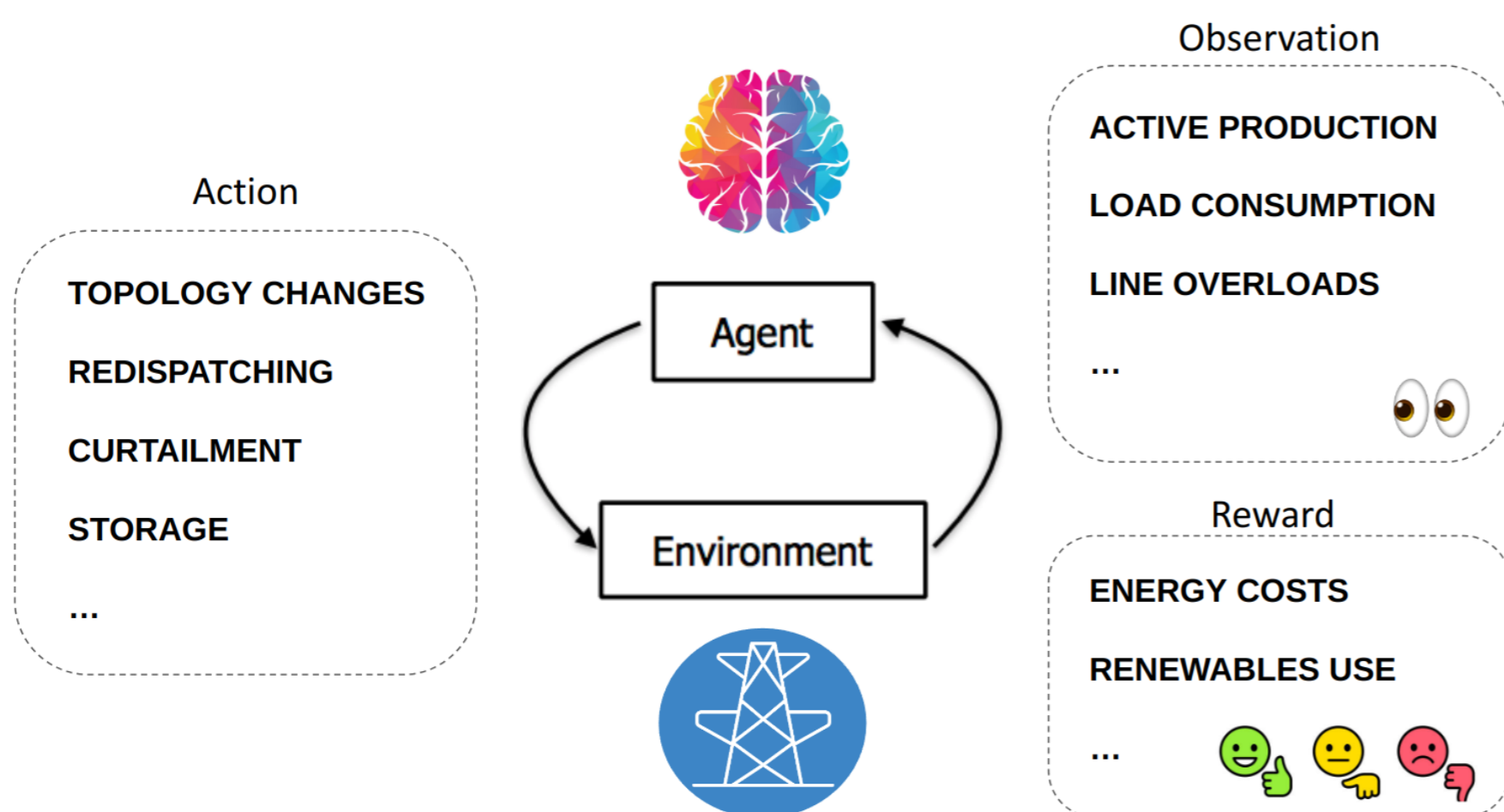# State and Action Factorization in Power Grids

Losapio G · Beretta D · Mussi M · Metelli AM · Restelli M

RL³ Lab, Politecnico di Milano

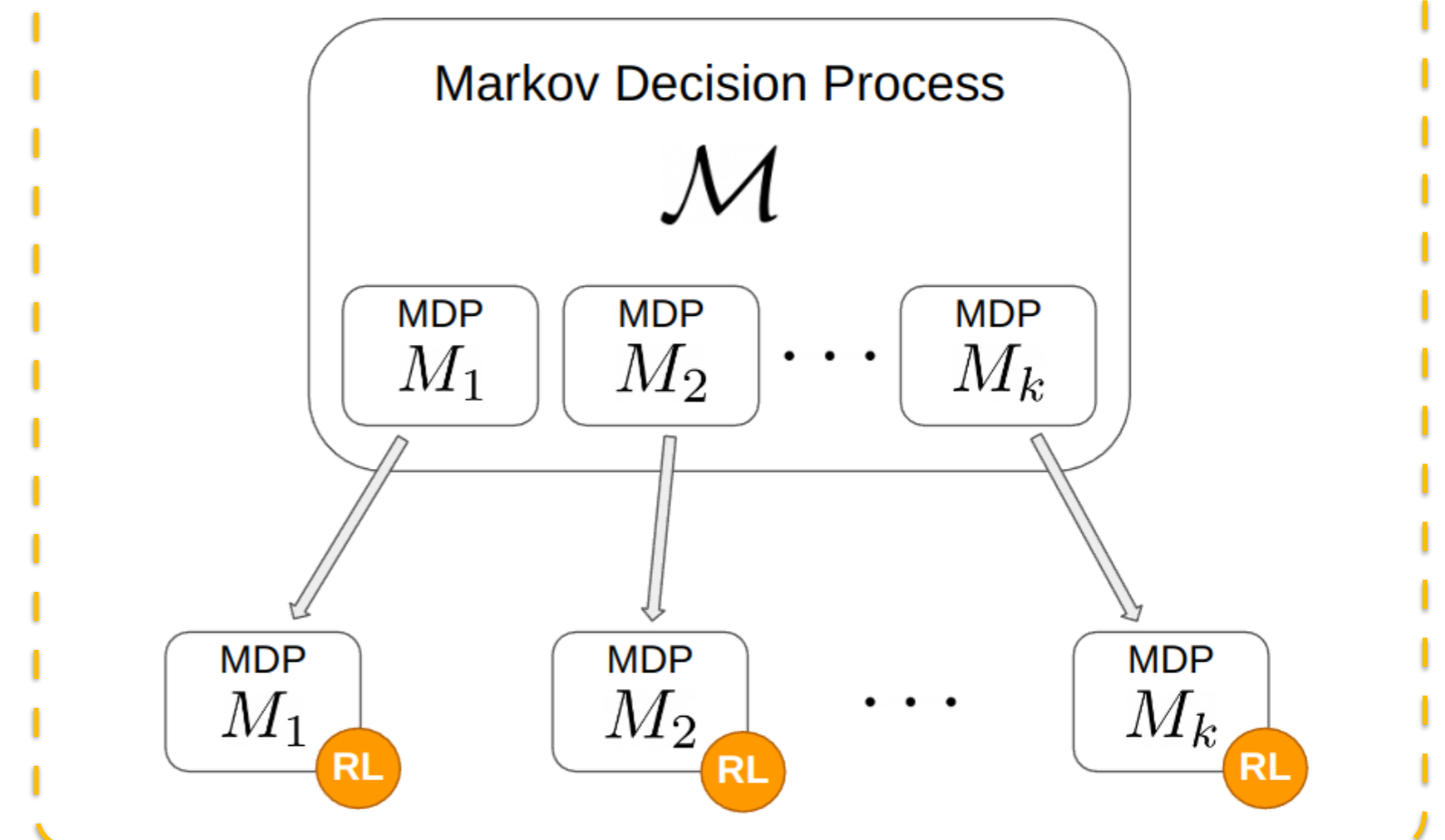## Motivation

Lots of recent works on RL for power grids
(L2RPN competitions)

Problem: curse of dimensionality
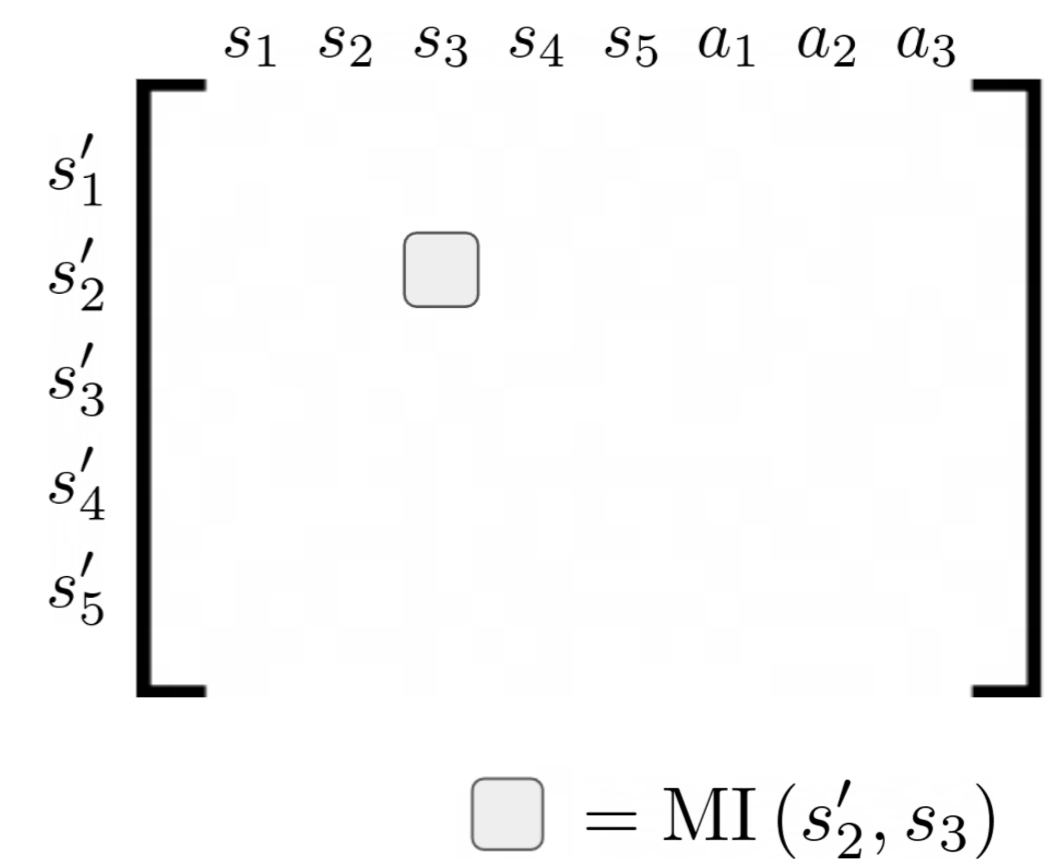
Factorization in smaller independent MDPs

Action
- TOPOLOGY CHANGES
- REDISPATCHING
- CURTAILMENT
- STORAGE
- ...

Agent

Environment

Observation
- ACTIVE PRODUCTION
- LOAD CONSUMPTION
- LINE OVERLOADS
- ...

Reward
- ENERGY COSTS
- RENEWABLES USE
- ...

Markov Decision Process $\mathcal{M}$

MDP M_1 MDP M_2 $\cdots$ MDP M_k

MDP M_1 RL   MDP M_2 RL $\cdots$ MDP M_k RL

Contribution

domain agnostic ✅

Algorithm 1: State and Actor Factorization

Input: MDP $\mathcal{M}$, Explorative policy π_e, Threshold δ

Output: Factorization $\left(\widehat{\mathcal{S}}_k, \widehat{\mathcal{A}}_k\right)_{k=1}^{\widehat{K}}$

Algorithm:

1. Collect a dataset $\mathcal{D}$ of transitions from $\mathcal{M}$ with policy π_e

2. Compute the adjacency matrix $\widehat{I}_{\mathcal{G}}$ approximating the mutual information on $\mathcal{D}$ and using δ as threshold

3. Transform $\widehat{I}_{\mathcal{G}}$ into a pseudo-block diagonal matrix and define the set of clusters $\left(\widehat{\mathcal{S}}_k, \widehat{\mathcal{A}}_k\right)_{k=1}^{\widehat{K}}$ corresponding to diagonal blocks

$$s_1 \; s_2 \; s_3 \; s_4 \; s_5 \; a_1 \; a_2 \; a_3$$

s_1' s_2' s_3' s_4' s_5'

$$\square = \mathrm{MI}\left(s_2', s_3\right)$$

Experiments

- Synthetic data (error = 0.02)
- IEEE case14 benchmark (Grid2Op simulator)

$\downarrow$

$\approx$ domain-expert analysis (*)
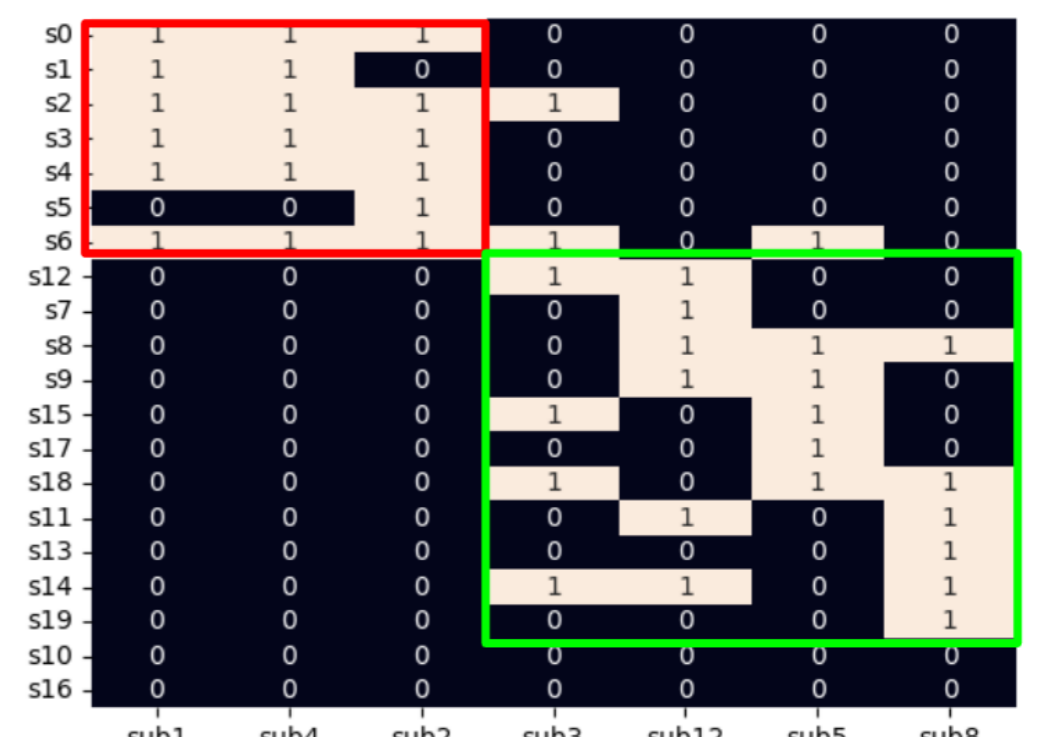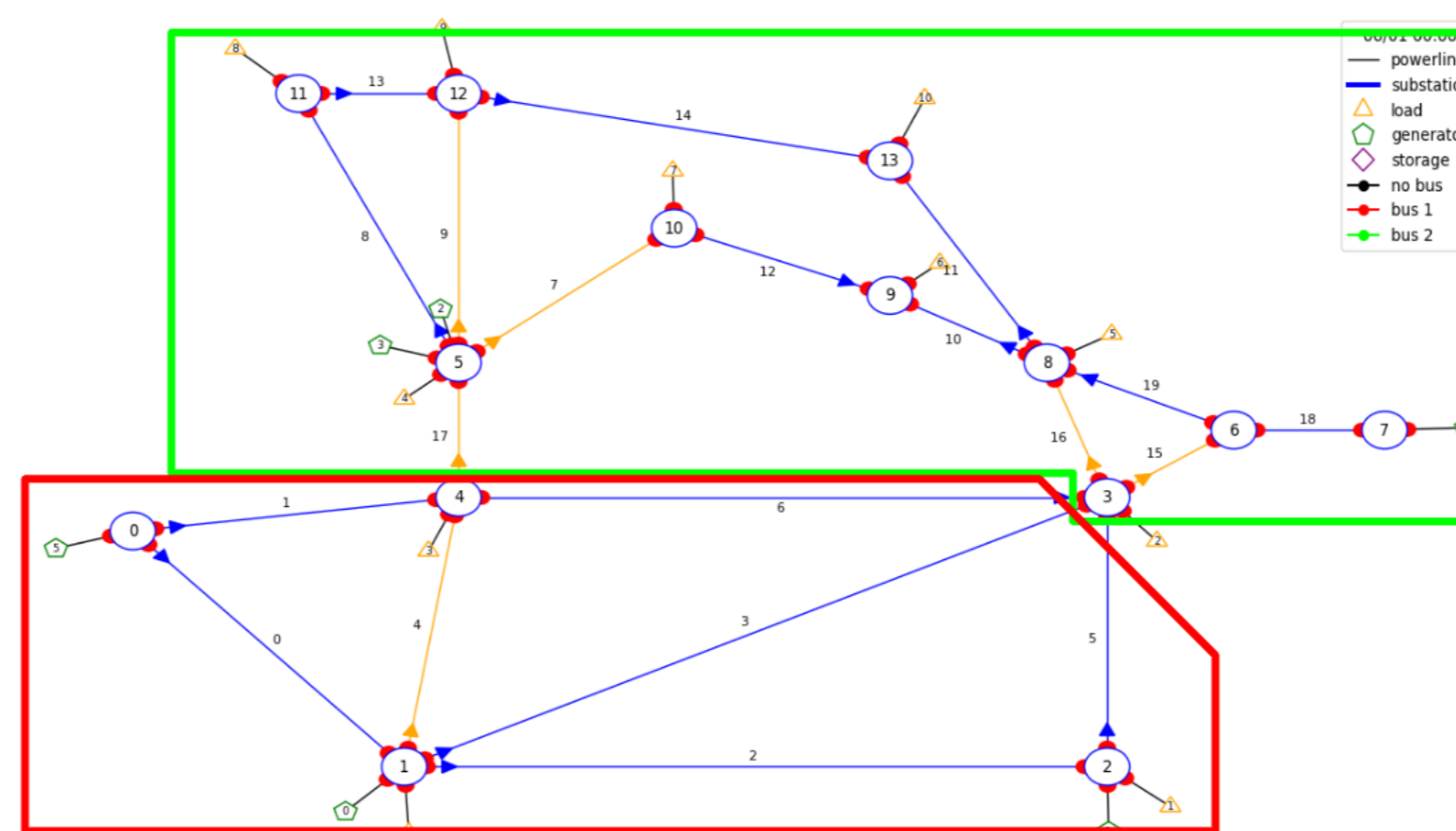
(*) Marot et al. "Guided machine learning for power grid segmentation" (2018)

Future work = {larger grids, hyperparams, correlation metrics, clustering, ... }