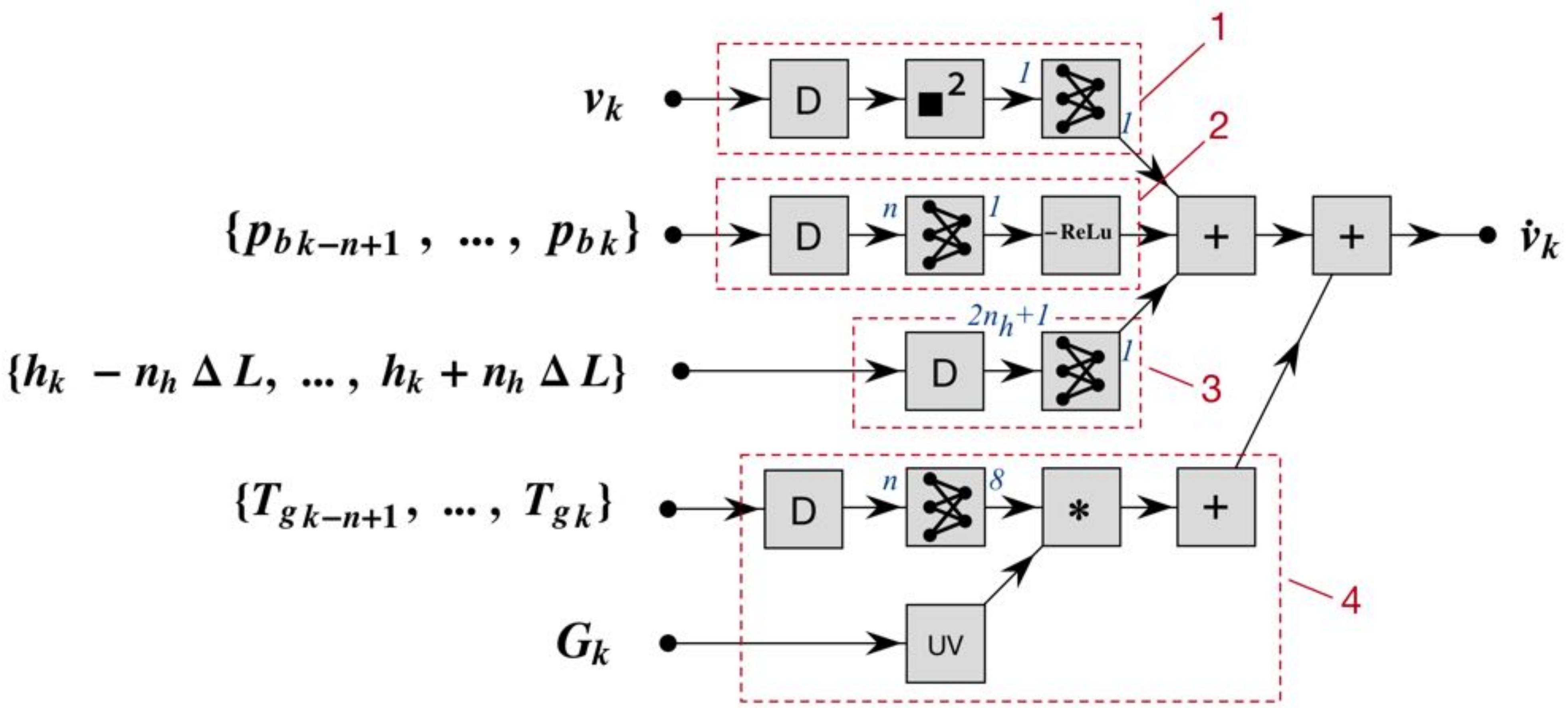
STRATEGY FOR BOOTSTRAPPING SENSORIMOTOR ABILITIES



LEVEL 1 – LEARNING FORWARD AND INVERSE MODELS

- The Cerebellum has a **specialized micro-structure**
  - Effective into learning (weak) **superposition of effects** and dynamical systems
  - **Sample efficient**, albeit **biased towards** learning **the real world physics**



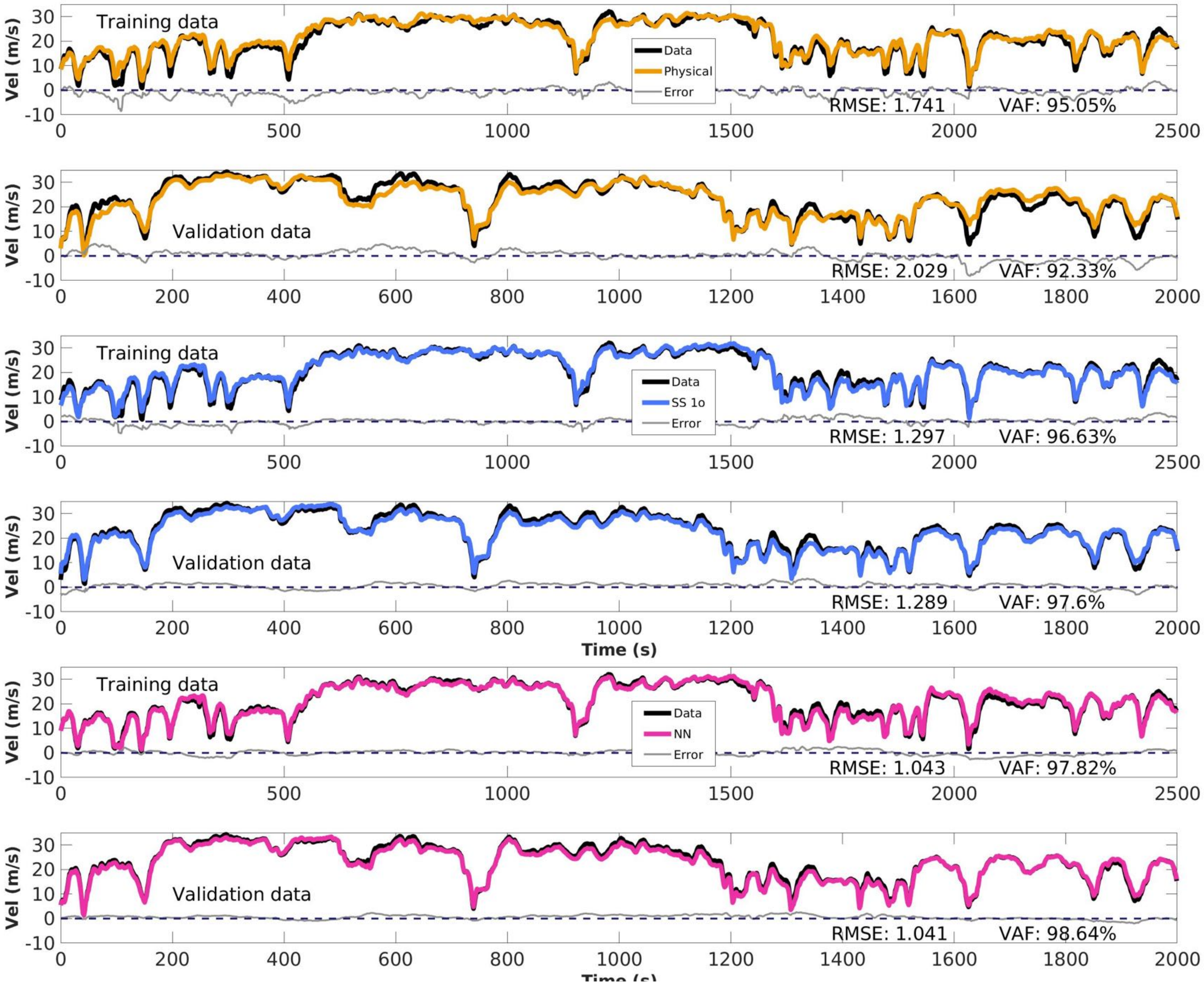


- S. James, S. Anderson, M. Da Lio, "Longitudinal Vehicle Dynamics: A Comparison of Physical and Data-Driven Models Under Large-Scale Real-World Driving Conditions," VSD (submitted), 2019

DREAMS4CARS

PERFORMANCE VS. OTHER TYPES OF MODELS

- Comparison with Analytical and SS models

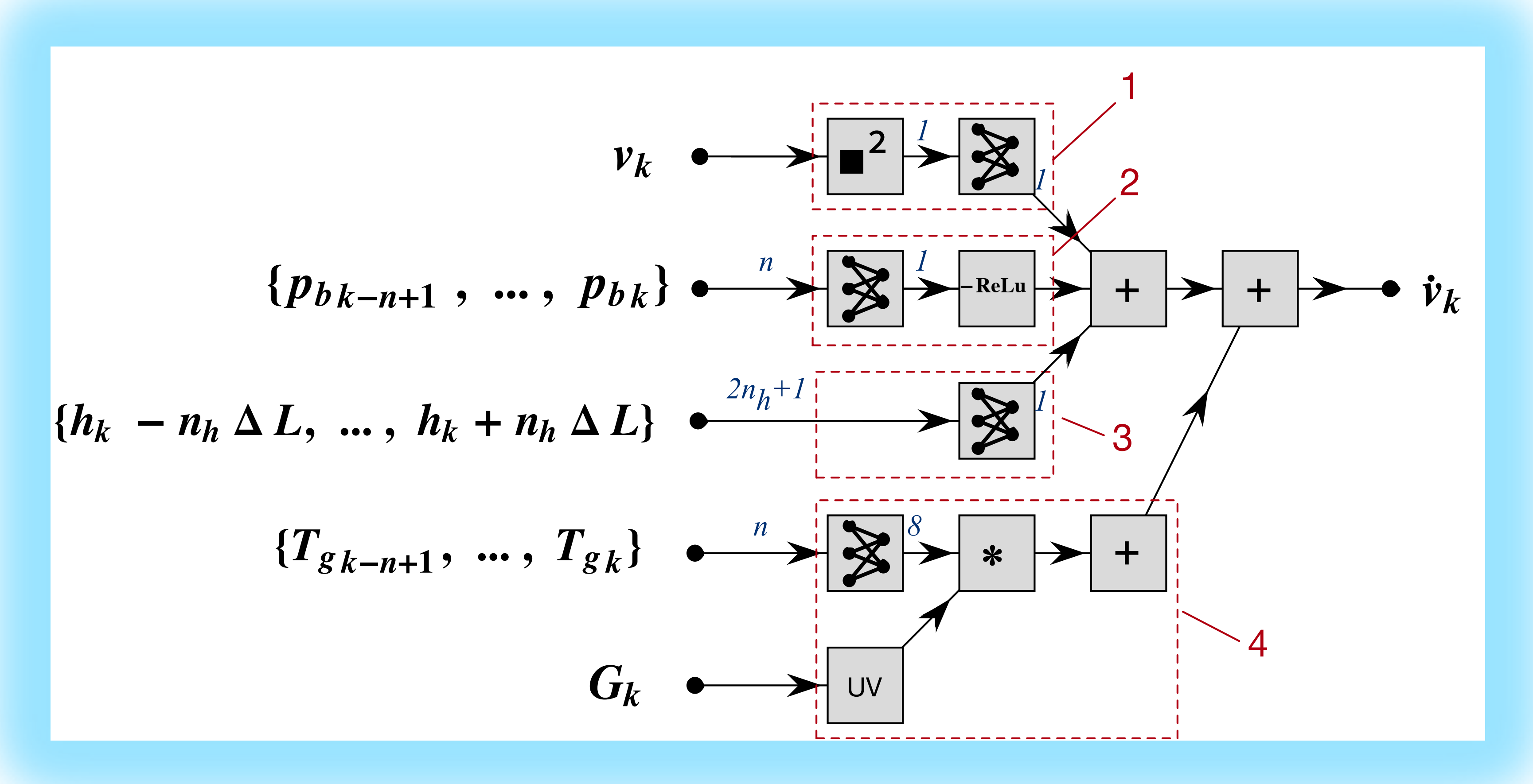


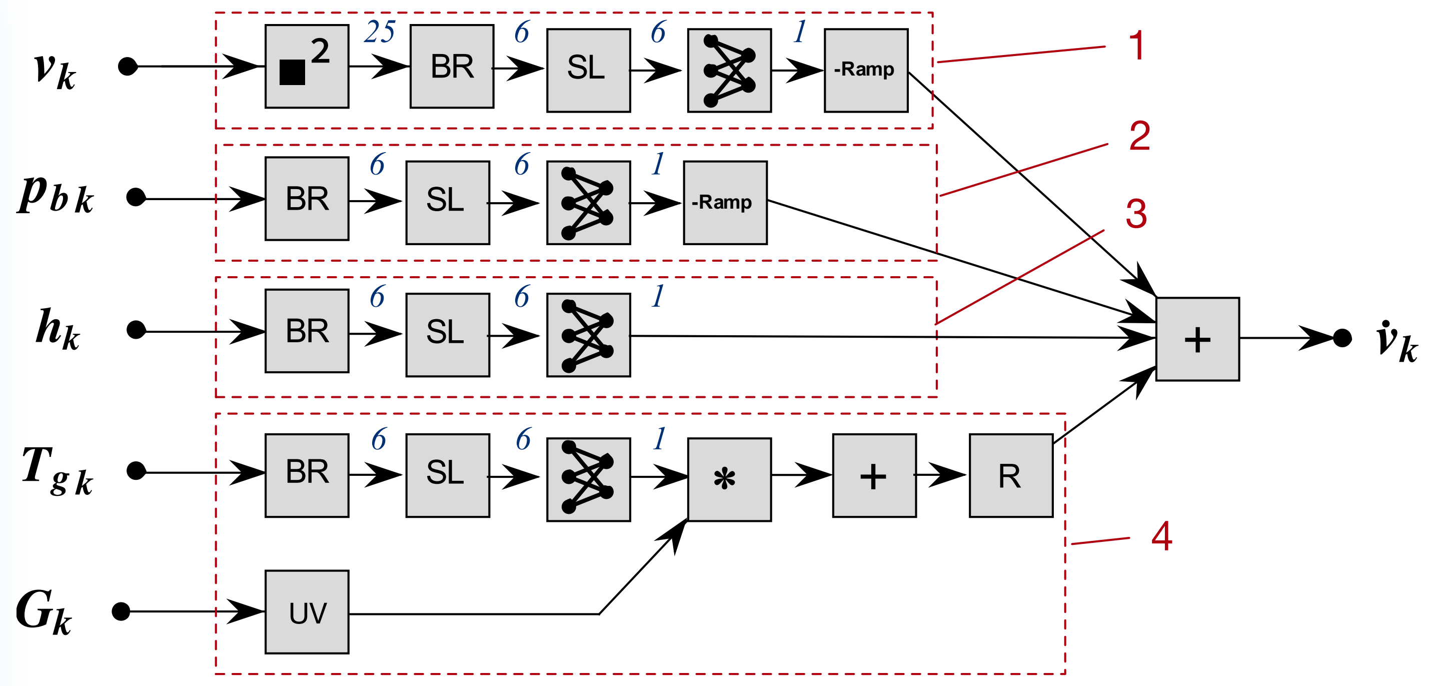
- S. James, S. Anderson, M. Da Lio, "Longitudinal Vehicle Dynamics: A Comparison of Physical and Data-Driven Models Under Large-Scale Real-World Driving Conditions," VSD (submitted), 2019

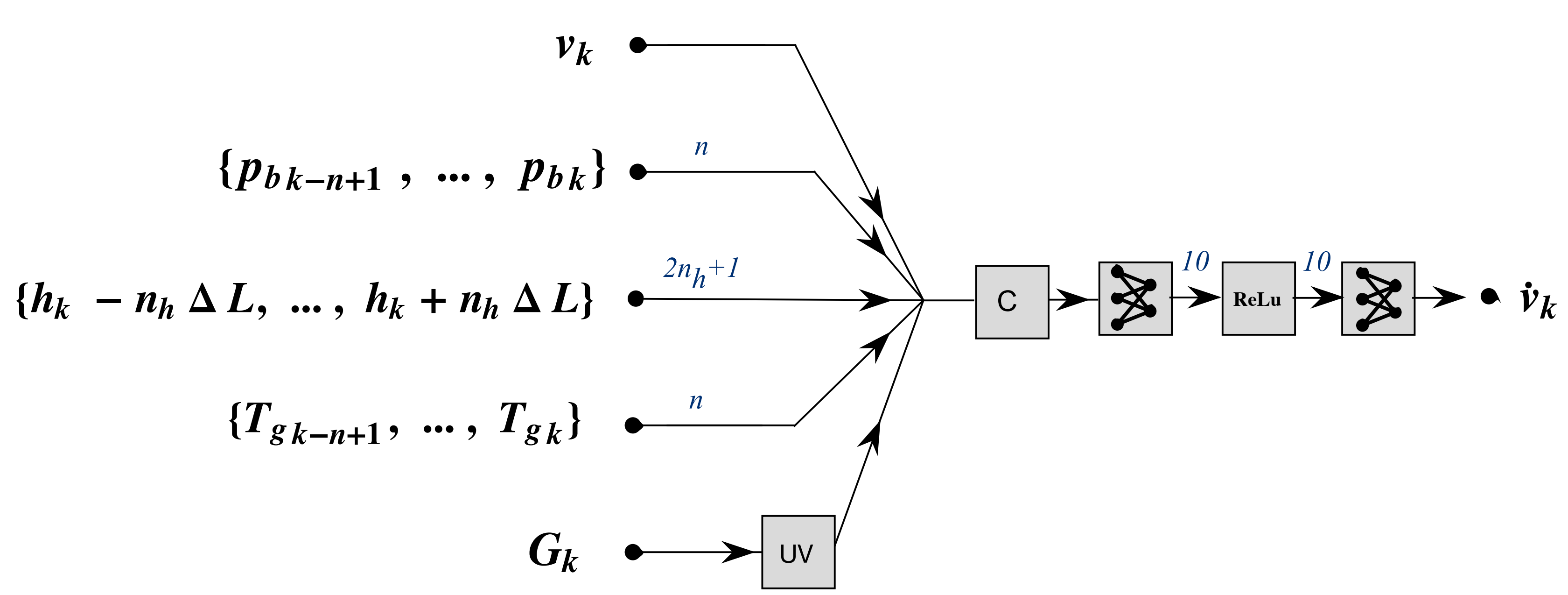
DREAMS4CARS

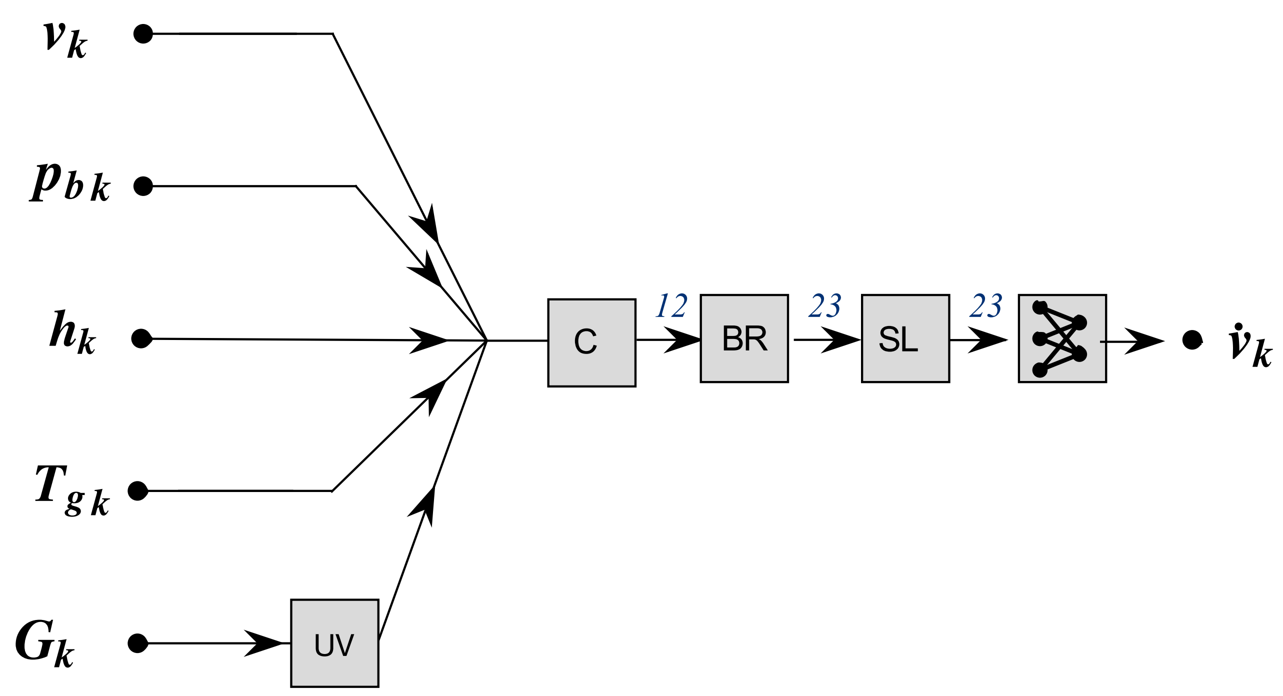
PERFORMANCE VS. OTHER TYPES OF NEURAL NETWORKS

- Comparison of networks of different architecture (recurrent versus **non-recurrent, with structure** vs with generic connections)









- M. Da Lio, D. Bortoluzzi, G.P. Rosati Papini "Modeling Longitudinal Vehicle Dynamics with Neural Networks," VSD, 2019

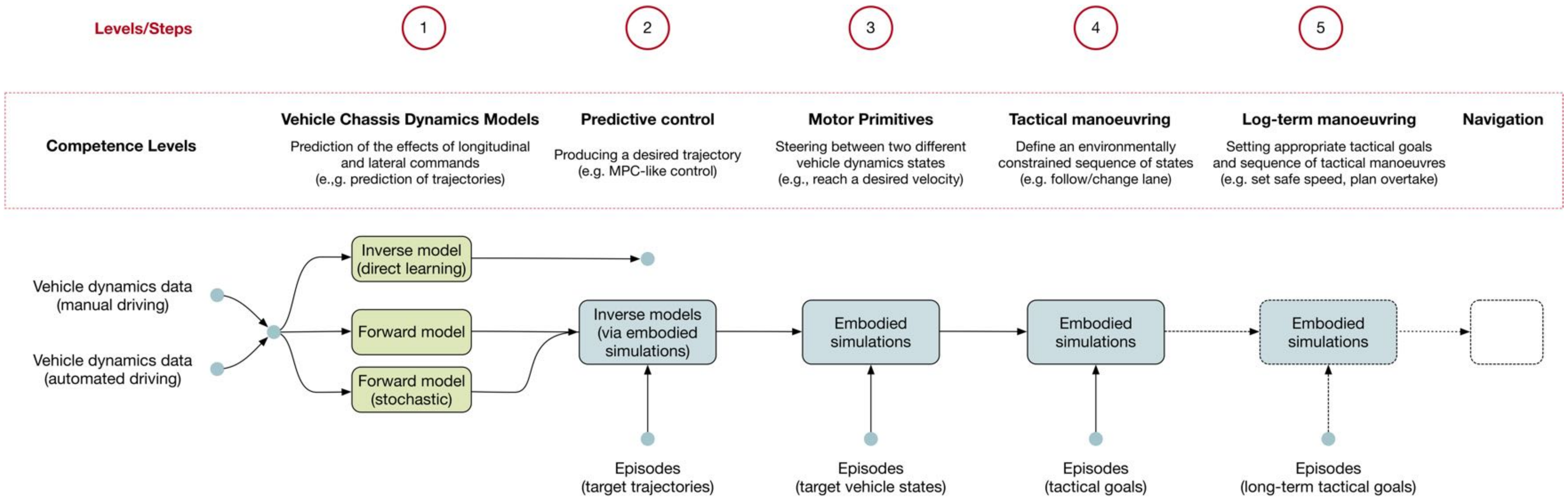
LEVELS 2+

- This part is omitted (still confidential)

EARLY MOVES...

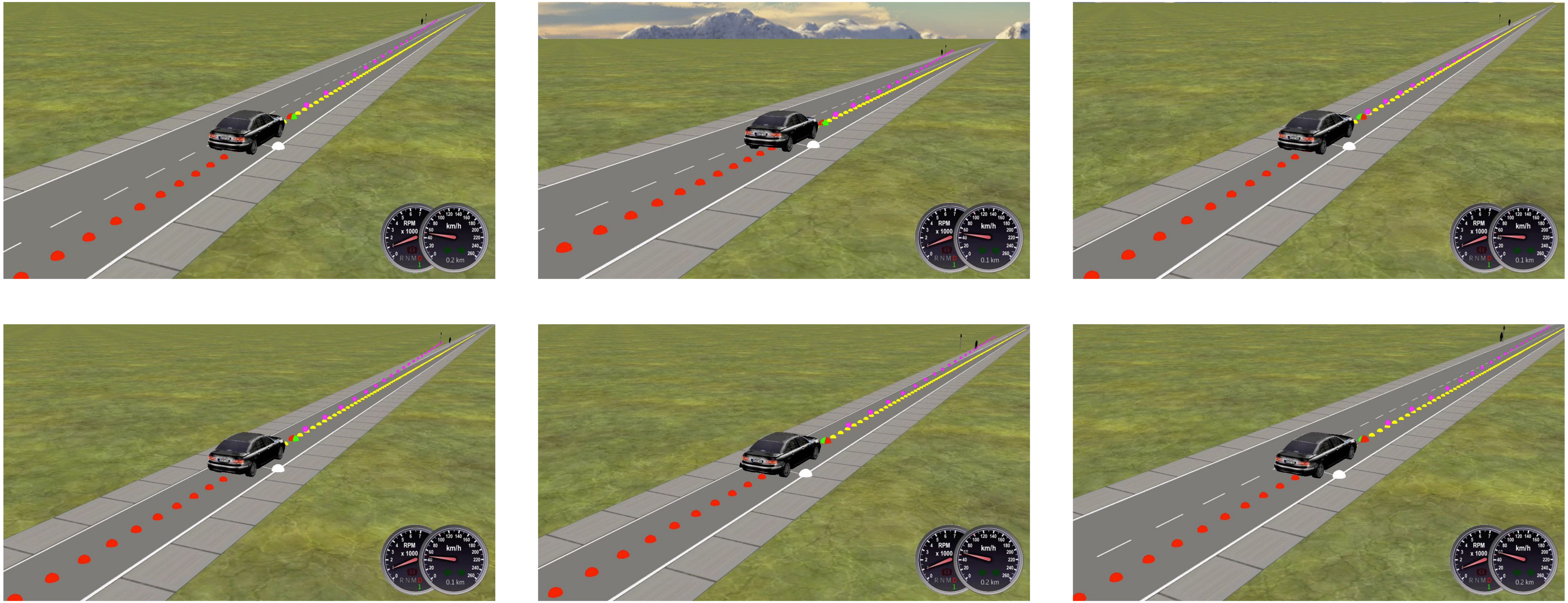
- **DFKI tests at ATC test track** (emergent behaviour)

LEVELS 4–5 VIA REINFORCEMENT LEARNING



- RL may be used to learn safe high-level behaviours
  - By learning biases for action selection
  - By learning some hyper-parameters for driving (e.g. the safe speed)

EXAMPLE LEARNING SAFE SPEED IN PEDESTRIAN CROSSINGS



DREAMS4CARS

CONCLUSIONS

1. Multi-loop agent architecture (functionally bioinspired)
2. Explainable AI (and robust, safe, modular, scalable, economical…)
3. Trained by learning predictive models that are then manipulated to synthetize sensorimotor behaviours at various levels
4. Constructs progressive abstractions of actions
5. RL efficiently integrated on top
6. More: see posters by Alice, Riccardo and Sara