



Contextual Background

Fairness

In this study, **fairness** in active power curtailment refers to ensuring that every customer is curtailed in an equal amount of energy.

To assess the effectiveness of fair curtailment, the **Gini index** is utilised as a quantitative measure.

Smart Inverter

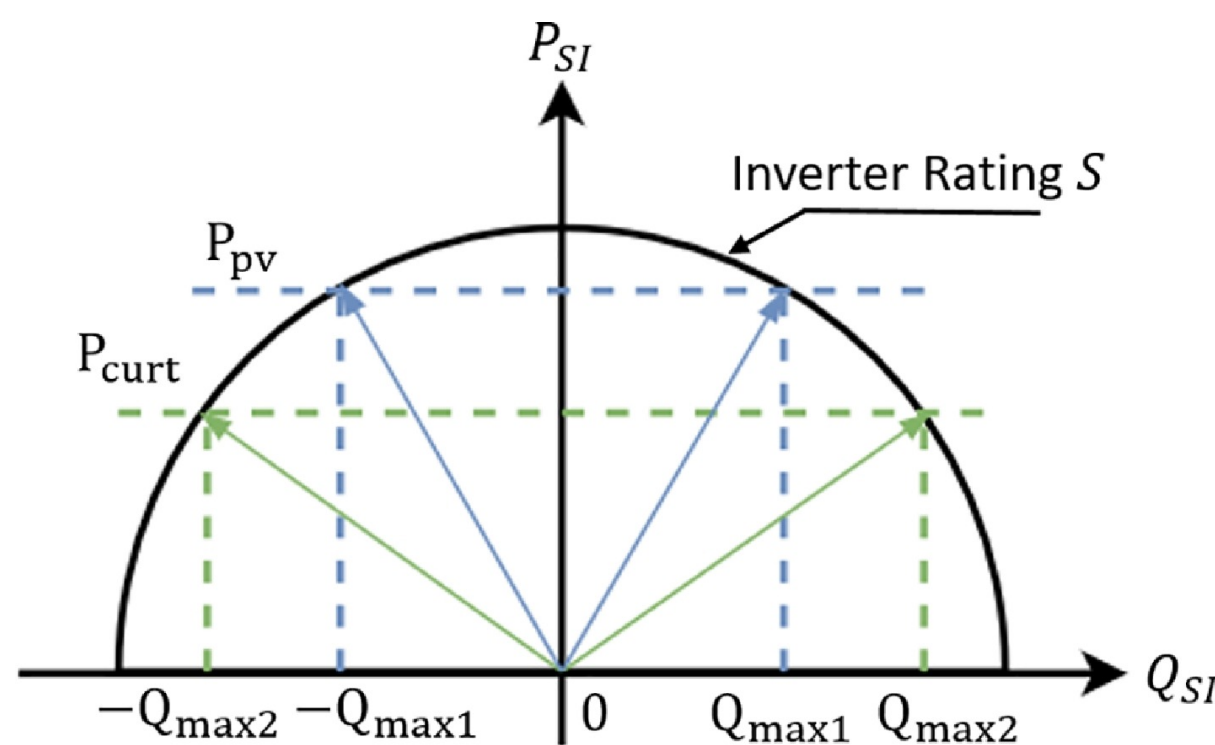


Fig 1. SI capability curve

The **active power**, P , and **reactive power**, Q , can be controlled to satisfy the formula $P^2 + Q^2 \leq S^2$.

Reinforcement Learning

- Environment
- State, s_t
- Action, a_t
- Agent
- Reward, r_t
- Discount factor, γ .

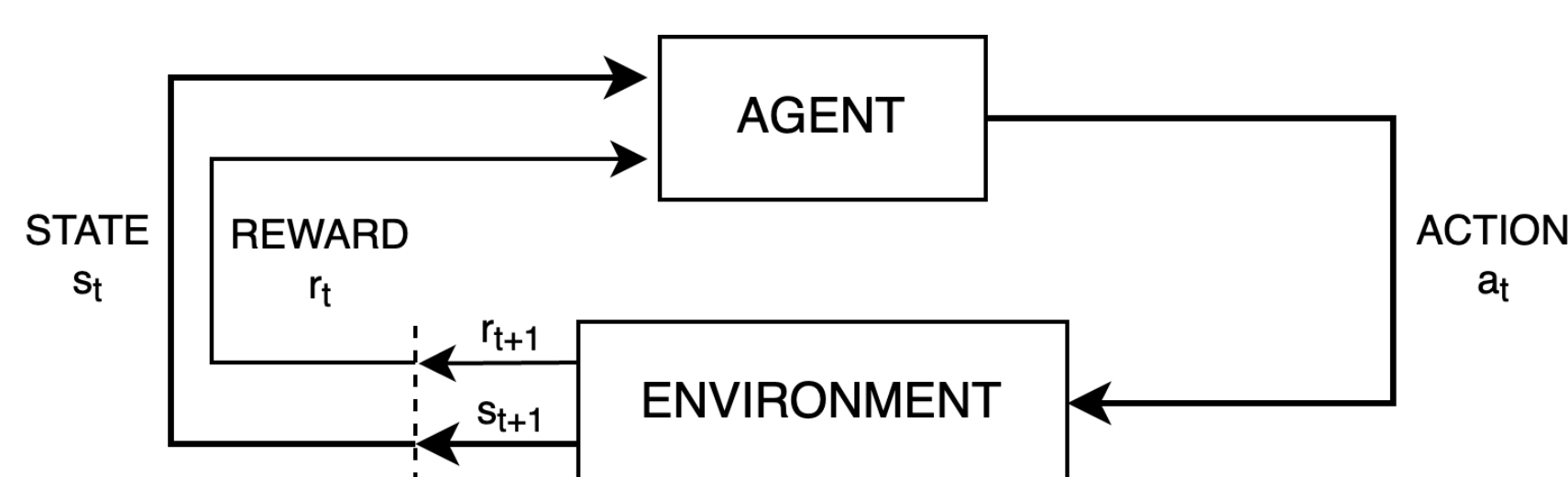


Fig 2. Action-reward feedback loop of a generic RL framework

The goal of the agent is to maximise the **expected return**:

$$R_t = \lim_{T \rightarrow \infty} \mathbb{E}_{\pi} \left[\sum_{i=t}^T \gamma^{i-t} r_i \right] \quad (1)$$

General RL solutions makes use of a recursive relationship known as the **Bellman equation**:

$$Q_{\pi}(s_t, a_t) = \lim_{T \rightarrow \infty} \mathbb{E} [r_t + \gamma \mathbb{E} [Q_{\pi}(s_{t+1}, a_{t+1})]] \quad (2)$$

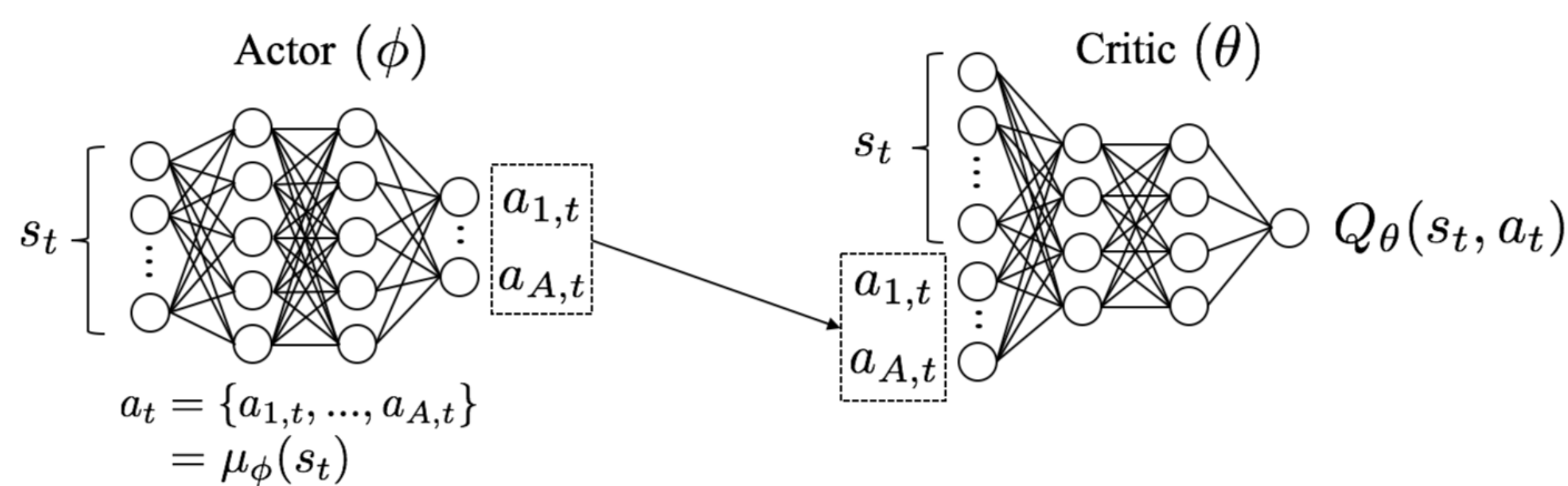


Fig 3. Actor-critic architecture in DDPG

Research question

How can the adoption of **reinforcement learning** techniques address **voltage issues** in distribution networks caused by the increasing use of photovoltaic panels, while ensuring **fairness** in active **power curtailment** among customers?

Proposed Method

- **Environment**: LV distribution network
- **Agent**: DDPG algorithm
- **Space**:

$$s_t = [\mathcal{L}_t^p, \mathcal{L}_t^q, \mathcal{DG}_t^p, V_t]$$

- **Action**: The agent can choose the reactive power output of each SI, from $-Q$ to $+Q$.
- **Reward**: The reward, r_t , consists of the sum of three terms:
 - A reward, r_v , regarding **voltage violation**.

$$r_v = - \sum_{i=0}^N f_v(V_{i,t}) \quad (3)$$

- A reward, r_a , limiting the **action magnitude**.

$$r_a = - \sum_{i=0}^{\mathcal{DG}} |a_{i,t}| \quad (4)$$

- A reward, r_f , for the **fair curtailment**.

$$r_f = - \sum_{i=0}^{\mathcal{DG}} |\bar{c}_t - c_{i,t}| \quad (5)$$

The total reward is given by:

$$r_t = \alpha \cdot r_v + \beta \cdot r_a + \omega \cdot r_f \quad (6)$$

The scaling factors satisfy the following relationships:

$$\alpha \cdot r_v \gg \beta \cdot r_a \simeq \omega \cdot r_f.$$

Results

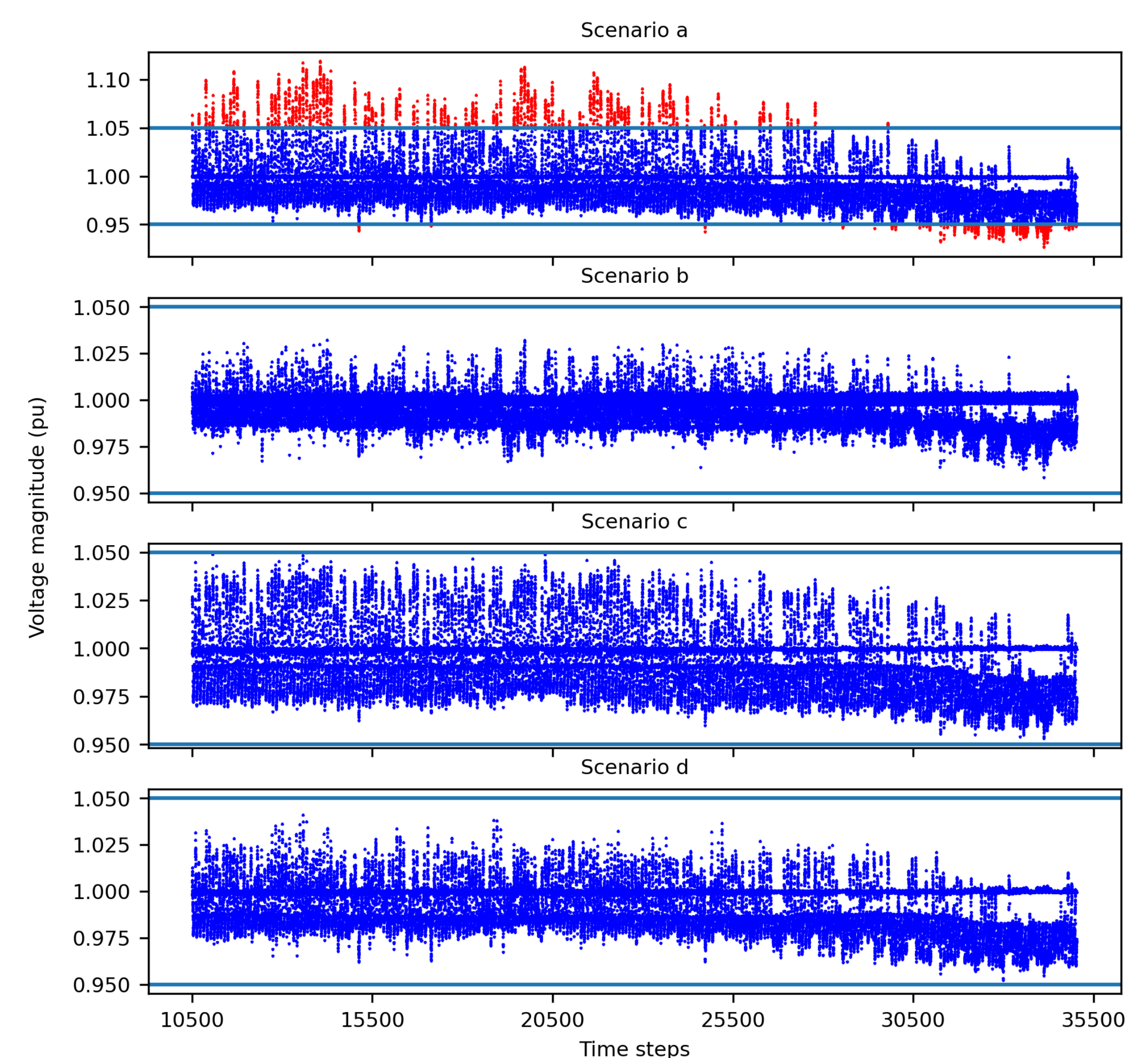


Fig 4. The buses' voltage situation for each scenario

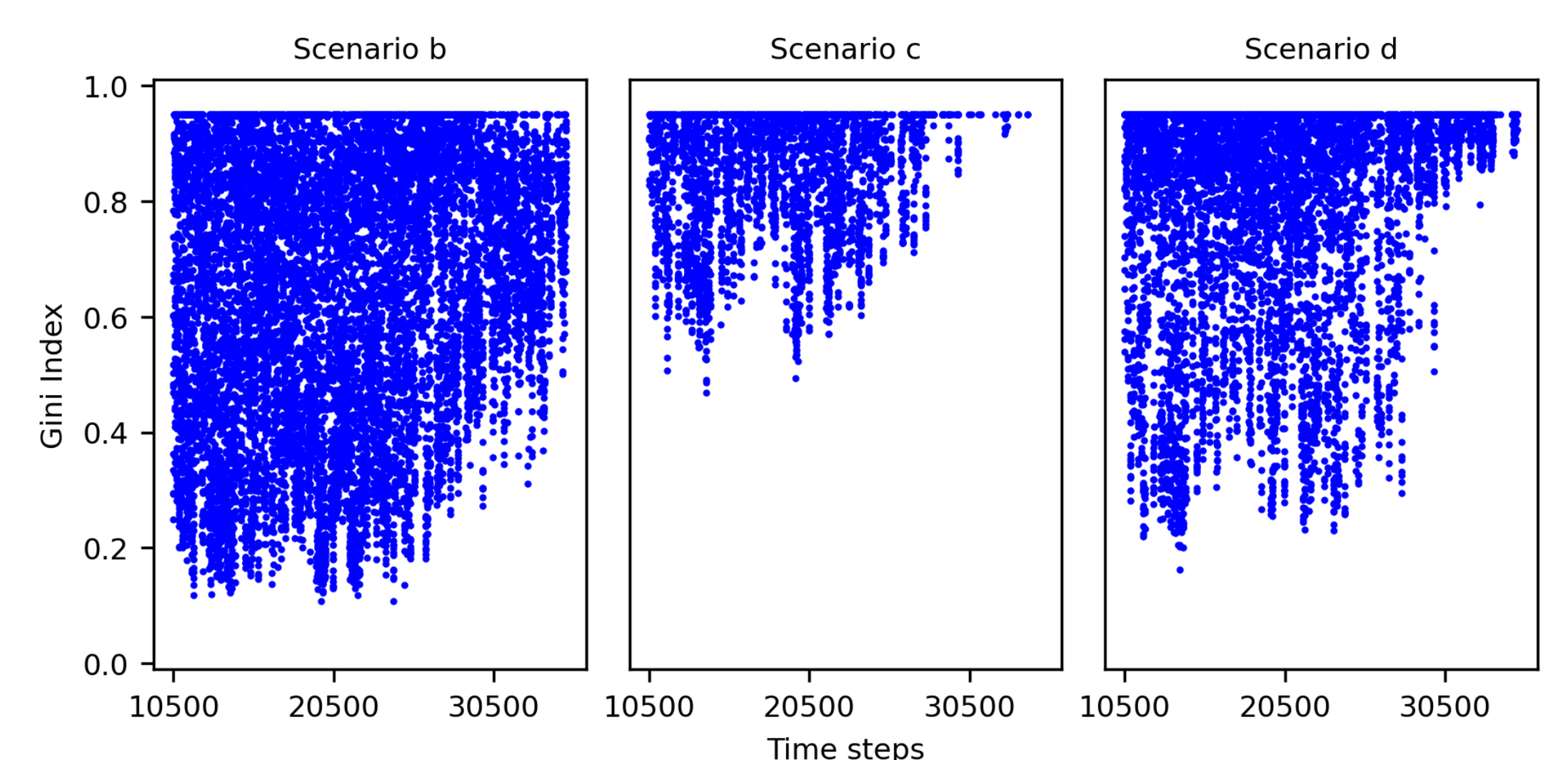


Fig 5. Gini index for each scenario where some control is applied

Table 1.

Summary of the results for each scenario

Scenario	Total PV generation (MWh)	Energy Curtailed (MWh)	# Under Voltages	# Over Voltages	Min V (pu)	Max V (pu)
a	47.5	0	493	1677	0.9263	1.1191
b		20.7	0	0	0.9582	1.0321
c		2.65	0	0	0.9530	1.0497
d		8.77	0	0	0.9520	1.0409

Conclusion

Contribution

The results show that it is possible to obtain a **fair active energy curtailment** for the customers, so to **avoid unequal curtailment**, but at the **cost of a higher curtailment**.

Future works

- Analyse the impact of the weights β and ω on the results.
- Use a decentralised reinforcement learning technique.
- Consider the limitations of SIs.