POLITECNICO MILANO 1863

RSE Ricerca Sistema Energetico | we move research

DOCTORAL PROGRAMME IN INFORMATION TECHNOLOGY

Empowering Digital Twins for Battery Energy Storage Systems with Machine Learning

Design, Development, and Applications of the Framework

Davide Salaorni

Supervisors: Prof. Marcello Restelli and Prof. Francesco Trovò

Politecnico di Milano - DEIB

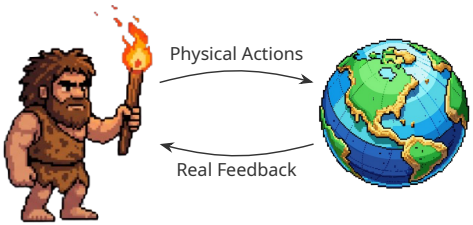
RSE S.p.A.

Milan, Italy / July 8, 2026

The Source of the Empirical Knowledge

A Clone of the Real World

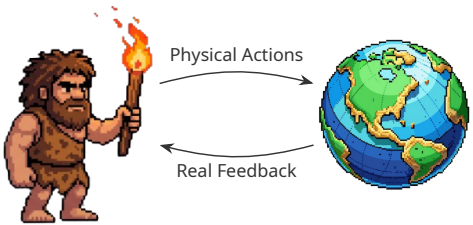
Before Computers



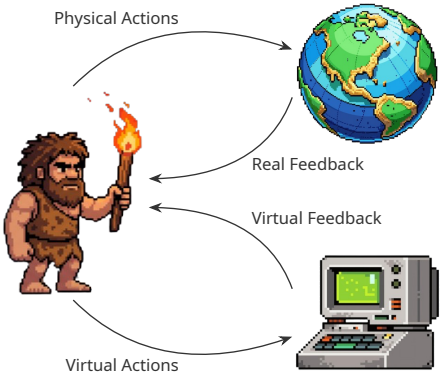
The Source of the Empirical Knowledge

A Clone of the Real World

Before Computers



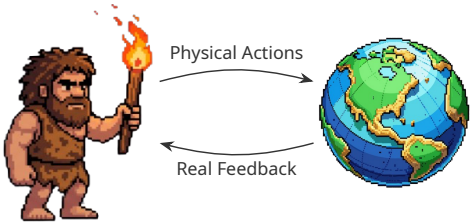
After Computers



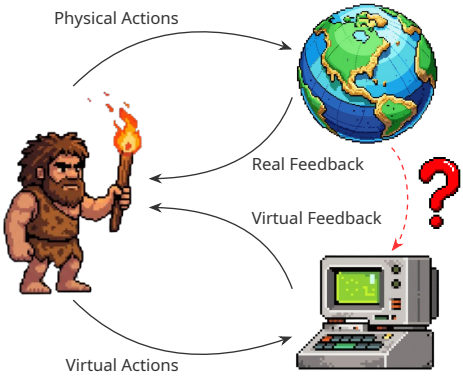
The Source of the Empirical Knowledge

A Clone of the Real World

Before Computers

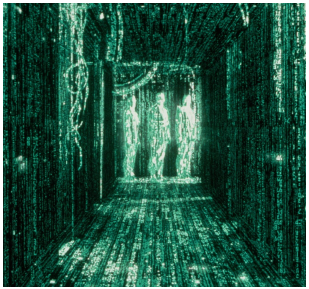


After Computers



The Source of the Empirical Knowledge

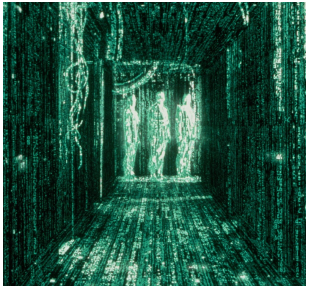
A Clone of the Real World



[1] *The Matrix* (1999)

The Source of the Empirical Knowledge

A Clone of the Real World

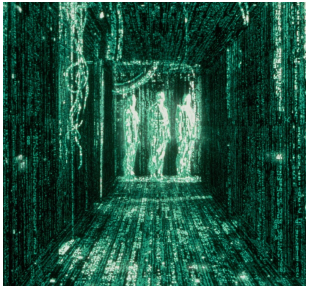




[1] *The Matrix* (1999)
[2] Maxis, *The Sims* (2000)

The Source of the Empirical Knowledge

A Clone of the Real World



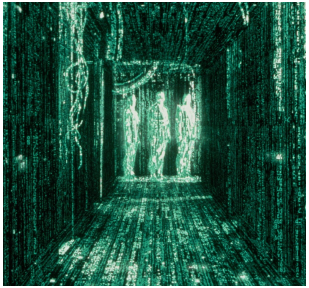




[1] *The Matrix* (1999)
[2] Maxis, *The Sims* (2000)

The Source of the Empirical Knowledge

A Clone of the Real World





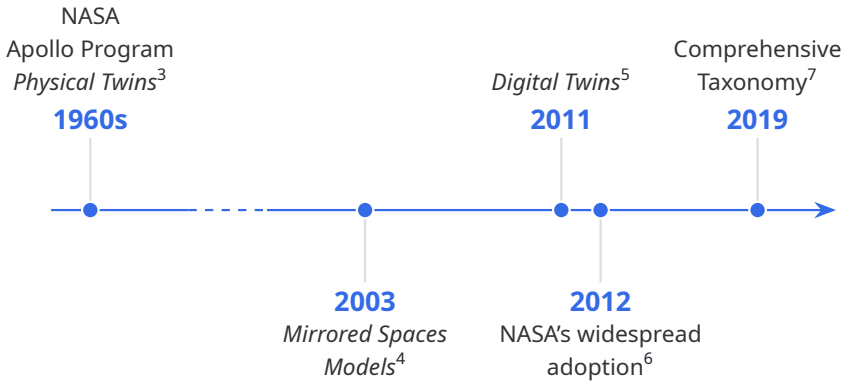


Is the Digital Twin (DT) a *clone* of (a part of) the real world?

Yes, but not only this...

[1] *The Matrix* (1999)
[2] Maxis, *The Sims* (2000)

The History of Digital Twins



[3] Rosen et al., "About The Importance of Autonomy and Digital Twins for the Future of Manufacturing" (2015).
[4] Grieves, "Digital twin: manufacturing excellence through virtual factory replication" (2014).
[5] Grieves, *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management* (2011).
[6] Glaessgen and Stargel, "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles" (2012).
[7] Barricelli et al., "A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications" (2019).

The History of Digital Twins

Definition (Digital Twin)

A virtual replica of a physical system, spanning the entire lifecycle of a product or process while remaining continuously synchronized with its real-world counterpart.[8]

[8] Aheleroff et al., “Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model” (2021)

The History of Digital Twins

Definition (Digital Twin)

A virtual replica of a physical system, spanning the entire lifecycle of a product or process while remaining continuously synchronized with its real-world counterpart.[8]



[8] Aheleroff et al., “Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model” (2021)

The History of Digital Twins

Definition (Digital Twin)

A virtual replica of a physical system, spanning the entire lifecycle of a product or process while remaining continuously synchronized with its real-world counterpart.[8]



- Reactive nature: **monitoring**, **diagnostics**, and **fault detection**
- Proactive nature: **optimization** and **decision-making** within *what-if* scenarios

[8] Aheleroff et al., "Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model" (2021)

The History of Digital Twins

Definition (Digital Twin)

A virtual replica of a physical system, spanning the entire lifecycle of a product or process while remaining continuously synchronized with its real-world counterpart.[8]

Machine Learning

"The study of computer algorithms that allow computer programs to automatically improve through experience" (Mitchell, *Machine learning*)

- Reactive nature: **monitoring**, **diagnostics**, and **fault detection**
- Proactive nature: **optimization** and **decision-making** within *what-if* scenarios

[8] Aheleroff et al., "Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model" (2021)

The History of Digital Twins

Several Application Domains





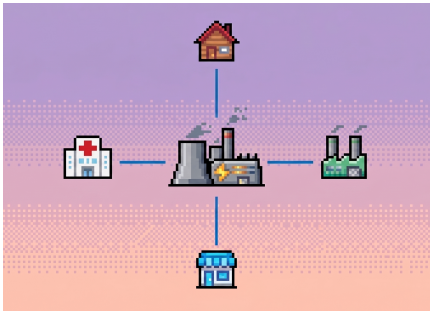




The Role of Batteries in the Energy Transition

A Radical Change of Paradigm

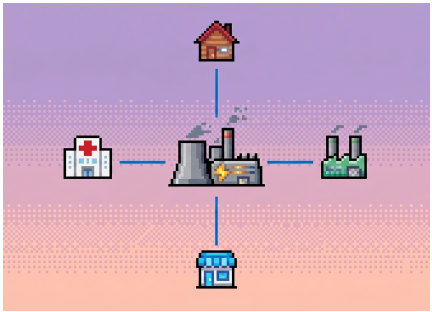
Centralized Power Grid



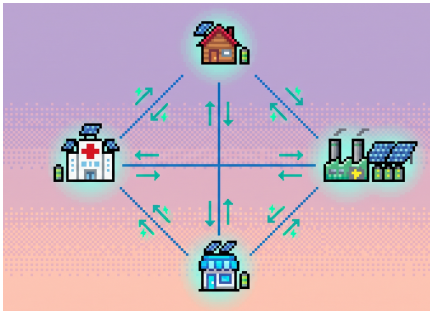
The Role of Batteries in the Energy Transition

A Radical Change of Paradigm

Centralized Power Grid



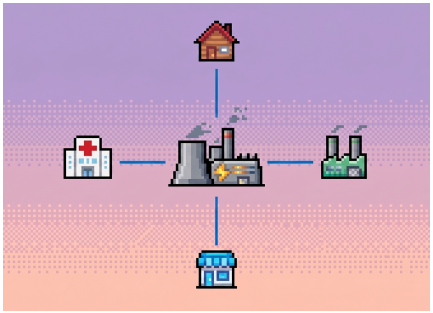
Decentralized Smart Grid



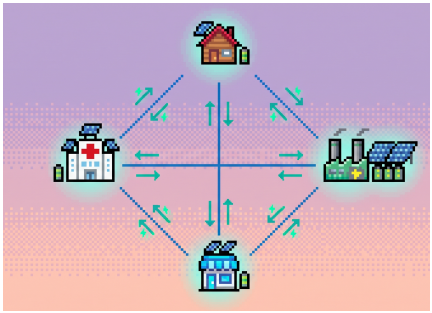
The Role of Batteries in the Energy Transition

A Radical Change of Paradigm

Centralized Power Grid



Decentralized Smart Grid



Batteries are crucial to guarantee **flexibility** and **stability** of the grid
DTs can enhance the **performance**, **reliability**, and **longevity** of batteries

Digital Twin of Battery Energy Storage Systems

What is this work about?

Challenges of Battery Simulation

- Non-linear **degradation**
- Complex electrochemical dynamics
- **Real-time** data acquisition
- **Computational efficiency** for real-time applications

Digital Twin of Battery Energy Storage Systems

What is this work about?

Challenges of Battery Simulation

- Non-linear **degradation**
- Complex electrochemical dynamics
- **Real-time** data acquisition
- **Computational efficiency** for real-time applications

Objectives and Contributions

- Design an **holistic DT simulator** for battery energy storage systems
- Integrate data-driven approaches for **real-time monitoring and adaptation**
- **Optimize battery energy management** within the context of smart grids

ROADMAP

Agenda

1. Design and Validation of the Digital Twin Framework
2. Machine Learning Integration for Real-Time Monitoring and Adaptation
3. Battery Energy Management through Reinforcement Learning
4. Conclusions and Discussion

1

Design and Validation of the Digital Twin Framework

A Generalistic Design for Smart Grids

Previous Works and Challenges

The current **literature** about DTs for the energy sector typically exhibits:

- a strict dichotomy between **data-driven** and **model-based** approaches[9]
- specificity to **single use cases** (e.g., SOC estimation, thermal management, structural integrity)
- lack of **generalized** architectures

[9] Semeraro et al., "Digital twin in battery energy storage systems: Trends and gaps detection through association rule mining" (2023).
[10] Park et al., "Digital Twin for Operation of Microgrid: Optimal Scheduling in Virtual Space of Digital Twin" (2020).
[11] Padmawansa et al., "Battery Energy Storage Capacity Estimation for Microgrids Using Digital Twin Concept" (2023).

The current **literature** about DTs for the energy sector typically exhibits:

- a strict dichotomy between **data-driven** and **model-based** approaches[9]
- specificity to **single use cases** (e.g., SOC estimation, thermal management, structural integrity)
- lack of **generalized** architectures

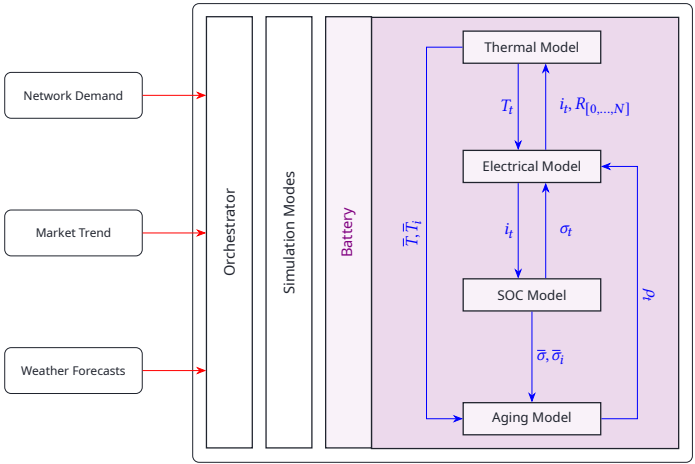
Key **challenges for Smart Grid applications** include:

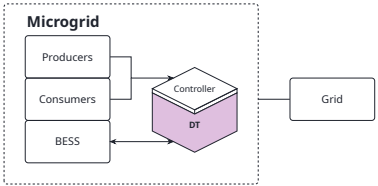
- modeling **coupled multi-physics phenomena**[10]
- managing **computational complexity** for real-time ML and optimization[11]

[9] Semeraro et al., "Digital twin in battery energy storage systems: Trends and gaps detection through association rule mining" (2023).
[10] Park et al., "Digital Twin for Operation of Microgrid: Optimal Scheduling in Virtual Space of Digital Twin" (2020).
[11] Padmawansa et al., "Battery Energy Storage Capacity Estimation for Microgrids Using Digital Twin Concept" (2023).

1 The ErNESTO Digital Twin

Grey-Box Modular Framework

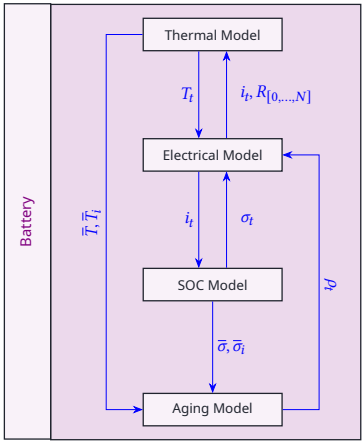




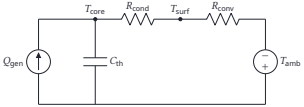
[12] Salaorni, Bianchi, Colnago, et al., "A novel digital twin for battery energy storage systems in micro-grids" (2025)

1 The ErNESTO Digital Twin

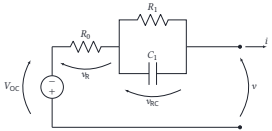
An Instance of the Framework



Thermal Model



Electrical Model



SOC Estimator

Coulomb Counting[13]

$$\sigma_t := \sigma_{t-1} - \frac{i_t}{C^{\text{nom}}} \Delta t$$

SOH Estimator

Semi-empirical aging model[14]

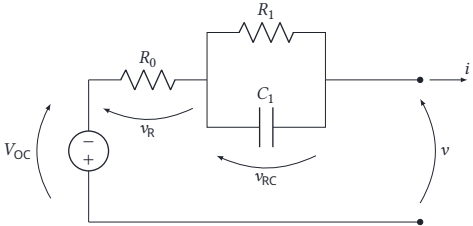
$$f_{\text{cal},t} = f_{\text{cal}}(t, \overline{\sigma}, \overline{T})$$

$$f_{\text{cyc},t} = \sum_{i}^{n_{\text{cyc}}} n_i f_{\text{cyc}}(\delta_i, \overline{\sigma}_i, \overline{T}_i)$$

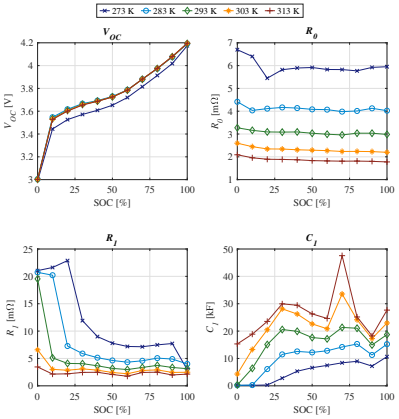
[13] Ng et al., "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries" (2009)
[14] Xu et al., "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment" (2018)

1 The ErNESTO Digital Twin

Parametrization of the Electrical Model



- V_{OC} is the open-circuit voltage
- R_0 is the ohmic resistance
- R_1 is the polarization resistance
- C_1 is the polarization capacitance



Reminder

R_0, R_1, and C_1 depend on the **SOC**, **temperature**, and **SOH** of the battery

The experimental campaign also includes:

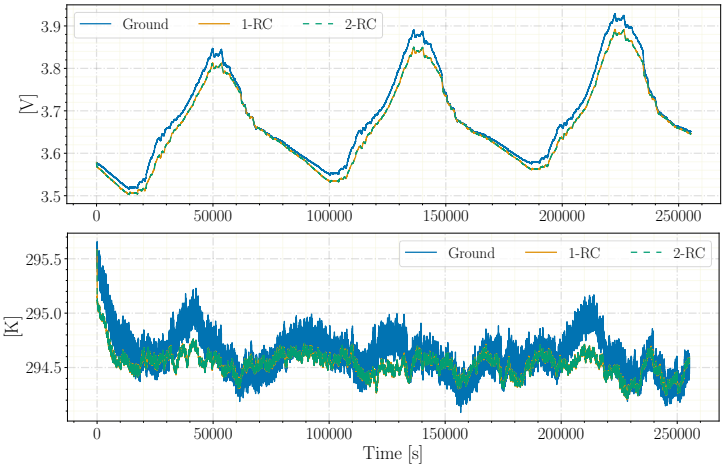
- Validation using a real-world **residential prosumer profile**
- Validation using a **microgrid power profile** from the RSE testbed facility
- Ablation studies and validation of the **thermal model** with experimental data
- Ablation studies and validation of the **aging model** with experimental data

Objectives of the Experiments

- Validate the **accuracy** of the DT sub-models
- Validate the **necessity** of the coupled multi-physics modeling (thermal and aging in addition to the electrical model)

1 Experiments and Results

Residential Prosumer Profile



1-RC and 2-RC models achieve similar performance

MAPE

$\approx 0.50\%$ (electrical),

$\approx 0.04\%$ (thermal)

MaAE

≈ 0.045 V (electrical),

≈ 0.54 K (thermal)

Contributions

ErNESTO is demonstrated to be:

- A **modular** and **generalized** architecture
- A **high-fidelity simulator** of battery dynamics
- Validated **interconnected sub-models** simulating electrical, thermal, and aging dynamics

Assumptions

It requires prior full knowledge of the battery's **model parameters** and their **dependencies**

2

Machine Learning Integration for Real-Time Monitoring and Adaptation

From a Simulator to a Digital Twin

Following the Dynamics Evolution

Definition (Digital Twin)

A virtual replica of a physical system, **spanning the entire lifecycle** of a product or process while **remaining continuously synchronized** with its real-world counterpart

Definition (Digital Twin)

A virtual replica of a physical system, **spanning the entire lifecycle** of a product or process while **remaining continuously synchronized** with its real-world counterpart

In a supervised learning setting, where $X_t \in \mathbb{R}^d$ is the **input feature vector**, y_t is the **target variable**, and $P_t(X, y)$ is their joint probability distribution generating the data stream $\{(X_t, y_t)\}_{t=1}^{\infty}$, **concept drift** occurs when:

$$\exists t \ : \ P_t(X, y) \neq P_{t+1}(X, y)$$

[15] Gama et al., "A survey on concept drift adaptation" (2014); Moreno-Torres et al., "A unifying view on dataset shift in classification" (2012).

Definition (Digital Twin)

A virtual replica of a physical system, **spanning the entire lifecycle** of a product or process while **remaining continuously synchronized** with its real-world counterpart

In a supervised learning setting, where $X_t \in \mathbb{R}^d$ is the **input feature vector**, y_t is the **target variable**, and $P_t(X, y)$ is their joint probability distribution generating the data stream $\{(X_t, y_t)\}_{t=1}^{\infty}$, **concept drift** occurs when:

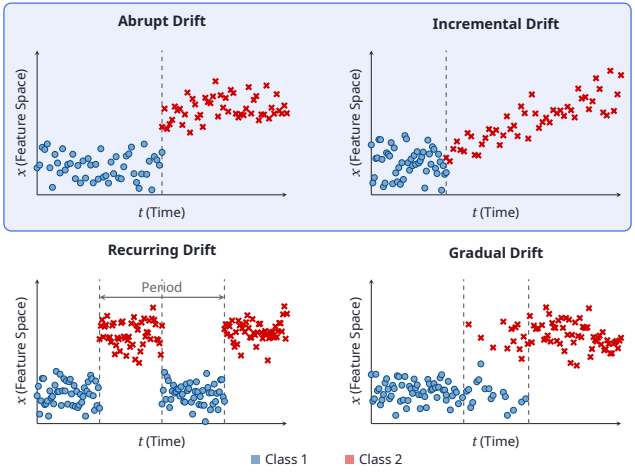
$$\exists t \ : \ P_t(X, y) \neq P_{t+1}(X, y)$$

- **Virtual Drift**: The change affects $P_t(X)$, i.e., the input distribution
- **Real Drift**: The change affects $P_t(y|X)$, i.e., the posterior distribution (system dynamics).[15]

[15] Gama et al., "A survey on concept drift adaptation" (2014); Moreno-Torres et al., "A unifying view on dataset shift in classification" (2012).

2 From a Simulator to a Digital Twin

Different Kind of Drifts



Parameter identification has to be made at **runtime**

- **Static** calibration (parameter identification and genetic algorithms)[16]
- **Online** estimation (MLP and sliding window learning)[17]

↪ Main problem: **verification latency**

[16] Hossain et al., "A parameter extraction method for the Thevenin equivalent circuit model of Li-ion batteries" (2019).
[17] Mussi et al., "A voltage dynamic-based state of charge estimation method for batteries storage systems" (2021).

Parameter identification has to be made at **runtime**

- **Static** calibration (parameter identification and genetic algorithms)[16]
- **Online** estimation (MLP and sliding window learning)[17]

↪ Main problem: **verification latency**

How to **assess drifts** and **detect faults**?

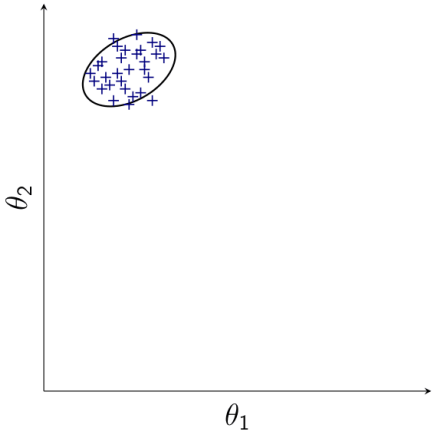
- supervised classification of **known fault patterns**[18]
- unsupervised **novelty detection** (*reconstruction-based* and *clustering-based*)[19]

↪ Main problem: focus on **rejecting anomalies** and **flagging for inspection**

[16] Hossain et al., "A parameter extraction method for the Thevenin equivalent circuit model of Li-ion batteries" (2019).
[17] Mussi et al., "A voltage dynamic-based state of charge estimation method for batteries storage systems" (2021).
[18] Zhang et al., "Deep learning-based prognostic approach for lithium-ion batteries with adaptive time-series prediction and on-line validation" (2020).
[19] Peng et al., "Unsupervised adaptive fleet battery pack fault detection with concept drift under evolving environment" (2024).

2 Real-Time Monitoring and Adaptation

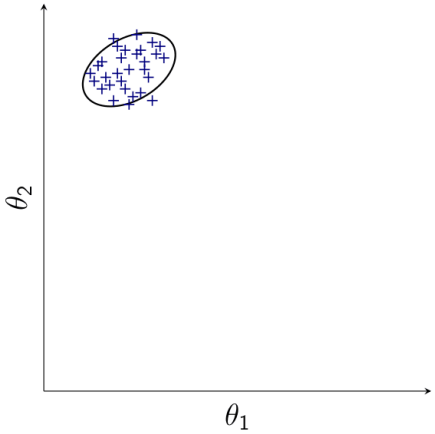
Core Idea



[20] Alippi et al., “A Self-Building and Cluster-Based Cognitive Fault Diagnosis System for Sensor Networks” (2014)

2 Real-Time Monitoring and Adaptation

Core Idea



[20] Alippi et al., “A Self-Building and Cluster-Based Cognitive Fault Diagnosis System for Sensor Networks” (2014)

2 Real-Time Monitoring and Adaptation

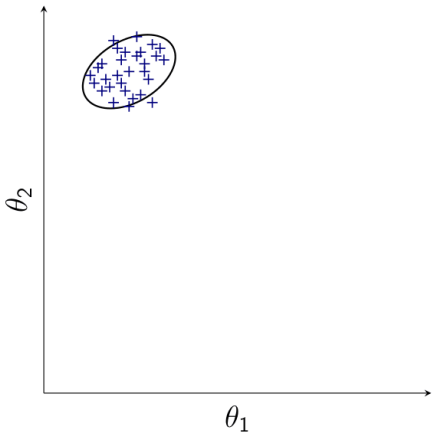
Core Idea

Tracking in the Parameter Space

While voltage residuals indicate that a mismatch in the system's dynamics exists, parameter shifts (R_0, R_1, C_1) reveal **why** and **where** it occurs

Online Fault Diagnosis System

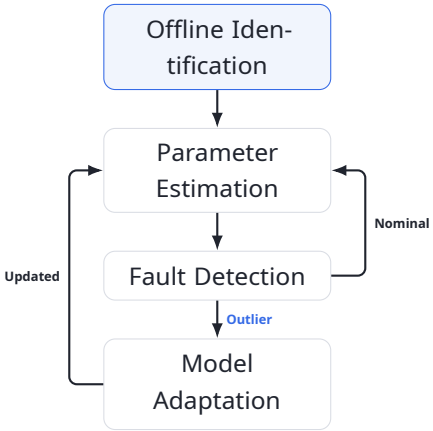
The fault dictionary is not pre-compiled but is *generated at runtime*, through an **evolving-clustering** method.[20] Explicitly accounting for the irreversible degradation process, each new faulty cluster is promoted to a **new valid operating mode**



[20] Alippi et al., "A Self-Building and Cluster-Based Cognitive Fault Diagnosis System for Sensor Networks" (2014)

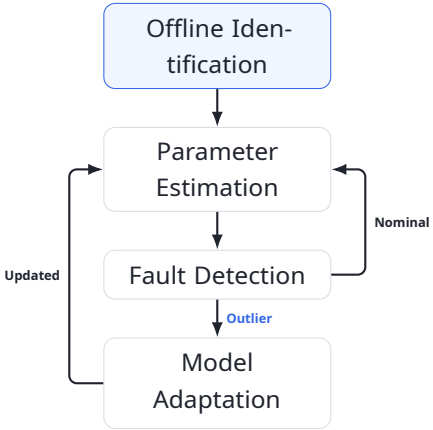
2 Real-Time Monitoring and Adaptation

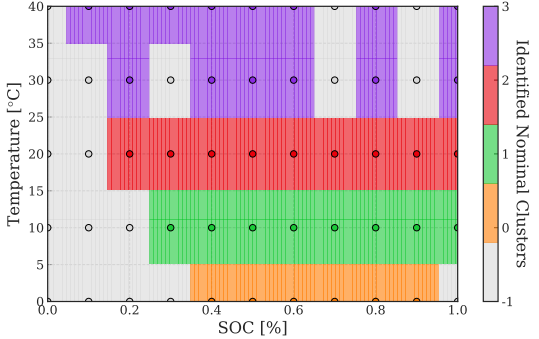
The Evolving-Clustering Routine



2 Real-Time Monitoring and Adaptation

The Evolving-Clustering Routine

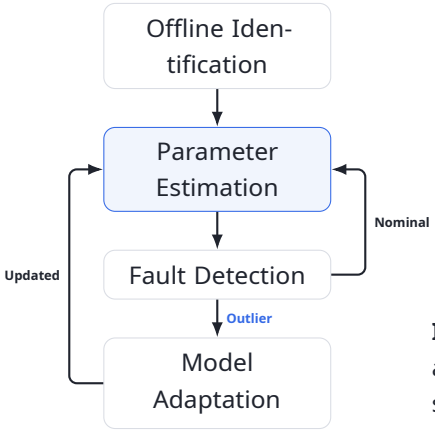




Ψ_0 is the initial **nominal cluster**

2 Real-Time Monitoring and Adaptation

The Evolving-Clustering Routine



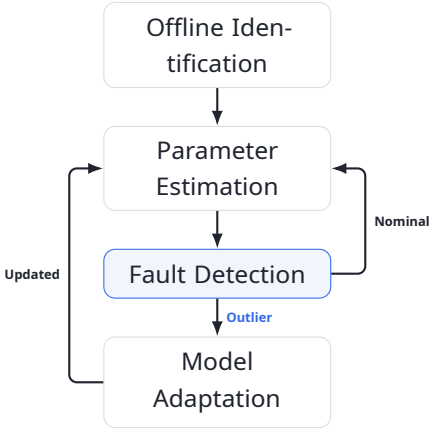
1. Normal DT simulation given a power profile
2. Collection of a window of samples $\mathcal{Z}_W = \{(\mathbf{x}_h, y_h)\}_{h=1}^{W}$
3. Estimation of a new candidate vector $\hat{\theta}_t$ such that:

$$\hat{\theta}_t = \arg\min_{\phi} \left[\frac{1}{W} \sum_{i=1}^{W} \left(y_i - \hat{y}_i(\phi, \mathbf{x}_i) \right)^2 + \xi \|\phi - \theta_{t-1}\|_2^2 \right]$$

In practice: We adopt the L-BFGS-B optimization algorithm and parameter normalization for numerical stability

2 Real-Time Monitoring and Adaptation

The Evolving-Clustering Routine



1. Compute the Mahalanobis distance of $\hat{\theta}_t$ from the centroid of Ψ_k:

$$D_M(\hat{\theta}_t) = \sqrt{(\hat{\theta}_t - \mu_k)^\top \Sigma_k^{-1} (\hat{\theta}_t - \mu_k)}\,.$$

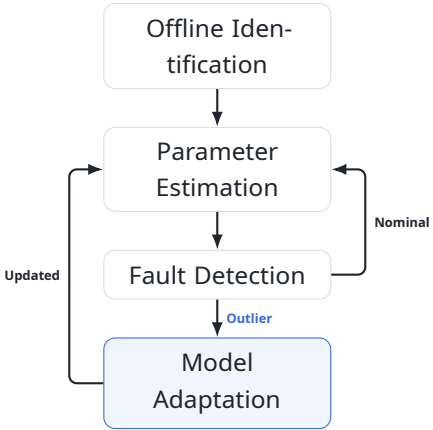
2. Compute a dynamical threshold τ_M accounting for the nominal cluster size

$$\tau_M = \sqrt{\frac{d(n_k^2 - 1)}{n_k(n_k - d)} F_{d,n_k-d,\alpha_s}}$$

3. If $D_M(\hat{\theta}_t) > \tau_M$, the estimate is an **outlier**; otherwise, it is accepted as an **inlier**

2 Real-Time Monitoring and Adaptation

The Evolving-Clustering Routine



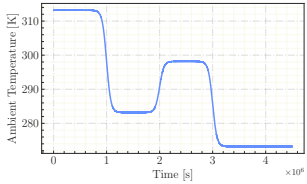
1. Statistical drift confirmation with a Two-Sample KS test. If $F_{\mathcal{O}} = F_{\Psi_k}$ is rejected at significance level α_{ks}, the drift is confirmed
2. Application of Mountain Method to select the cluster center that maximizes the potential density function:

$$\Omega(c_j) = \sum_{h=1}^{|\mathcal{O}|} \exp\left(-\frac{\|c_j - \hat{\theta}_h\|_\lambda^2}{2r^2}\right)$$

3. Minimum Covariance Determinant to reduce residual noise and identify the subset of samples that minimize the determinant of the covariance matrix

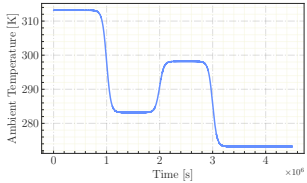
2 Experiments and Results

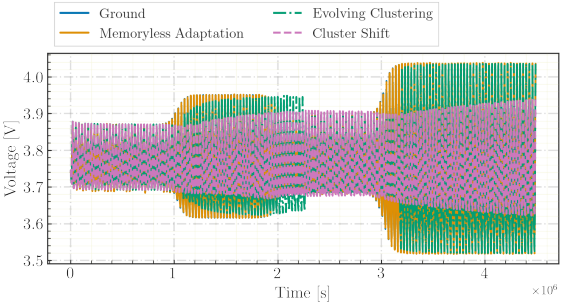
Simulated Scenario



2 Experiments and Results

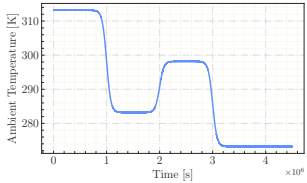
Simulated Scenario

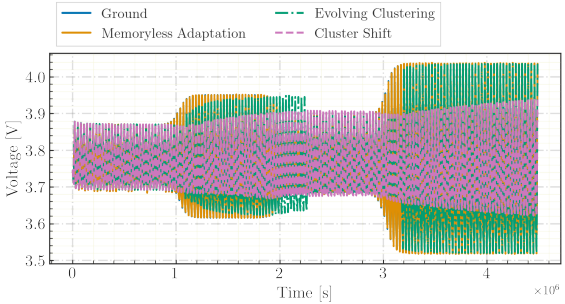


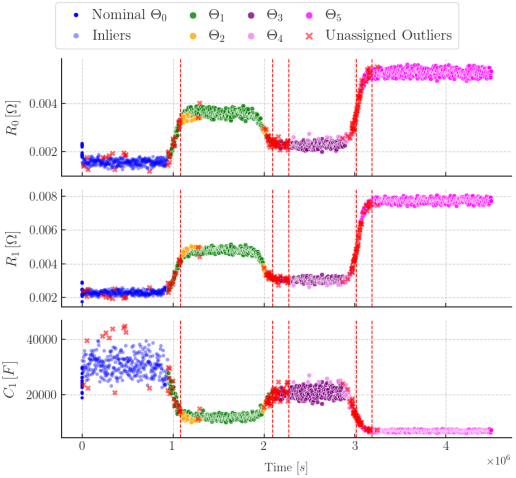


2 Experiments and Results

Simulated Scenario

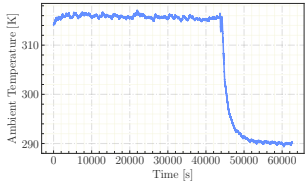






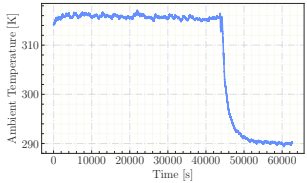
2 Experiments and Results

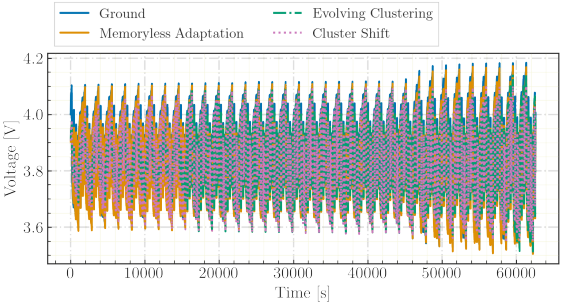
Real-World Scenario



2 Experiments and Results

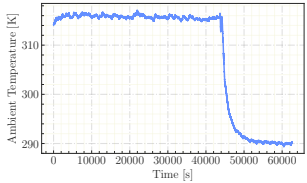
Real-World Scenario

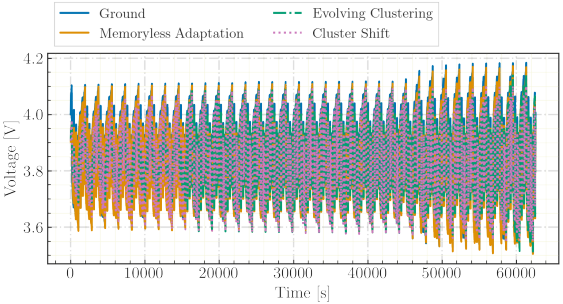


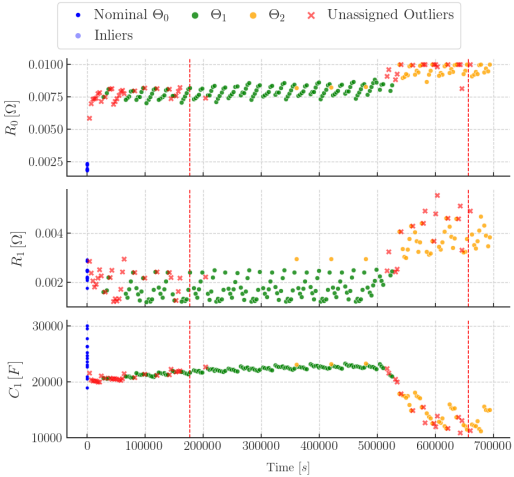


2 Experiments and Results

Real-World Scenario







Strengths and Achievements

The evolving-clustering routine demonstrates:

- **Performance** almost comparable with the optimal baseline (*memoryless*)
- **Superior interpretability** compared to traditional methods

Strengths and Achievements

The evolving-clustering routine demonstrates:

- **Performance** almost comparable with the optimal baseline (*memoryless*)
- **Superior interpretability** compared to traditional methods

Current Limitations and Research Directions

The proposed framework exhibits a few practical constraints:

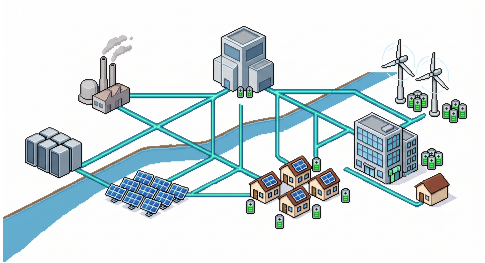
- Historical nominal clusters are currently **not retained or reused**
- Subject to **verification latency**, which can reduce real-time responsiveness

3

Battery Energy Management through Reinforcement Learning

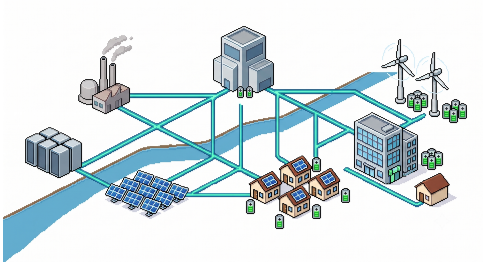
3 Energy Management in Smart Grids

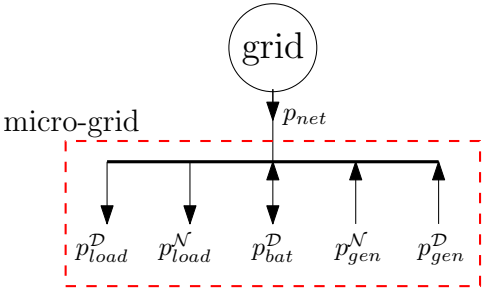
An Inventory Management Problem



3 Energy Management in Smart Grids

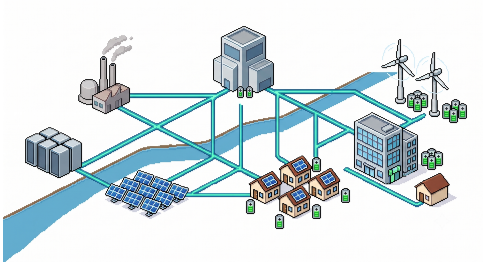
An Inventory Management Problem

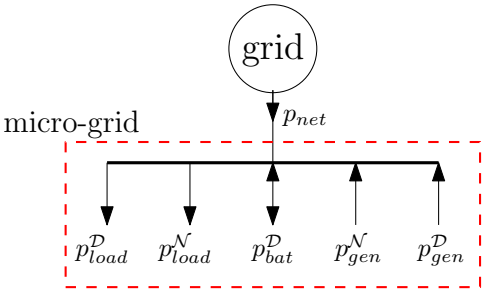




3 Energy Management in Smart Grids

An Inventory Management Problem





Battery as an inventory

The battery is a **buffer** that can be used to store/withdraw the surplus energy in order to deal with **future uncertainties** in generation and demand

Energy Management in Microgrids

A Single Agent Control Problem

Optimizing battery behavior to **maximize the revenue** from the trading with the energy market and **minimizing battery costs** associated with degradation

$$J_{\mathcal{T}}(\pi) = \mathbb{E}_{\tau \sim \pi}\left[\sum_{t=0}^{T-1} \left(r_{\text{trad},t} + r_{\text{deg},t}\right)\right]$$

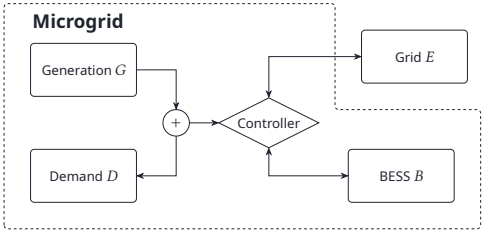
[21] Salaorni, Bianchi, Restelli, et al., "A Reinforcement Learning Approach for Optimal Control in Microgrids" (2025)

Energy Management in Microgrids

A Single Agent Control Problem

Optimizing battery behavior to **maximize the revenue** from the trading with the energy market and **minimizing battery costs** associated with degradation

$$J_{\mathcal{J}}(\pi) = \mathbb{E}_{\tau\sim\pi}\left[\sum_{t=0}^{T-1}\left(r_{\text{trad},t} + r_{\text{deg},t}\right)\right]$$



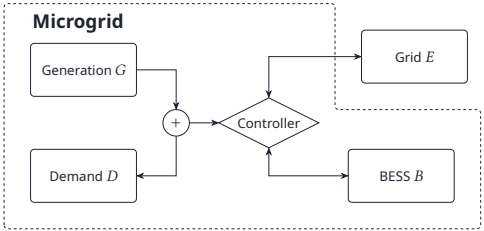
[21] Salaorni, Bianchi, Restelli, et al., "A Reinforcement Learning Approach for Optimal Control in Microgrids" (2025)

3 Energy Management in Microgrids

A Single Agent Control Problem

Optimizing battery behavior to **maximize the revenue** from the trading with the energy market and **minimizing battery costs** associated with degradation

$$J_{\mathcal{T}}(\pi) = \mathbb{E}_{\tau\sim\pi}\left[\sum_{t=0}^{T-1}\left(r_{\text{trad},t} + r_{\text{deg},t}\right)\right]$$



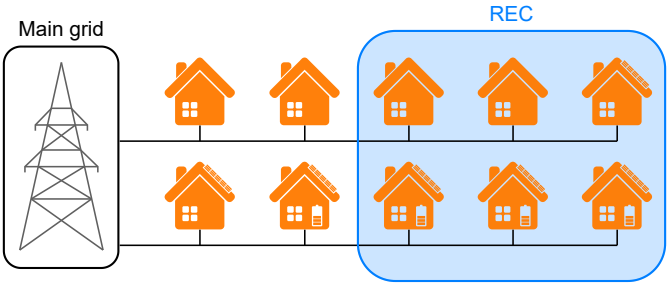
In a preliminary work[21] we:

- Formulated the problem as a **Markov Decision Process**
- Employed **reinforcement learning** (PPO) to learn an optimal policy
- Validated the approach against SOTA and rule-based methods using real-world data

[21] Salaorni, Bianchi, Restelli, et al., "A Reinforcement Learning Approach for Optimal Control in Microgrids" (2025)

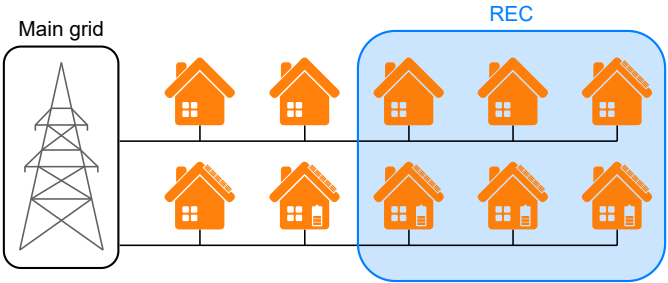
3 Extension to a Renewable Energy Community

Balancing Collaboration and Self-Interest



3 Extension to a Renewable Energy Community

Balancing Collaboration and Self-Interest



Social Welfare (for fairness)

$$\mathscr{W} := \sum_{n \in \mathcal{N}_A} J_n$$

Self-Consumption (for performance)

$$U_t := \min \left\{ \sum_{n \in \mathcal{N}} p^+_{n,t}, \sum_{n \in \mathcal{N}} p^-_{n,t} \right\}$$

Aligning Prosumers' Interests

Two Phase Stochastic Game

$$\mathscr{M}_{\text{TPSG}} := \langle I, \mathcal{S}, \mathcal{A}, P, \mathcal{R}, \Omega, \mathcal{O}, \gamma \rangle$$

- $I := \mathcal{N}_A \cup \{C\}$ is the set of agents
- Standard stochastic game formulations typically assume assume **simultaneous** actions[22]
- Node agents collect an additional $r^{\text{inc}}_{n,t}$ reward term, which is leveraged by the REC agent to **align prosumers' interests** and **foster collaboration**

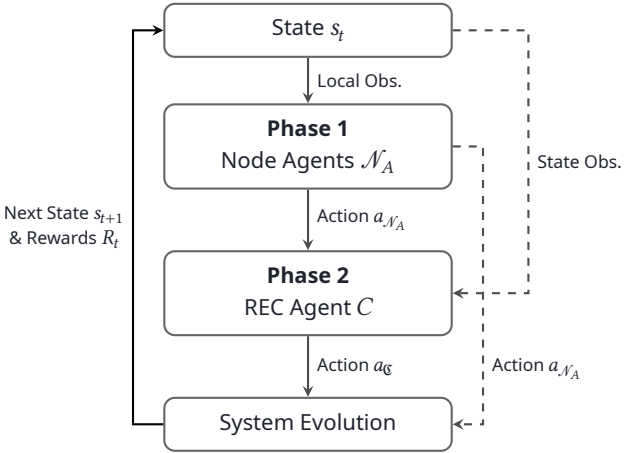
[22] Yang et al., *Game-Theoretic Multiagent Reinforcement Learning* (2025)

3 Aligning Prosumers' Interests

Two Phase Stochastic Game

$$\mathscr{M}_{\text{TPSG}} := \langle I, \mathcal{S}, \mathscr{A}, P, \mathscr{R}, \Omega, \mathcal{O}, \gamma \rangle$$

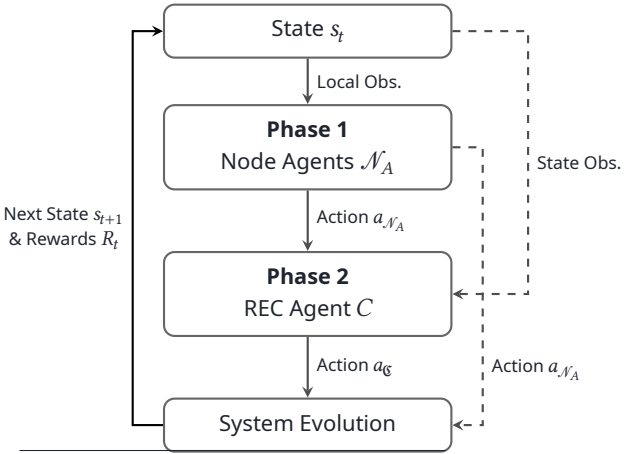
- $I := \mathcal{N}_A \cup \{C\}$ is the set of agents
- Standard stochastic game formulations typically assume assume **simultaneous** actions[22]
- Node agents collect an additional $r_{n,t}^{\text{inc}}$ reward term, which is leveraged by the REC agent to **align prosumers' interests** and **foster collaboration**



[22] Yang et al., *Game-Theoretic Multiagent Reinforcement Learning* (2025)

3 The INAIA Algorithm

Structure and Key Innovations

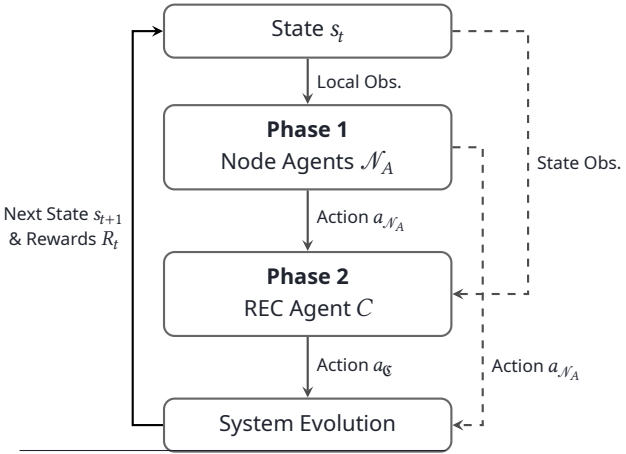


[23] Foerster et al., *Learning with Opponent-Learning Awareness* (2018)

[24] Delpero et al., "Aligning Prosumer Interests in Renewable Energy Communities through Multi-Agent Reinforcement Learning and Reward Shaping" (2026)

3 The INAIA Algorithm

Structure and Key Innovations



Opponent-Awareness

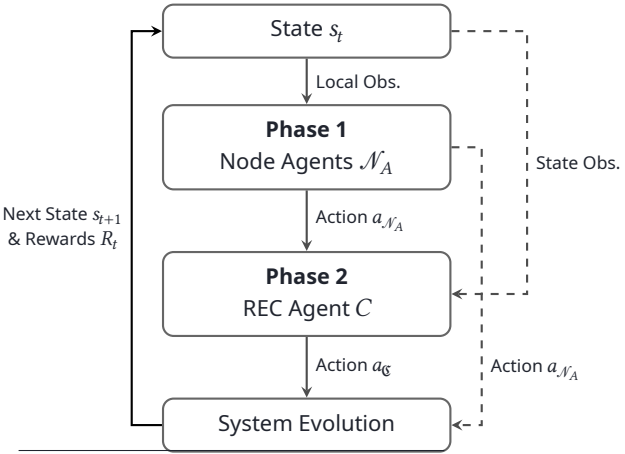
Combining the LOLA-style **lookahead** with the hierarchical structure of the REC setting allows the REC agent to **anticipate** the node agents' behavior and **shape** their rewards accordingly[23]

[23] Foerster et al., *Learning with Opponent-Learning Awareness* (2018)

[24] Delpero et al., "Aligning Prosumer Interests in Renewable Energy Communities through Multi-Agent Reinforcement Learning and Reward Shaping" (2026)

3 The INAIA Algorithm

Structure and Key Innovations



Opponent-Awareness

Combining the LOLA-style **lookahead** with the hierarchical structure of the REC setting allows the REC agent to **anticipate** the node agents' behavior and **shape** their rewards accordingly[23]

Myopic REC Agent

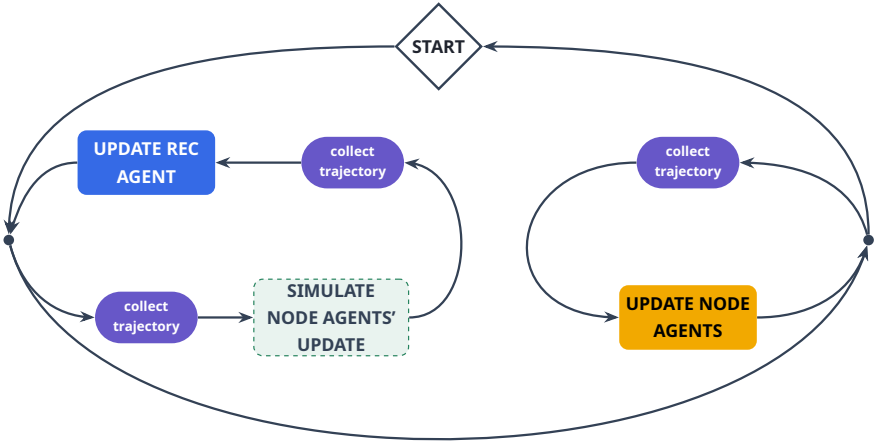
The REC agent **indirectly controls** the node agents through incentive allocation and acts **myopicly** to guide the node agents' learning

[23] Foerster et al., *Learning with Opponent-Learning Awareness* (2018)

[24] Delpero et al., "Aligning Prosumer Interests in Renewable Energy Communities through Multi-Agent Reinforcement Learning and Reward Shaping" (2026)

3 The INAIA Algorithm

Two-Phase Learning Dynamics



INAIA-OPT: different set of parameters for each node agent (more flexible and more unfair)
INAIA-SHARED: single set of parameters for all node agents (more scalable)

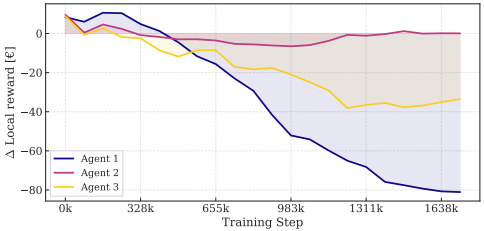
INAIA-OPT: different set of parameters for each node agent (more flexible and more unfair)
INAIA-SHARED: single set of parameters for all node agents (more scalable)

	Social Welfare [€]			Self-Consumption [MWh]		
Algorithm	**3**	**5**	**8**	**3**	**5**	**8**
RULE-BASED	-14.47 ± 3.64	-47.13 ± 3.43	-128.74 ± 31.59	-0.20 ± 0.05	-1.57 ± 0.12	-14.82 ± 0.42
PPO	-18.85 ± 3.99	-64.68 ± 4.97	-172.20 ± 36.03	-0.24 ± 0.05	-1.77 ± 0.15	-15.00 ± 0.52
LOLA	-59.16 ± 6.79	-171.28 ± 20.50	$-^{*}$	-0.28 ± 0.08	-2.15 ± 0.19	$-^{*}$
INAIA-SHARED	-3.84 ± 2.20	-30.48 ± 8.00	-88.96 ± 29.13	-0.14 ± 0.04	-1.43 ± 0.15	-7.56 ± 0.40

* Results are missing due to insufficient RAM availability, which prevents the completion of the model training

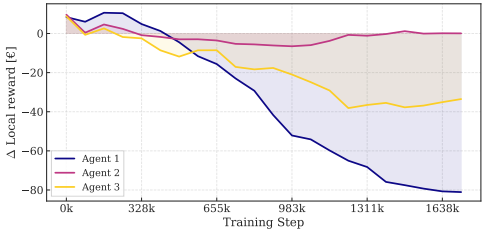
3 Experiments and Results

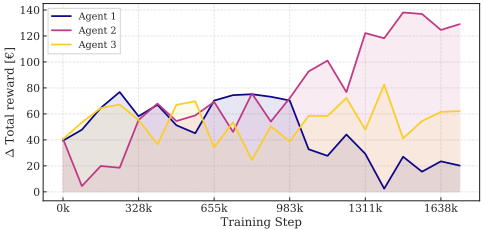
Convenience in Joining the Community



3 Experiments and Results

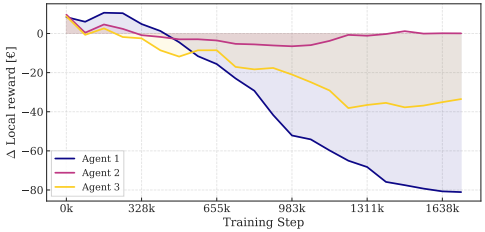
Convenience in Joining the Community

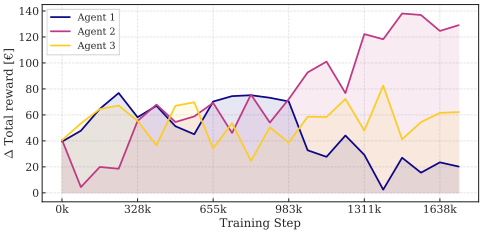




3 Experiments and Results

Convenience in Joining the Community





- The agents are **incentivized to join the community** as their reward increases with respect to the case in which they act independently
- This can lead to **unfairness** in the community, as some agents may benefit more than others
- The REC agent can be implemented with **fairness mechanisms**

Key Contributions

- Novel **Two-Phase Stochastic Game formulation** for the REC setting
- Novel **INAIA** algorithm for incentive allocation and reward shaping in an opponent-aware fashion
- Demonstrated the **scalability** and **performance** of the proposed approach

Key Contributions

- Novel **Two-Phase Stochastic Game formulation** for the REC setting
- Novel **INAIA** algorithm for incentive allocation and reward shaping in an opponent-aware fashion
- Demonstrated the **scalability** and **performance** of the proposed approach

Assumptions

- Simplified energy market model (day-ahead market)
- Power line capacity constraints are not considered
- The REC regulations are specific of the Italian context

4

Conclusions and Discussion

4 Wrapping Up

Contributions

The achievements of this work can be summarized as follows:

1. Design of **ErNESTO** Digital Twin
    - A **modular** and **generalized** architecture
    - A **high-fidelity simulator**
    - Validated **interconnected sub-models** through an extensive experimental campaign

4 Wrapping Up

Contributions

The achievements of this work can be summarized as follows:

1. Design of **ErNESTO** Digital Twin
   - A **modular** and **generalized** architecture
   - A **high-fidelity simulator**
   - Validated **interconnected sub-models** through an extensive experimental campaign
2. Integration of ErNESTO with **online learning** for real-time monitoring and adaptation
   - Effective and interpretable **evolving clustering** routine
   - Validated on simulated and real-world data

4 Wrapping Up

Contributions

The achievements of this work can be summarized as follows:

1. Design of **ErNESTO** Digital Twin
   - A **modular** and **generalized** architecture
   - A **high-fidelity simulator**
   - Validated **interconnected sub-models** through an extensive experimental campaign
2. Integration of ErNESTO with **online learning** for real-time monitoring and adaptation
   - Effective and interpretable **evolving clustering** routine
   - Validated on simulated and real-world data
3. Optimization of **energy management** within the context of smart grids
   - Design of a **reinforcement learning** solution for the microgrid energy management problem
   - Formalization of the problem as a **Two-Phase Stochastic Game** and design the novel **INAIA** algorithm for incentive allocation
   - Validation of scalability and performance employing real-world data

4 Wrapping Up

Future Works

1. **Deployment** of ErNESTO on a real-world microgrid testbed

4 Wrapping Up

Future Works

1. **Deployment** of ErNESTO on a real-world microgrid testbed
2. Integration of the monitoring pipeline with **fault detection** and **isolation systems**

4 Wrapping Up

Future Works

1. **Deployment** of ErNESTO on a real-world microgrid testbed
2. Integration of the monitoring pipeline with **fault detection** and **isolation systems**
3. Inclusion of **sustainability objectives** and introduction of specific **structure in the MDP formulation**[25]

[25] Maran et al., "Learning in Markov Decision Processes with Exogenous Dynamics" (2026)

4 Wrapping Up

Future Works

1. **Deployment** of ErNESTO on a real-world microgrid testbed
2. Integration of the monitoring pipeline with **fault detection** and **isolation systems**
3. Inclusion of **sustainability objectives** and introduction of specific **structure in the MDP formulation**[25]
4. Integration of **fairness mechanisms** and giving **theoretical guarantees** regarding the structural capabilities of RECs

[25] Maran et al., "Learning in Markov Decision Processes with Exogenous Dynamics" (2026)

4 Wrapping Up

Acknowledgements

- **Marcello** and **Francesco** for their wise guidance and supervision
- **Federico**, **Silvia**, and **Andrea** (RSE S.p.A.) for their dedication and support
- **My co-authors and colleagues** for their valuable contributions

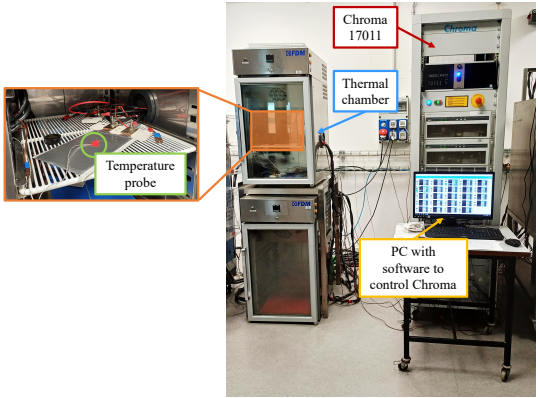


Thank You

Questions?

5 Digital Twin Validation

Experimental Setup



Parameter	Value
Nominal capacity (C_N)	20 Ah
Maximum voltage (V_{max})	4.2 V
Minimum voltage (V_{min})	3 V
Dimensions $L \times W \times H$	$196 \times 127 \times 7$ mm
Weight (M)	480 g

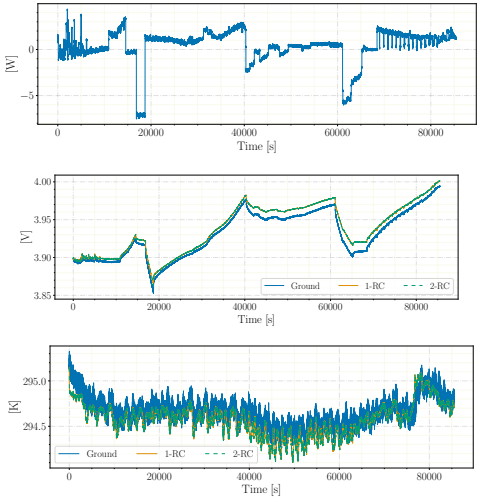
5 Digital Twin Validation

Microgrid Power Profile Profile

Experiment		Voltage			Temperature		
Type	**Test**	MAE[V]	MAPE[%]	MaAE[V]	MAE[K]	MAPE[%]	MaAE[K]
PV	1-RC	0.0185	0.50	0.0448	0.1231	0.04	0.5443
PV	2-RC	0.0186	0.50	0.0446	0.1231	0.04	0.5443
DER-TF	1-RC	0.0072	0.18	0.0166	0.0994	0.03	0.4673
DER-TF	2-RC	0.0072	0.18	0.0165	0.0994	0.03	0.4673
TJ	1-RC w/o temp	0.01568	0.41	0.0460	-	-	-
TJ	2-RC w/o temp	0.0161	0.42	0.0461	-	-	-
TJ	1-RC	0.0102	0.27	0.0302	0.5309	0.17	6.4563
TJ	2-RC	0.0108	0.28	0.0295	0.5309	0.17	6.4563

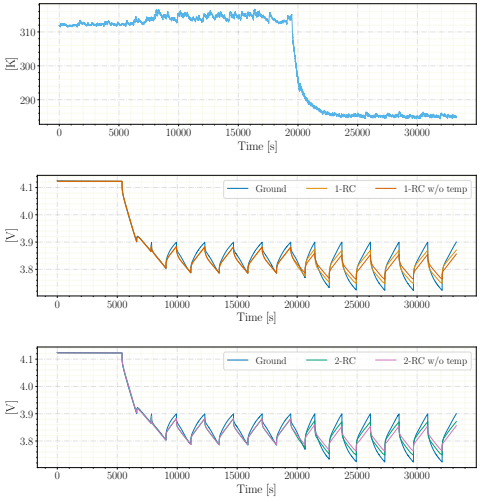
5 Digital Twin Validation

Microgrid Power Profile Profile



5 Digital Twin Validation

Thermal Model Validation



Evolving Clustering Algorithm

Algorithm 1 ERNESTO's Evolving Clustering Routine

Input: Initial training data $\mathcal{Z}_M$

1: Estimate initial parameter cluster $\boldsymbol{\Theta}_0 \leftarrow \text{OFFLINEIDENTIFICATION}(\mathcal{Z}_M)$
2: Compute statistics $\Psi_0 \leftarrow (\text{mean}(\boldsymbol{\Theta}_0), \text{cov}(\boldsymbol{\Theta}_0), M, t_0)$
3: $\boldsymbol{\Psi} \leftarrow \{\Psi_0\}$ ▷ *List of known states*
4: $k \leftarrow 0$ ▷ *Index of active state*
5: $t \leftarrow 0$ ▷ *Adaptation epoch index*
6: $\mathcal{O} \leftarrow \emptyset$ ▷ *Outlier buffer*
7: Update DT model $\boldsymbol{\theta} \leftarrow \boldsymbol{\mu}_{\Psi_0}$
8: **while** DT is running **do**
9: Collect data window $\mathcal{Z}_W = \{(\mathbf{x}_h, y_h)\}_{h=1}^{W}$
10: $\hat{\boldsymbol{\theta}}_t \leftarrow \text{PARAMETERESTIMATION}(\mathcal{Z}_W)$
11: is_outlier $\leftarrow \text{COGNITIVEFAULTDETECTION}(\hat{\boldsymbol{\theta}}_t, \Psi_k)$
12: **if not** is_outlier **then**
13: $\boldsymbol{\Theta}_k \leftarrow \boldsymbol{\Theta}_k \cup \{\hat{\boldsymbol{\theta}}_t\}$ ▷ *Update nominal cluster*
14: Update statistics $\Psi_k \leftarrow (\text{mean}(\boldsymbol{\Theta}_k), \text{cov}(\boldsymbol{\Theta}_k), |\boldsymbol{\Theta}_k|, t)$
15: **else**
16: $\mathcal{O} \leftarrow \mathcal{O} \cup \{\hat{\boldsymbol{\theta}}_t\}$ ▷ *Buffer outlier*
17: $\Psi_{\text{new}} \leftarrow \text{ONLINEMODELADAPTATION}(\mathcal{O})$
18: **if** $\Psi_{\text{new}} \neq$ NULL **then**
19: $\boldsymbol{\Psi} \leftarrow \boldsymbol{\Psi} \cup \{\Psi_{\text{new}}\}$ ▷ *Update known states*
20: $k \leftarrow k + 1$ ▷ *Switch to new active state*
21: $\mathcal{O} \leftarrow \emptyset$ ▷ *Clear buffer (or partial clear)*
22: **end if**
23: **end if**
24: $\boldsymbol{\theta} \leftarrow \boldsymbol{\mu}_k$ ▷ *Recalibrate DT*
25: $t \leftarrow t + 1$ ▷ *Update adaptation epoch index*
26: **end while**

Algorithm 2 ONLINEMODELADAPTATION

Input: Outlier buffer $\mathcal{O}$, Nominal state Ψ_k, Threshold α_{ks}, Grid parameters (r, λ)

Project data: $\mathcal{Z}_{\text{proj}} \leftarrow \text{PCA}(\mathcal{O} \cup \text{samples}(\Psi_k))$
$p_{\text{val}} \leftarrow \text{KS-TEST}(\mathcal{Z}_{\text{proj}}[\mathcal{O}], \mathcal{Z}_{\text{proj}}[\Psi_k])$
if $p_{\text{val}} > \alpha_{ks}$ **then**
  return NULL ▷ *No significant drift detected*
else
  $\tilde{\mathbf{\Theta}} \leftarrow \text{MOUNTAINMETHOD}(\mathcal{O}, r, \lambda)$
  if $|\tilde{\mathbf{\Theta}}| \leq d$ **then**
    return NULL ▷ *Cluster too sparse*
  else
    $\bar{\mathbf{\Theta}} \leftarrow \text{MCD}(\tilde{\mathbf{\Theta}})$
    if $|\bar{\mathbf{\Theta}}| > d + 1$ **then**
      $\Psi_{\text{new}} \leftarrow (\text{mean}(\bar{\mathbf{\Theta}}), \text{cov}(\bar{\mathbf{\Theta}}), |\bar{\mathbf{\Theta}}|, t_{\text{now}})$
      return Ψ_{new}
    else
      return NULL ▷ *Cluster too sparse or singular*
    end if
  end if
end if

Optimal Control within a Microgrid

Markov Decision Process Formulation

Controller Objective

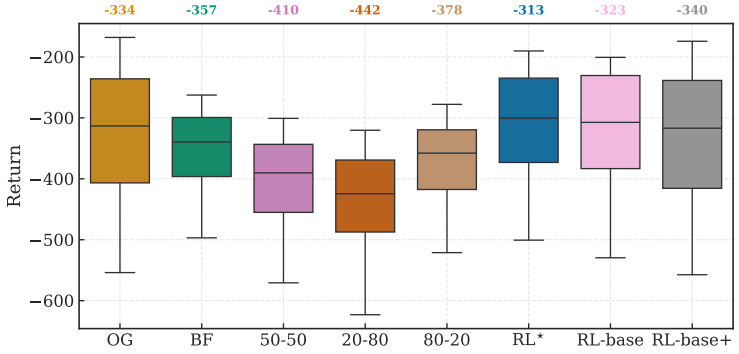
Maximize the revenue from the trading and minimizing the battery costs

$$\mathscr{M} = \langle \mathcal{S}, \mathcal{A}, P, R, \gamma, \mu_0 \rangle$$

- $s_t \in \mathcal{S}$ includes battery **state** variables, **generation** and **demand** estimations, **market prices**, and temporal signals
- $a_t \in [0, 1]$ represents the **proportion of surplus power** to dispatch (or withdraw) to (from) the BESS
- $r_t = r_{\text{trad},t} + r_{\text{deg},t} + \omega r_{\text{clip},t}$ rewards the **gain for trading** and **penalizes degradation** and overshooted actions

5 Optimal Control within a Microgrid

Demonstration of Reinforcement Learning Capabilities



Our formulation ($RL^{\star}$), exploiting PPO, demonstrates **superior performance** compared to the rule-based and state-of-the-art methods[25]

[25] Schulman et al., "Proximal policy optimization algorithms" (2017).

INAIA Pseudocode

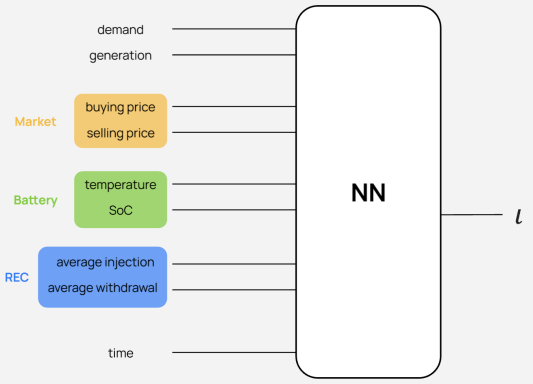
Algorithm 4 Pseudocode of INAIA algorithm.

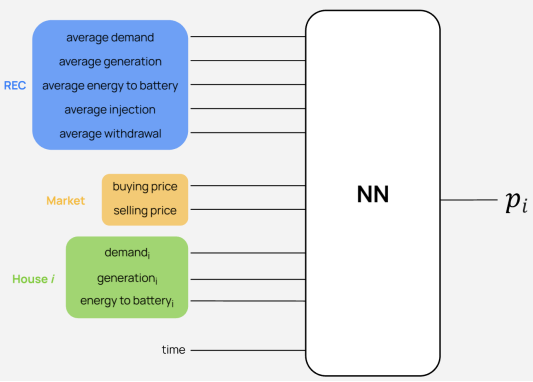
Input: Iterations I, REC episodes $E_{\mathfrak{C}}$, nodes episodes $E_{\mathcal{N}_A}$

Initialize REC parameters ϕ, node parameters $\{\theta_n\}_{n\in\mathcal{N}_A}$
for 1 to I **do**
  for 1 to $E_{\mathfrak{C}}$ **do**
    $\{\hat{\theta}_n\} \leftarrow \{\theta_n\}$
    for $l \in \{1, \dots, L\}$ **do**
      Collect trajectory $\hat{\tau} \sim \left(\{\pi_{\hat{\theta}_n}\}, \pi_{\phi}^{\mathfrak{C}}\right)$
      Simulate update on $\{\hat{\theta}_n\}$: $\hat{\theta}_n \leftarrow \hat{\theta}_n - \hat{\alpha}_n \nabla_{\hat{\theta}_n} \mathcal{L}_n(\hat{\tau}), \quad \forall n \in \mathcal{N}_A$
      Collect new trajectory $\hat{\tau}' \sim \left(\{\pi_{\hat{\theta}_n}\}, \pi_{\phi}^{\mathfrak{C}}\right)$
      Compute $J_{\mathfrak{C}}$: $J_{\mathfrak{C}}(\hat{\tau}') = \frac{1}{\widehat{T}} \sum_{t=0}^{\widehat{T}} R_{\mathfrak{C}}(\hat{s}_t, \hat{\boldsymbol{a}}_t)$
      Update ϕ: $\phi \leftarrow \phi + \alpha_{\mathfrak{C}} \nabla_{\phi} J_{\mathfrak{C}}(\hat{\tau}'))$
    end for
  end for
  for 1 to $E_{\mathcal{N}_A}$ **do**
    Collect trajectory $\tau \sim \left(\{\pi_{\theta_n}\}, \pi_{\phi}^{\mathfrak{C}}\right)$
    Update $\{\theta_n\}$: $\theta_n \leftarrow \theta_n - \alpha_n \nabla_{\theta} \mathcal{L}_n(\tau), \quad \forall n \in \mathcal{N}_A$
  end for
end for

5 Multi-Agent for Renewable Energy Community

Policy Architectures





Multi-Agent for Renewable Energy Community

Fairness Analysis

Algorithm	3 Agents		5 Agents		8 Agents	
	Gini (↓)	**Jain's** (↑)	**Gini** (↓)	**Jain's** (↑)	**Gini** (↓)	**Jain's** (↑)
RULE-BASED	**0.0802 ± 0.0166**	**0.9744 ± 0.0111**	0.1425 ± 0.0553	0.9254 ± 0.0588	**0.1541 ± 0.0394**	**0.9152 ± 0.0367**
PPO	0.3632 ± 0.0190	0.6867 ± 0.0240	0.2954 ± 0.0267	0.7745 ± 0.0305	0.3250 ± 0.0050	0.7510 ± 0.0066
LOLA	0.2020 ± 0.0585	0.8622 ± 0.0725	0.4226 ± 0.0248	0.6114 ± 0.0372	–	–
INAIA-OPT	0.1135 ± 0.0682	0.9456 ± 0.0522	0.4004 ± 0.0455	0.6359 ± 0.0688	0.2111 ± 0.0162	0.8389 ± 0.0218
INAIA-SHARED	0.1369 ± 0.0674	0.9244 ± 0.0474	**0.1247 ± 0.0394**	**0.9472 ± 0.0350**	0.4325 ± 0.0304	0.5817 ± 0.0377

5 Multi-Agent for Renewable Energy Community

Ablation Study

