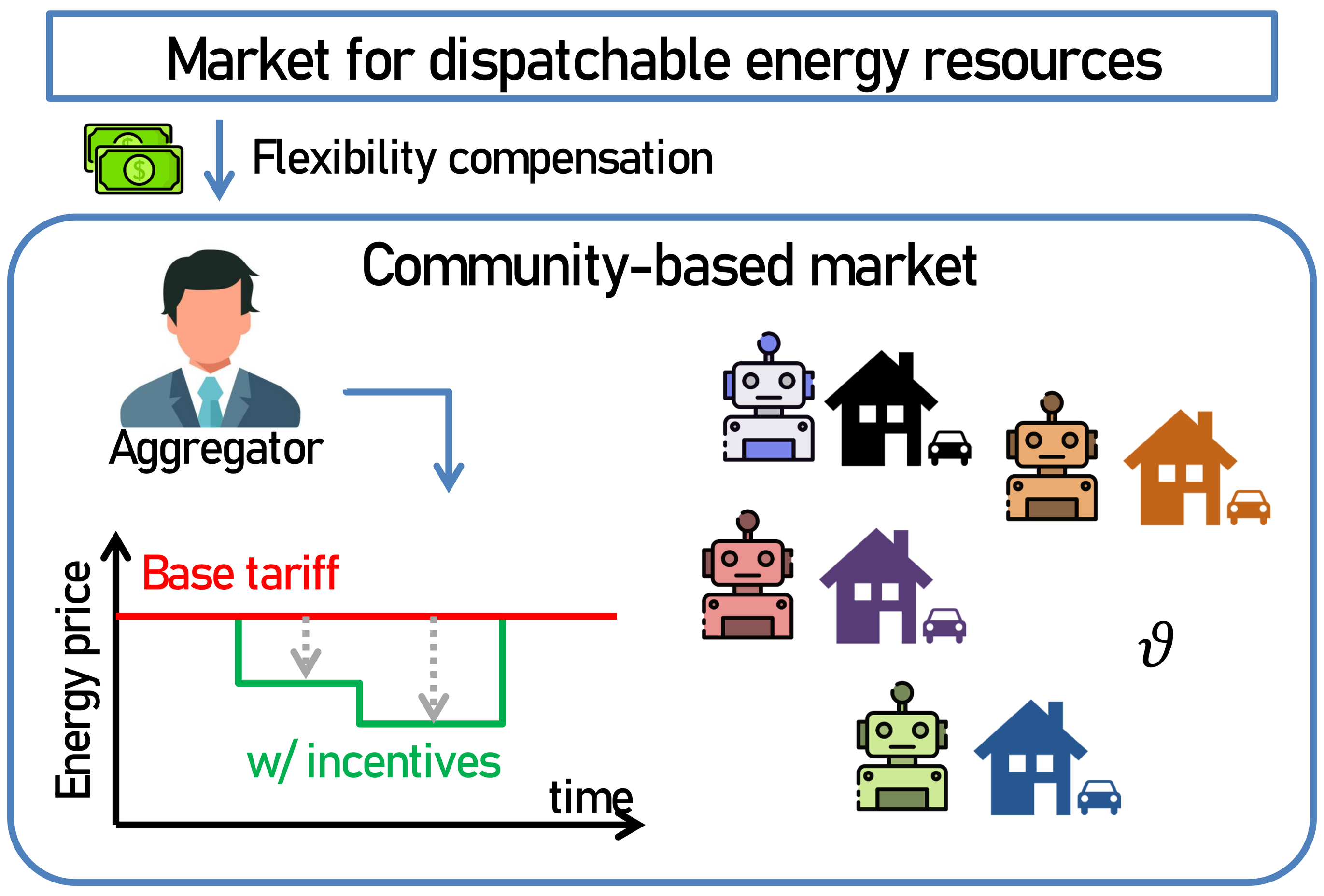
TMSG - Training Multiagent Strategies for the Grid

Federico Rosato *(federico.rosato@supsi.ch)*
Institute for Sustainability Applied to the Built Environment, DACD, SUPSI

In the TMSG project, we seeked to formalize the automatic concurrent management of flexible electric assets in a residential smart grid as a Markov Game. Discomfort tolerance and over-coordination avoidance are modeled as parameters. We introduced specialized techniques to obtain sensible solutions for the agents. The strategic behavior obtained was analyzed and the quantification of flexibility investigated. In the final phase, power flow simulations including the trained algorithmic agents were performed to gauge the effects.

Introduction

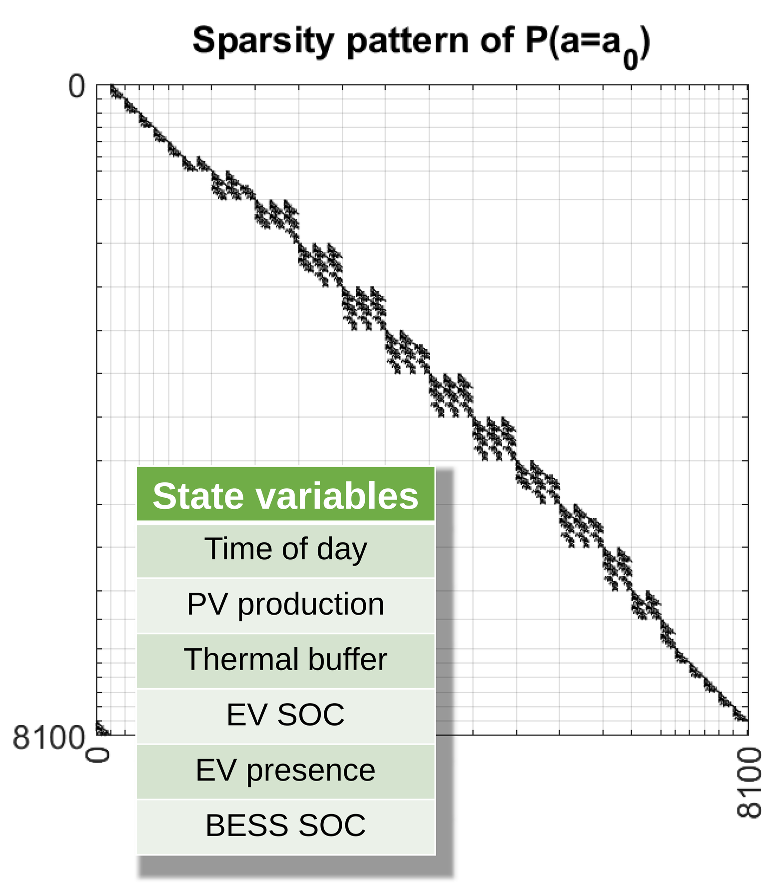
The general setup used is that of an Energy Community with a managed, community-based market with an **internal time-of-use tariff**.



The core research question pertains the **quantification of the flexibility** of the community, modeled as a group of rational self-interested agents parametrized by a vector ϑ including quantitative penalties for user discomfort, coordination penalties, etc. The agents seek to maximize the benefit offered to the user, while also avoiding excessive coordination that would disrupt the grid.

Markov Game modeling

The framework above is modeled as a Markov Game (S, A, P, R, γ) over one day. The state space S is composed of **values common to all agents** (time of day, weather), and a private part comprising appliances state. The agents can decide **whether to charge the batteries and/or to fill the thermal inertia**, giving a space A of **4 possible actions**. The rewards (penalties) on which the automatic agents are trained include the dissatisfaction of the user, the price of electricity and a **quadratic term for excessive coordination.**



Solution Method

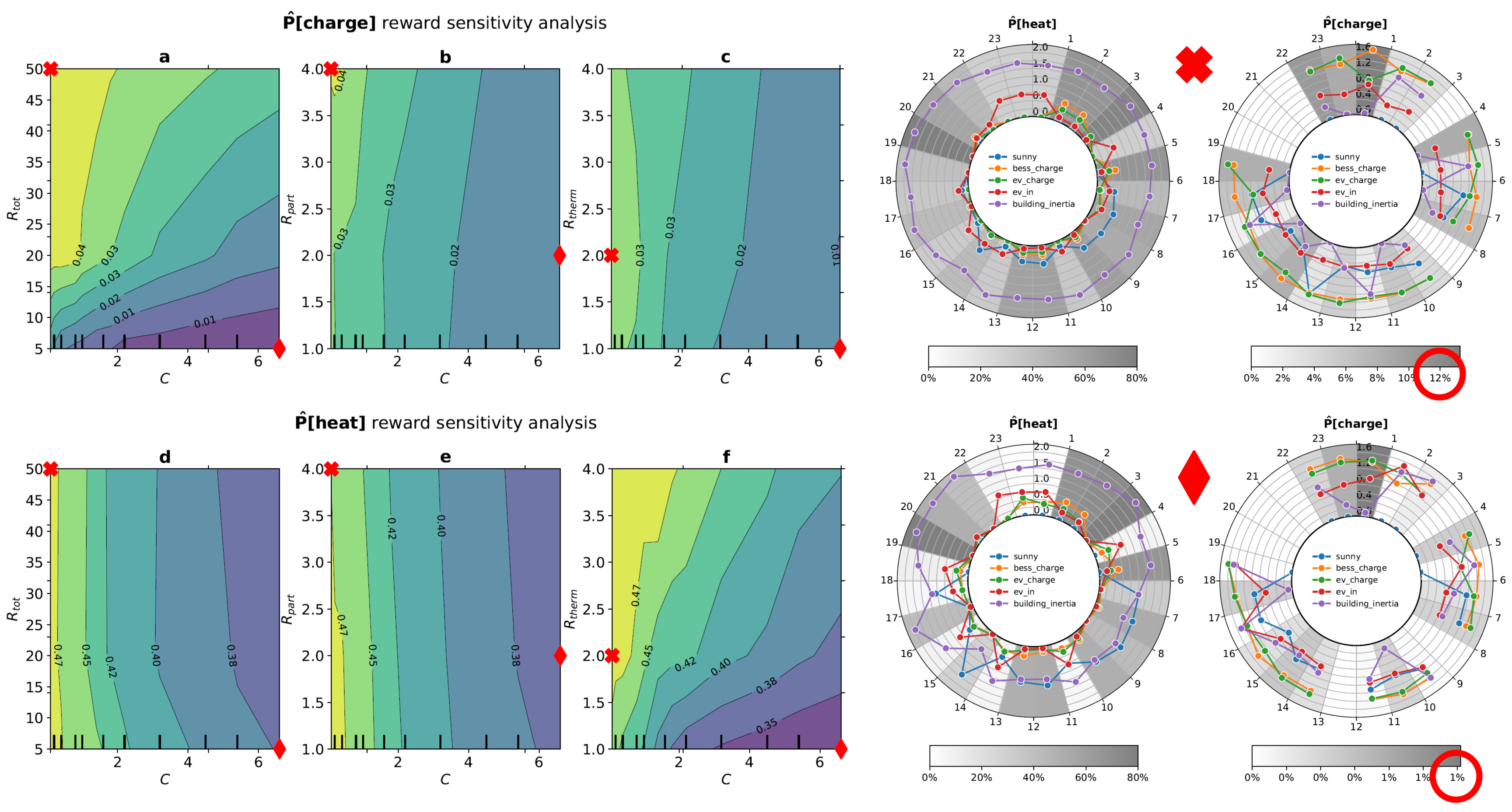
A two-step method was devised, resembling Policy Iteration with an additional coordination step. The method uses the simplyfing stipulation of **agent simmetry and the informational structure of the problem** (agents are unaware of others' private state).

$$\begin{aligned}\max_{x} \quad & \sum_{s\in S_{p,h}} \Big(\sum_{a\in A} Q(s,a)x_{s,a} - c\sum_{s'\in S_{p,h}} \mathbb{P}(s'|s)\big(\sum_{a\in A} x_{s',a}\hat{\mathcal{E}}(s',a)\big)^2\Big)\\ \text{s.t.} \quad & x \in \mathbb{R}_+^{|S_{p,h}|\times|A|}\\ & \sum_{a\in A} x_{s,a} = 1 \ \ \forall s \in S_{p,h}\end{aligned}$$

- *Find state values w/ current policy*
- *Recompute policy by solving convex formulation at each public state*

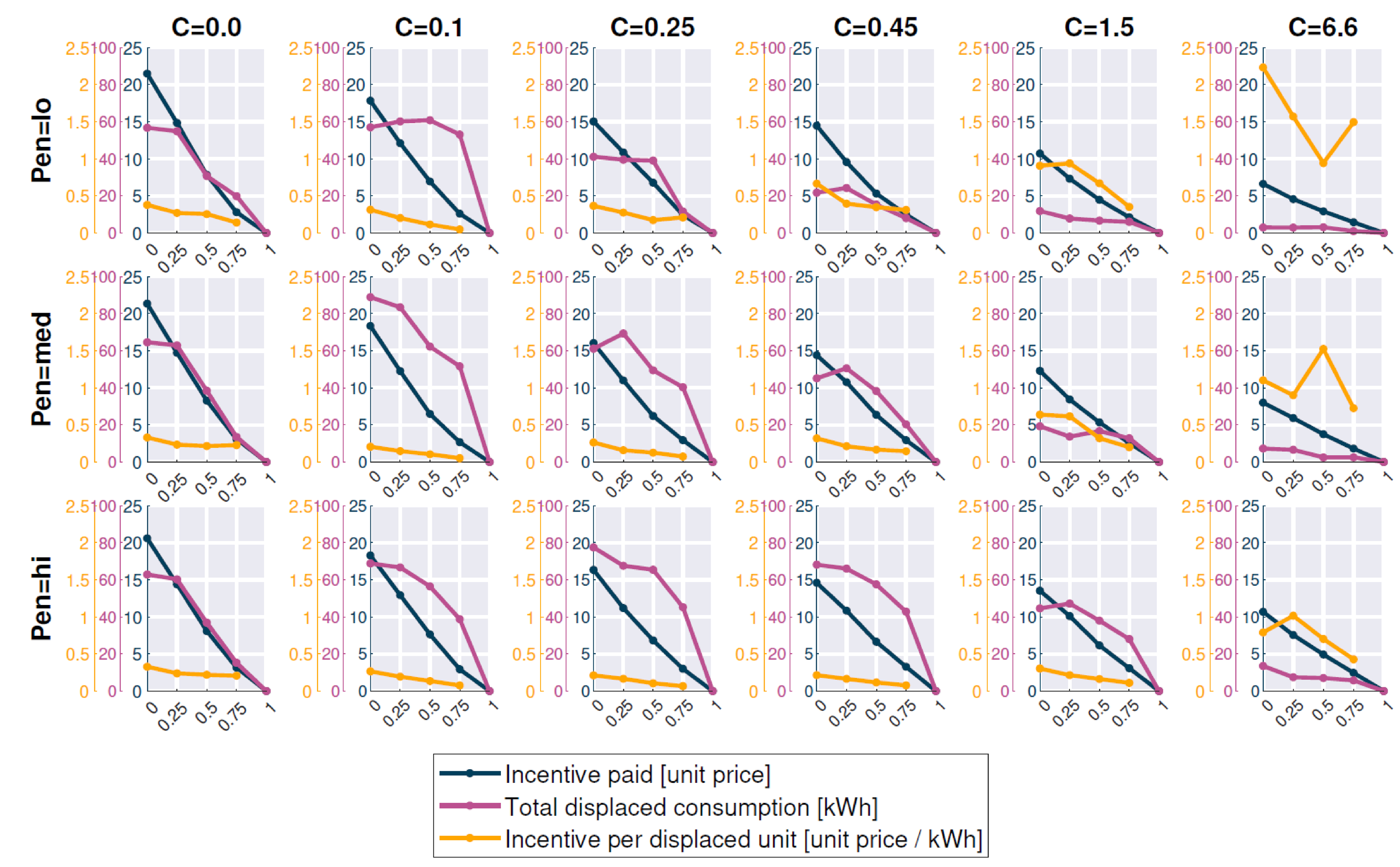
While not guaranteeing theoretical convergence due to the unfavorable dynamics of markov games, the method showed **good convergence to «sensible» solutions** in reasonable time.

Results



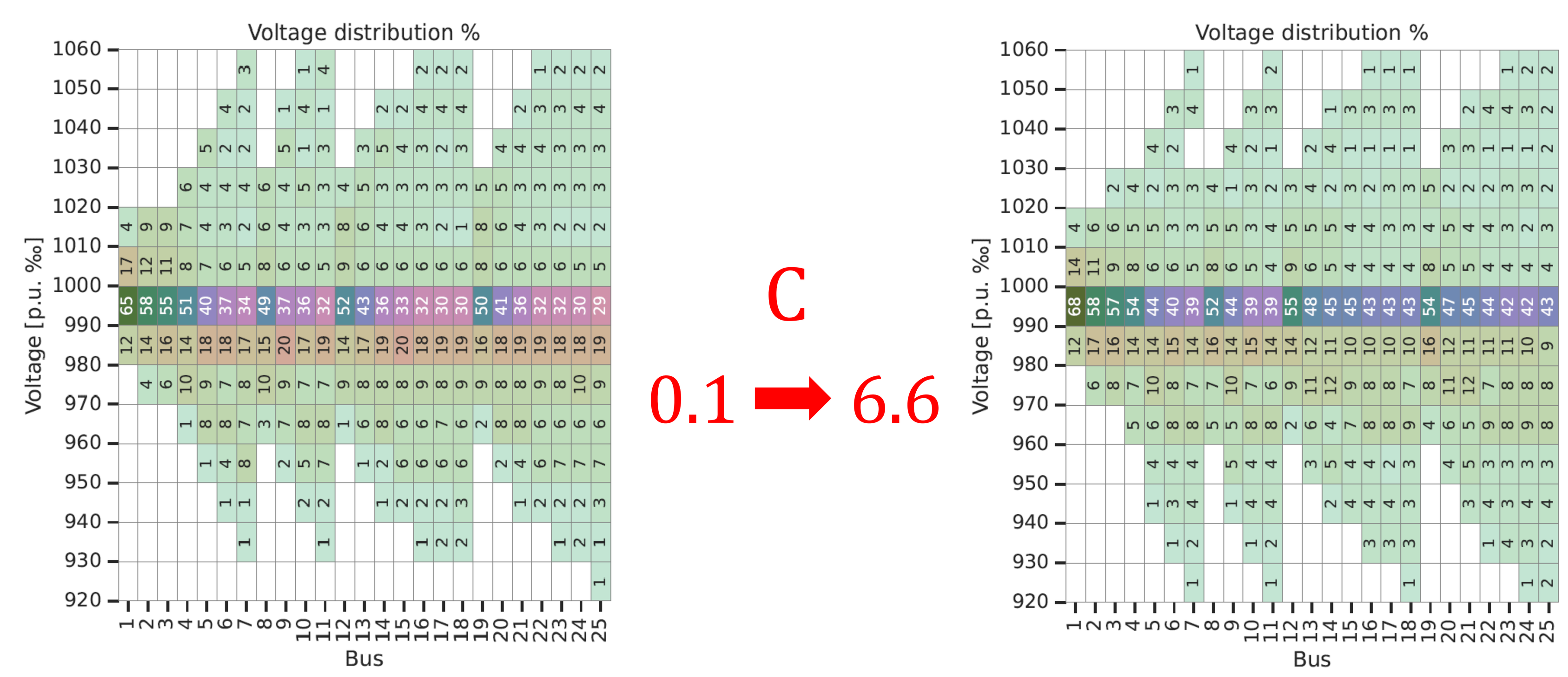
The solution method was run on a wide variety of parametrizations. In a case with no user consumption, **the effect of penalties is clear on the "aggressiveness" in taking the relative action**. With a custom visualization method for showing feature importance over time, we see how **changing the reward structure influences the decisional process**.

Flexibility



By establishing a flat price incentive in the central hours of the day of low residential consumption (12:00-17:00), we can shift the load. The tool allows us to **quantify the displaced load and the relative and absolute incentive expenditure**; this is just a case study, as any incentive shaping can be analyzed. This showcases the usefulness of the tool as an Energy Community evaluation and feasibility estimator.

Power flow



The C coefficient is an important part of the model, coupling the agents and allowing them to take into account the overloading of the grid. By performing power flow simulations on a test grid with the automated agents, we see how increasing C can **stabilize the voltage levels**. The effect of other parameters on voltage was also explored.

The research presented here is supported by the Swiss Federal Office of Energy SFOE as part of SOUR project TMSG.