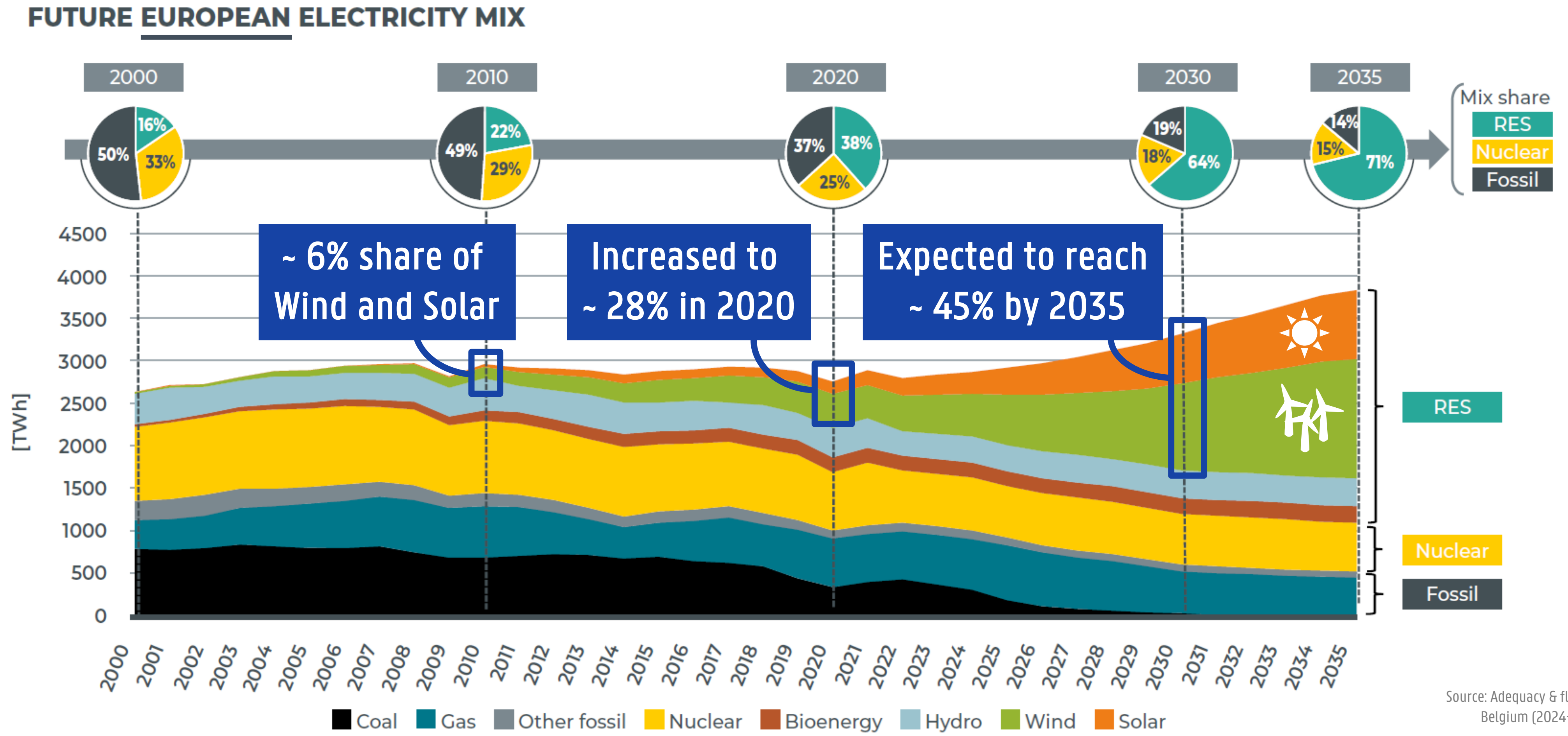
Scalable Attention-based Reinforcement Learning Method for Multi-asset Control

Seyed Soroush Karimi Madahi, Giuseppe Gabriele, Bert Claessens, Chris Develder
AI4E Team, IDLab, Ghent University – imec

GHENT UNIVERSITY

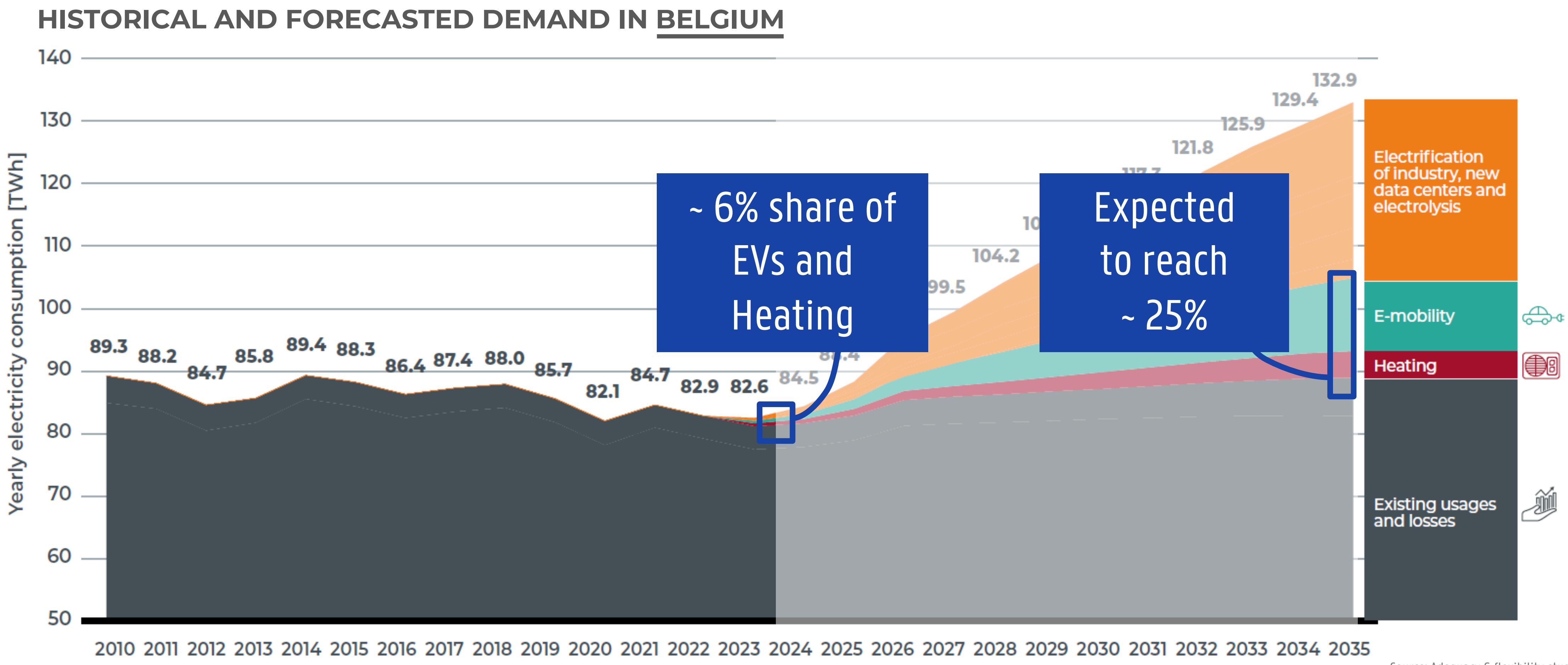
IDLab
INTERNET & DATA LAB
imec

Energy Transition – Supply



Source: Adequacy & flexibility study for Belgium (2024-2034), ELIA

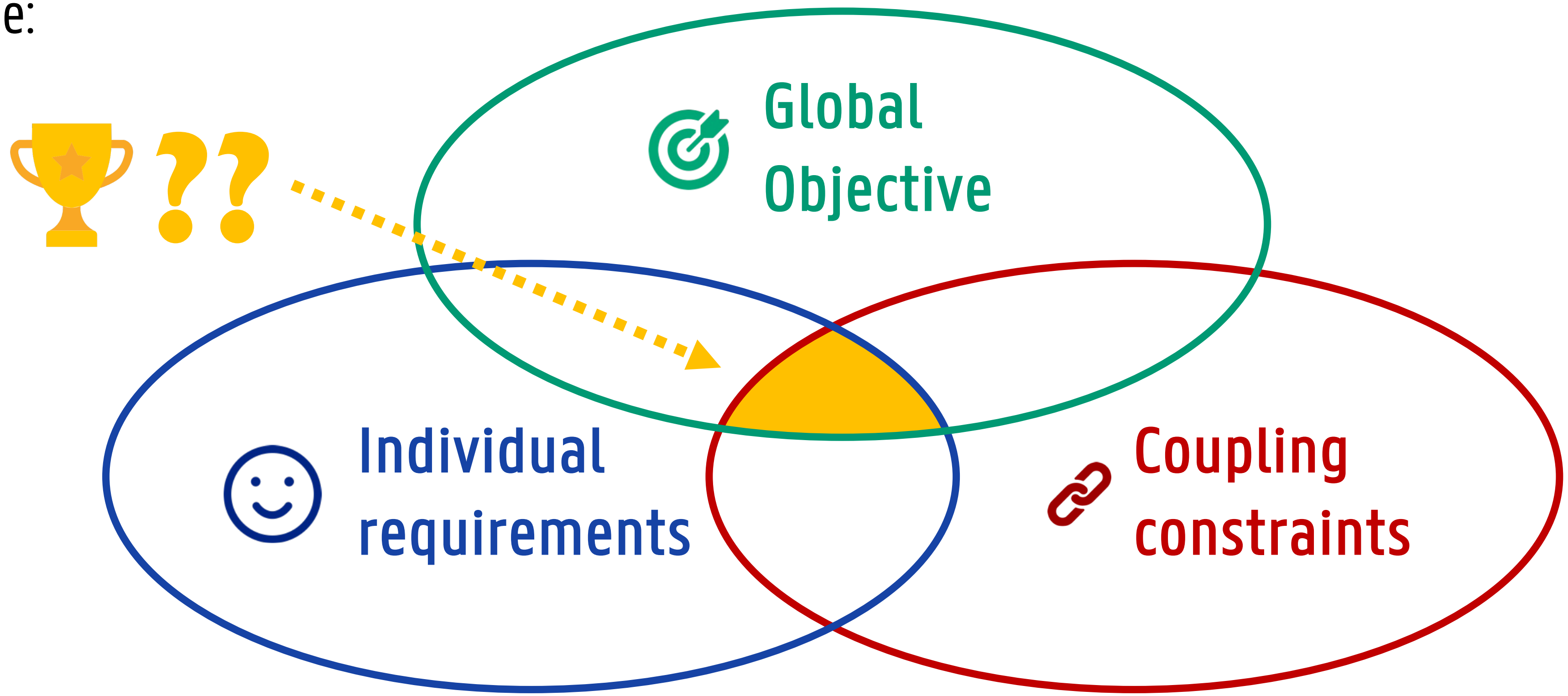
Energy Transition – (Residential) Demand



Source: Adequacy & flexibility study for Belgium (2024-2034), ELIA

Opportunity: Exploit flexibility of (aggregated!) residential assets

- We need algorithms for multi-asset control
- Challenge:



S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Opportunity: Exploit flexibility of (aggregated!) residential assets

- We need algorithms for multi-asset control
- Challenge: Achieve global objective + Joint constraints + Individual constraints
- Candidate algorithms:

Requirements	Optimization-based	Centralized RL	Distributed RL
View	Global ✓	Global ✓	Local ✗
Scalable to large pools	No ✗	No ✗	Yes ✓
System model required	Yes ✗	No ✓	No ✓
Forecaster required	Yes ✗	No ✓	No ✓
Inference time	High ✗	Low ✓	Low ✓

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Opportunity: Exploit flexibility of (aggregated!) residential assets

- We need algorithms for multi-asset control
- Challenge: Achieve global objective + Joint constraints + Individual constraints
- Candidate algorithms:

Requirements	Optimization-based	Centralized RL	Distributed RL	??
View	Global ✓	Global ✓	Local ✗	Global ✓
Scalable to large pools	No ✗	No ✗	Yes ✓	Yes ✓
System model required	Yes ✗	No ✓	No ✓	No ✓
Forecaster required	Yes ✗	No ✓	No ✓	No ✓
Inference time	High ✗	Low ✓	Low ✓	No ✓

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Opportunity: Exploit flexibility of (aggregated!) residential assets

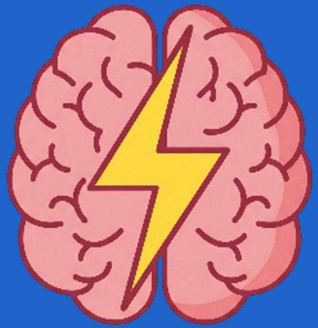
- We need algorithms for multi-asset control
- Challenge: Achieve global objective + Joint constraints + Individual constraints
- Candidate algorithms:

Requirements	Optimization-based	Centralized RL	Distributed RL	??
View	Global √	Global √	Local ✗	Global √
Scalable to large pools	No ✗	No ✗	Yes √	Yes √
Syste				No √
Fore				No √
Infer				No √

Our contributions:

1. Attention-based RL = scalable control for asset pool
2. Proof-of-principle validation for EV charging case

Attention-Based RL



Attention-based state representations for scalable RL

Idea:

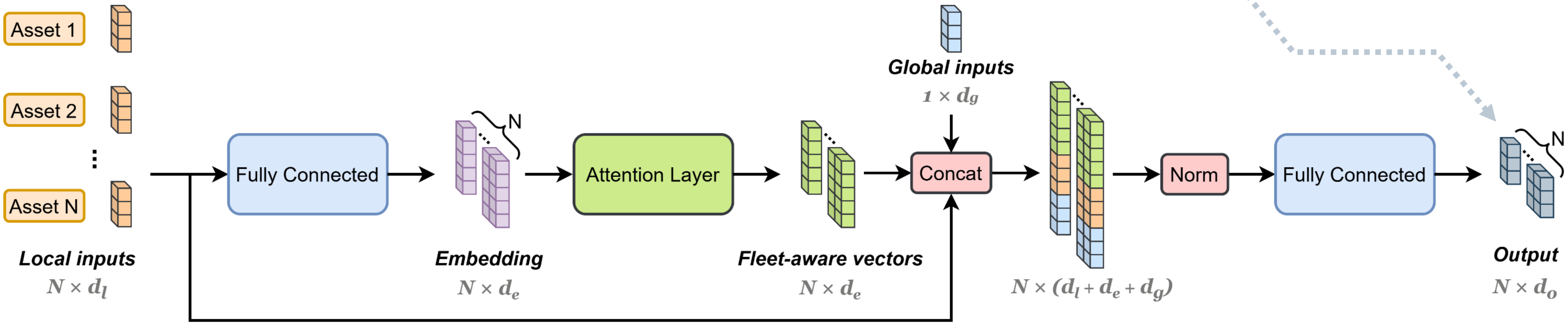
- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, …)

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Attention-based state representations for scalable RL

Idea:

- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, …)

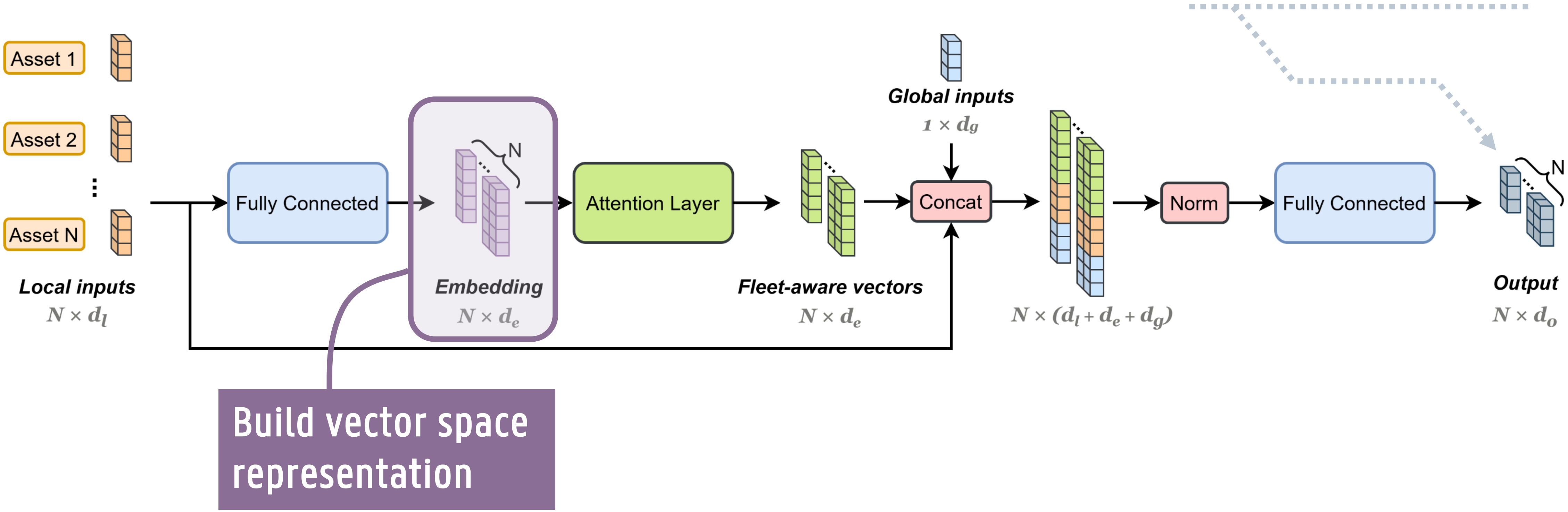


S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Attention-based state representations for scalable RL

Idea:

- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, ...)



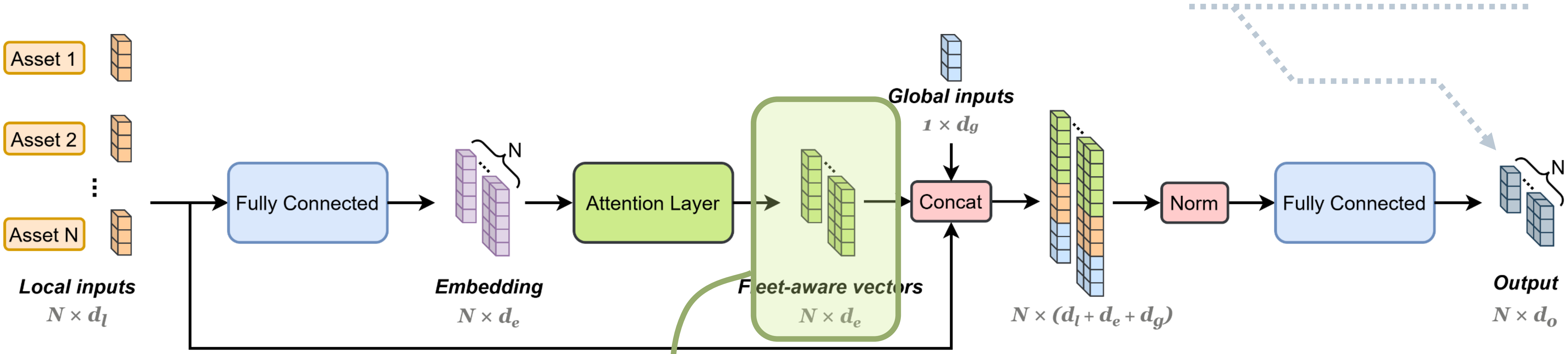
GHENT UNIVERSITY

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Attention-based state representations for scalable RL

Idea:

- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, …)



Adjust representation in light of other assets

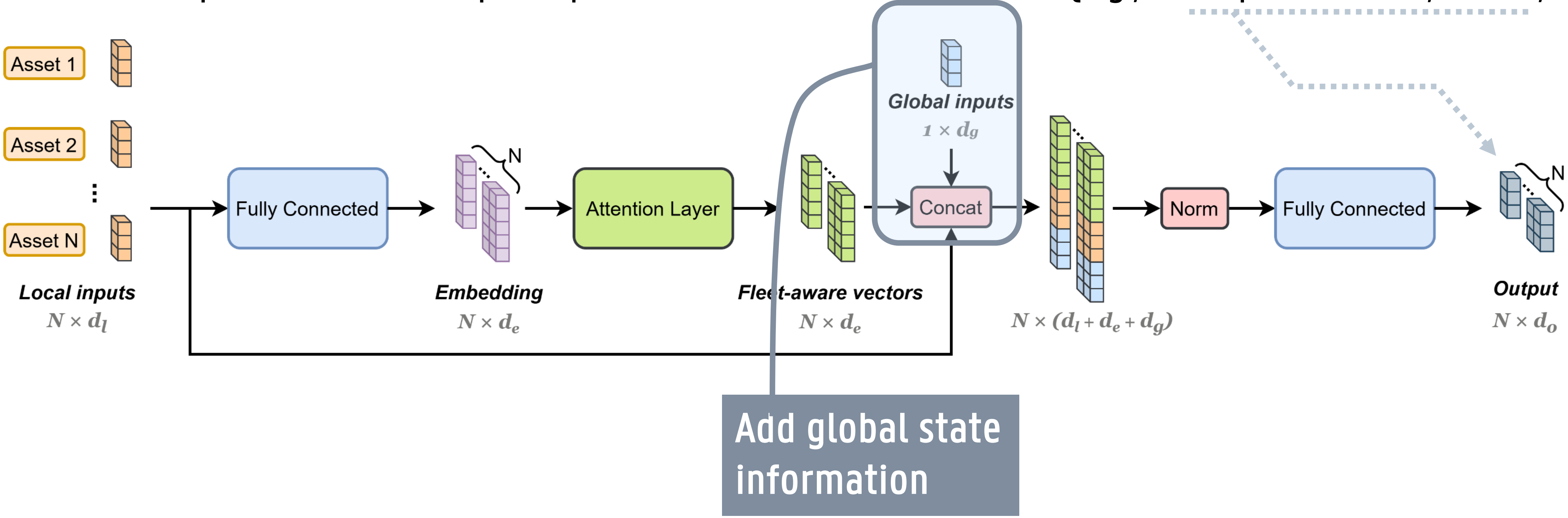
Note: Attention layer can deal with varying number of assets involved!

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Attention-based state representations for scalable RL

Idea:

- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, …)

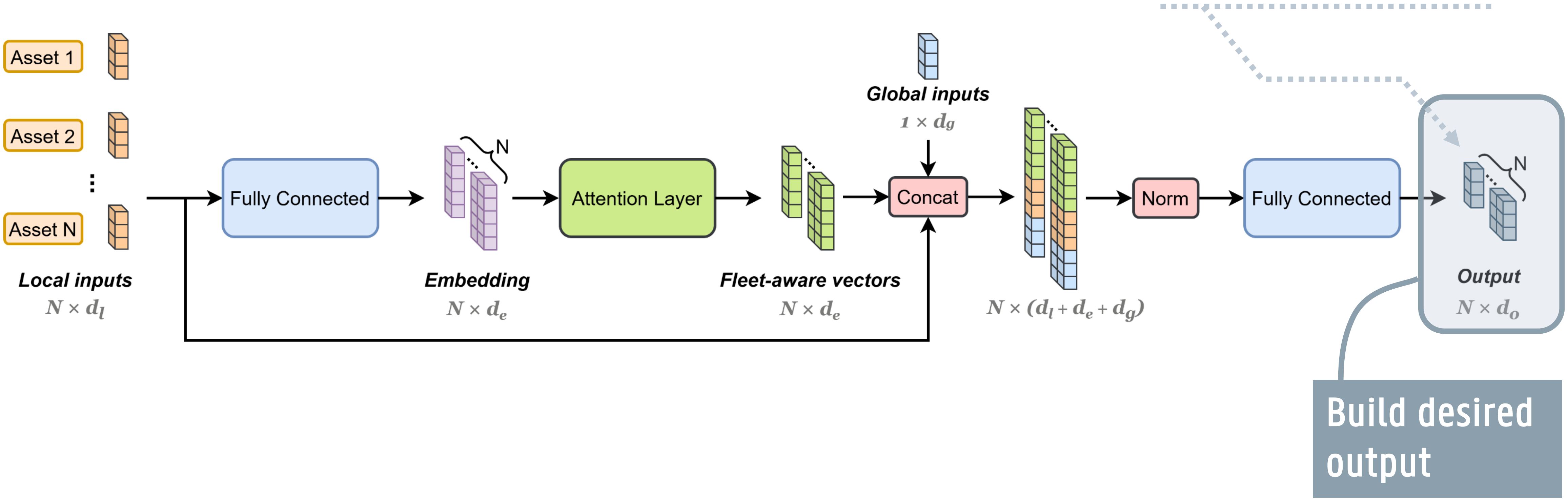


S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Attention-based state representations for scalable RL

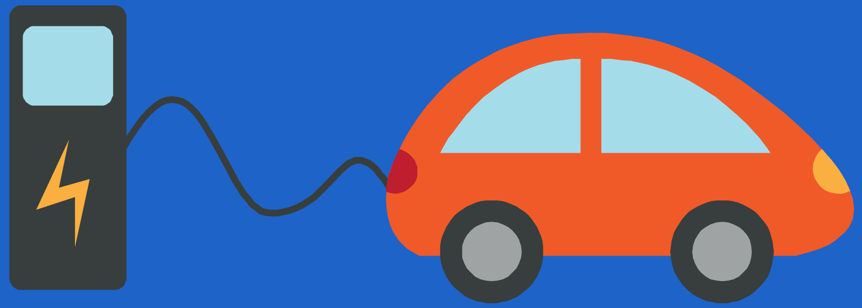
Idea:

- Construct state representation combining own asset + relative to other assets + global state
- Use state representation as input representation in conventional RL (e.g., to map to Q-value, action, …)



S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Case study: EV charging



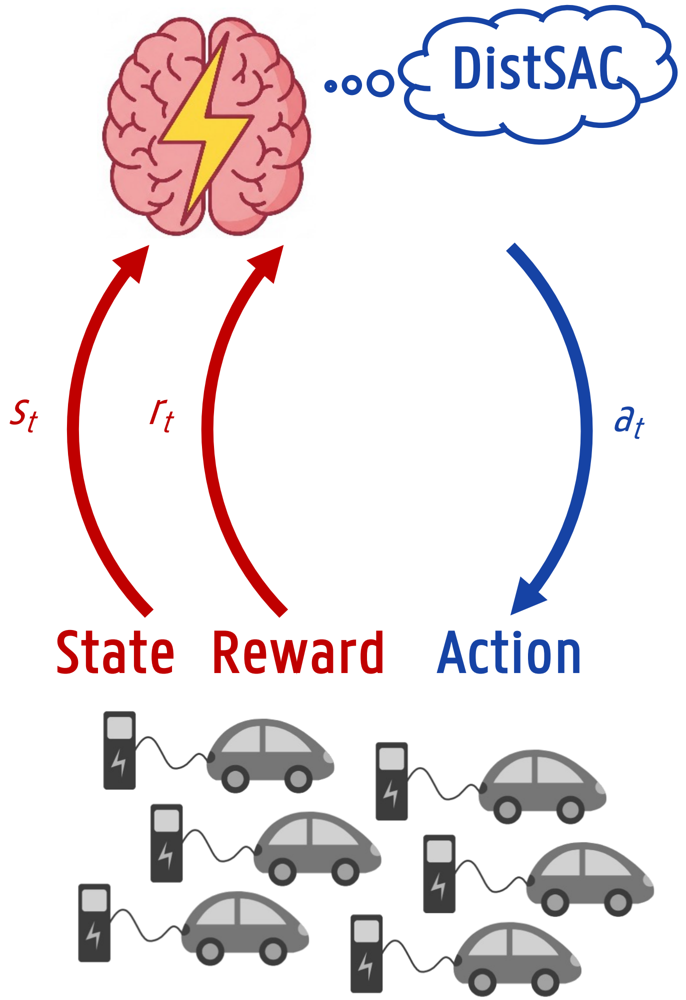
S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

MDP Formulation for EV Problem

- At each timestep t the agent
  - Gets **state** observation $s_t = \bigoplus_{i=1}^{N_t} (\underbrace{t_{i,t}^{\text{arr}}, t_{i,t}^{\text{dep}}, E_{i,t}^{\text{req}},}_{\text{local info}} \underbrace{t, N_t}_{\text{global info}})$
  - Decides to take **action** $a_t = \bigoplus_{i=1}^{N_t} a_{i,t}, \quad a_{i,t} \in \{0, P_i^{\max}\}$
- The environment
  - Receives action a_t
  - Transitions to **state** s_{t+1}
  - Returns scalar **reward**

$$r_t = \underbrace{-\left(\sum_{i=1}^{N_t} a_{i,t}\right)^2}_{\text{load flattening}} \underbrace{- \alpha \sum_{i=1}^{N_t} \mathbb{1}\left\{E_{i,t}^{\text{req}} > t_{i,t}^{\text{dep}} \cdot P_i^{\max}\right\}}_{\text{penalty for unfinished charging}}$$

RL goal: learn policy to flatten the load of the whole parking, while satisfying the needs of EV users

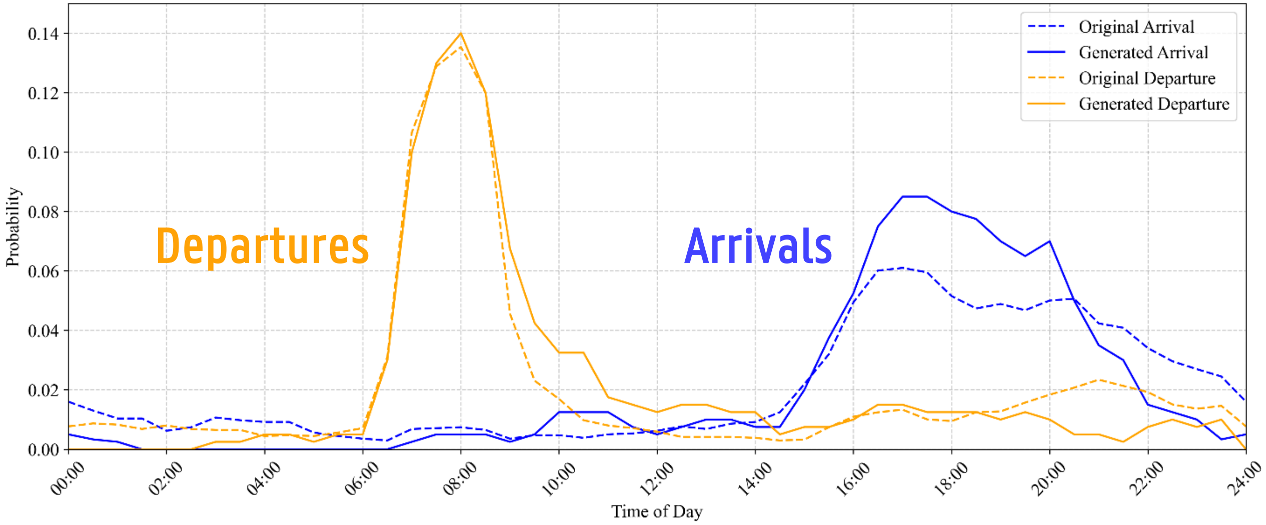


S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Simulation Setup

Datasets

- Historical real-world dataset from 20 EVs
  - Belgian residential charging sessions
  - Data split: 40d train / 5d eval / 5d test
- Scaled dataset from 100 EVs
  - Fit Gaussian mixture model on historical dataset
  - Data split: 30d train / 7d eval / 7d test

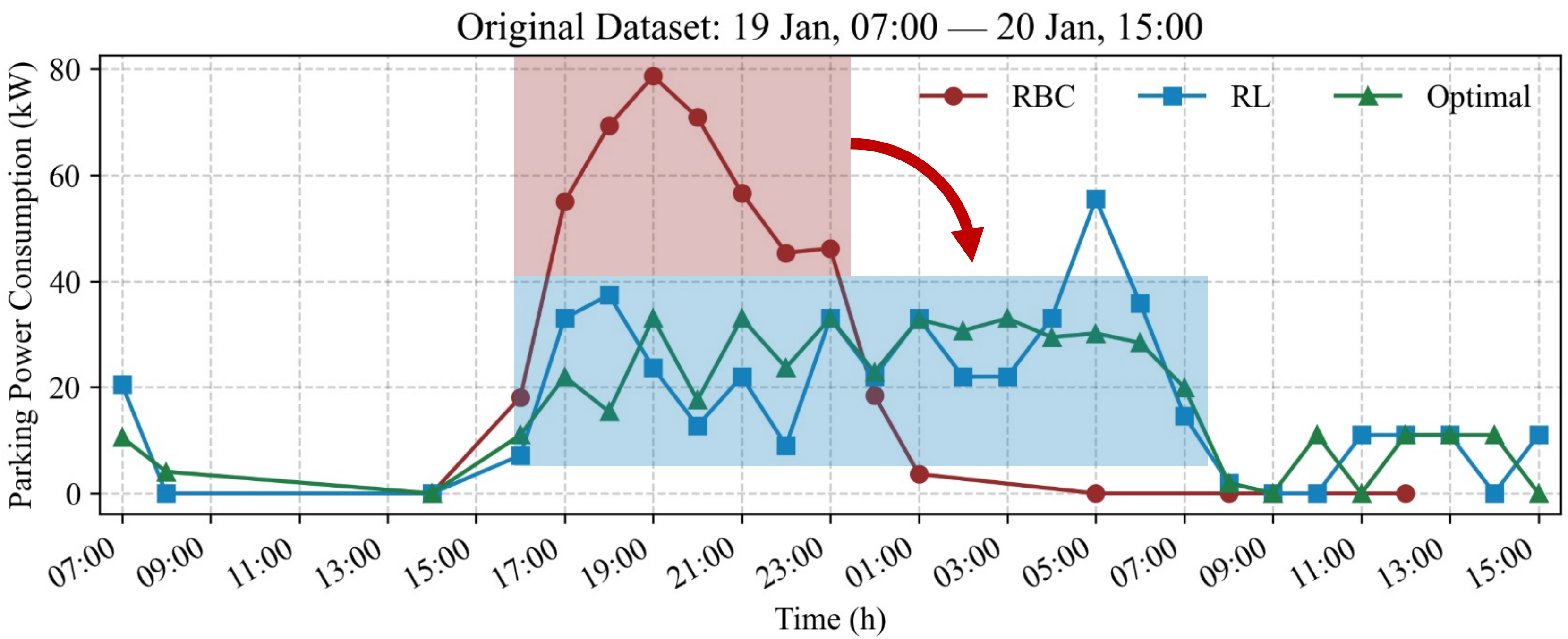


Benchmark controllers

- Business-as-usual (BAU): Start charging upon arrival
- Optimal: MPC optimization with perfect foresight over full day

GHENT UNIVERSITY

Simulation Results – Our RL outperforms BAU



Attention-based RL

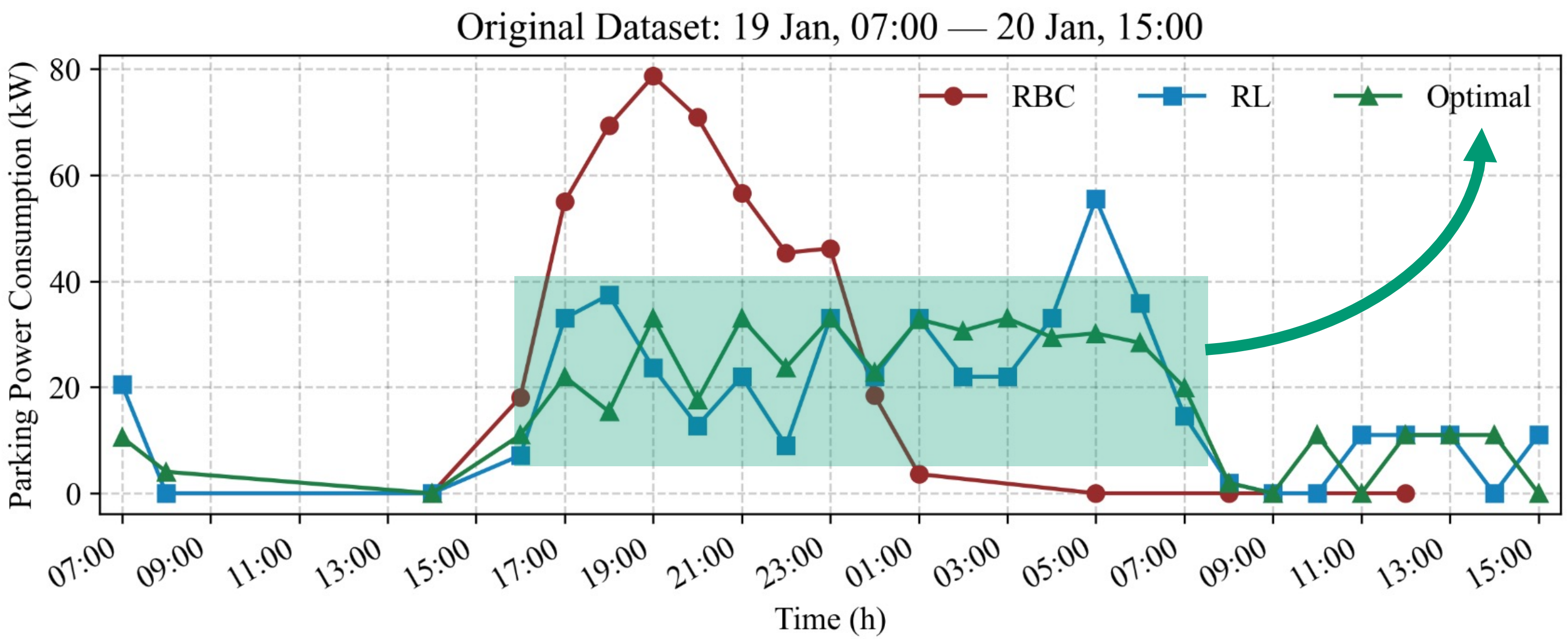
- Flattens load (peak reduced by 27%)*

*: 27% is the average over full test set

GHENT UNIVERSITY

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Simulation Results – Our RL outperforms BAU



Attention-based RL

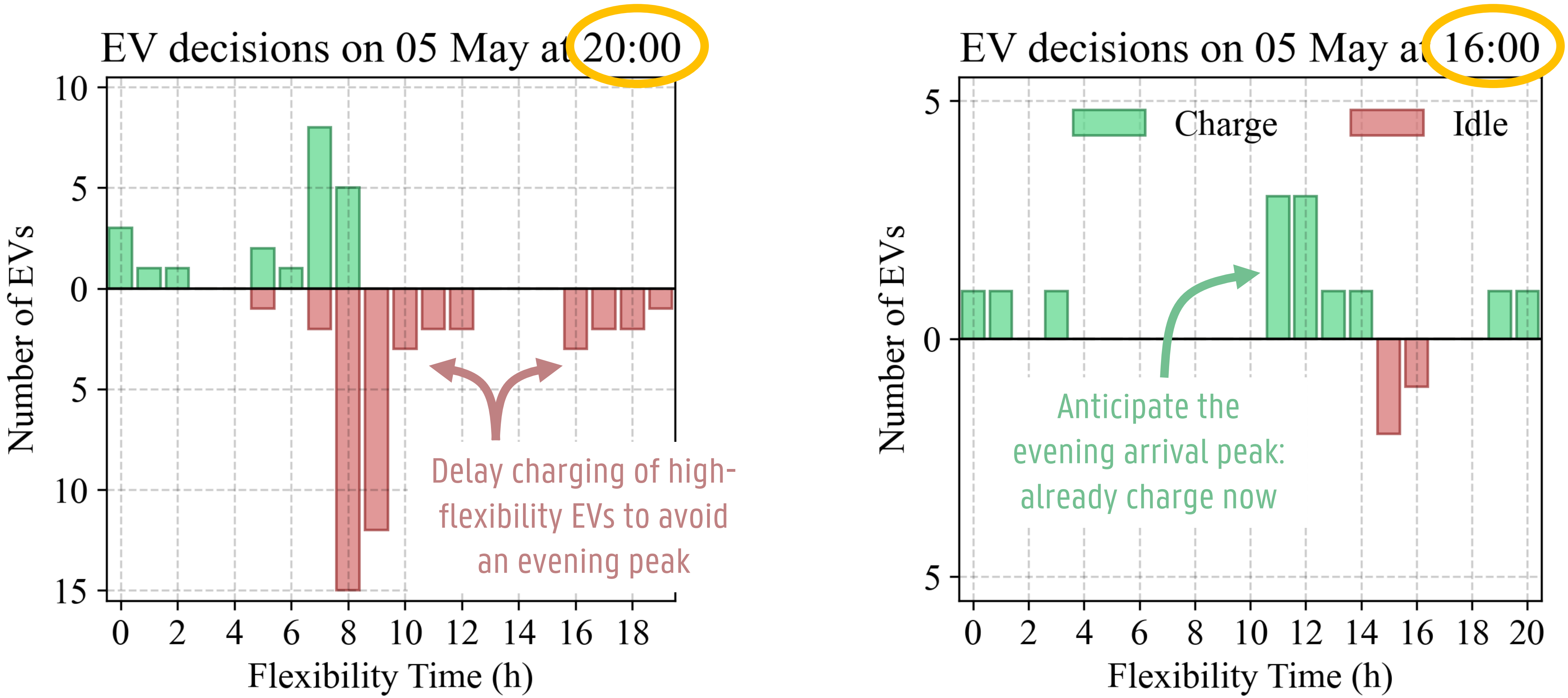
- Flattens load (peak reduced by 27%)*
- Performs close to optimization based on perfect foresight

*: 27% is the average over full test set

GHENT UNIVERSITY

S.S.K Madahi, G. Gabriele, B. Claessens, and C. Develder, "Scalable Attention-based Reinforcement Learning Method for Multi-asset Control", CO-BUILD @ ICML 2025

Simulation Results – Our RL learns to prioritize EVs correctly



Wrap-up

Conclusion

- Introduced a **scalable centralized** RL method to jointly control a fleet of flexible assets
    - Asset-specific info aggregation using attention layer
    - Model parameter size independent of number of assets
- Case study: load flattening of EV parking lot
    - Outperforms BAU by 28%
    - Average peak reduction of 27%

Future work

- Application to other asset types in multi-agent energy problems
- Joint control of a fleet of <u>heterogeneous</u> flexible assets

Thank you!
Any questions?

It's not easy being green ...

BLUE BIRD

Funded by the European Union

Flanders State of the Art

GHENT UNIVERSITY

Thank you!
Any questions?

UGENT
AI4E

Thank you!
Any questions?

PAPER

TEAM

It's not easy being green ...

GHENT UNIVERSITY