

# Apprentissage par renforcement pour le contrôle haut niveau de systèmes énergétiques distribués

## Contexte et objectifs

### Contrôle et dimensionnement du micro-réseau

- ▶ Un **micro-réseau** est un réseau électrique géographiquement localisé incluant production d'électricité renouvelable, stockage d'électricité et demande de charge
- ▶ L'Energy Management System (**EMS**) pilote l'électrolyseur et la pile à combustible (PAC) d'un micro-réseau (Figure 1) en temps réel.
- ▶ L'outil développé concerne à la fois la **conception** et le **pilotage** de ces systèmes énergétiques décentralisés.
- ▶ Le travail en cours cherche à étudier la possibilité d'utiliser des **modèles par apprentissage automatique** pour l'intégration des composantes aléatoires telles que la production d'électricité d'origine renouvelable et la consommation.

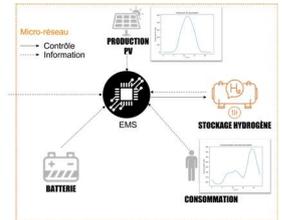


Figure 1: Représentation schématique du micro-réseau constitué de panneaux PV, d'une batterie électrochimique, d'une PAC et d'un électrolyseur.

## Contrôle optimal d'un micro-réseau

### Conception de l'EMS par apprentissage par renforcement

- ▶ La méthode d'apprentissage automatique utilisée pour le contrôle est l'**apprentissage par renforcement** [1] (Figure 2). L'agent apprend à minimiser la quantité d'électricité soutirée du réseau central pour gagner en autonomie et diminuer le coût d'exploitation (Figure 3).

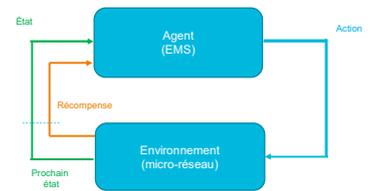


Figure 2: représentation des interactions entre agent (EMS) et environnement (micro-réseau) en apprentissage par renforcement.

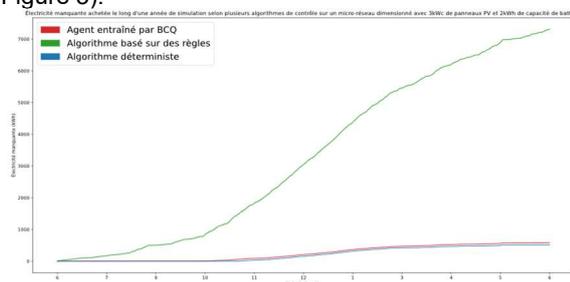


Figure 3: Électricité manquante achetée selon plusieurs algorithmes de contrôle sur un micro-réseau dimensionné avec 3kWc de puissance nominale de panneaux PV et 2kWh de capacité de batterie au long d'une année simulée.

- ▶ Lors de son apprentissage, l'agent choisit des **actions** et reçoit des **récompenses** dont les valeurs dépendent de l'écart entre les **états** de l'environnement et les objectifs attendus.

## Dimensionnement optimal d'un micro-réseau

### Dimensionnement optimal sous contrôle optimal

- ▶ Les variables du dimensionnement sont la puissance PV installée et la capacité de la batterie. Une fois la politique de contrôle apprise, le coût d'un dimensionnement est calculé (Figure 4) en intégrant les coûts d'exploitation (OPEX) et les coûts d'investissement (CAPEX).
- ▶ Le dimensionnement optimal est obtenu avec un **algorithme d'optimisation** (recuit simulé)

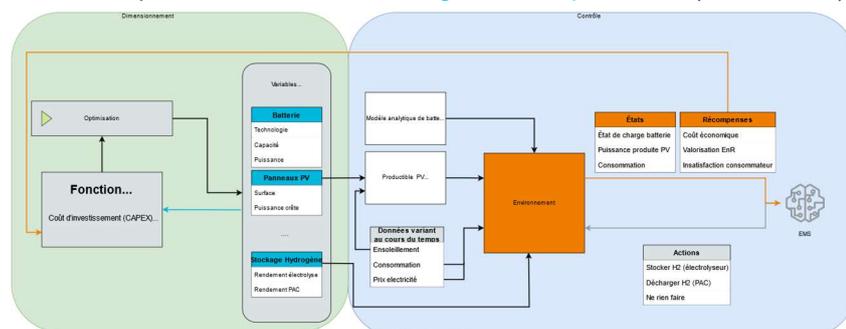


Figure 4: Représentation de la méthodologie adoptée : un algorithme d'optimisation bi-niveau minimise l'OPEX qui est déterminé avec le RL et le CAPEX.

## Conclusion et perspectives

- ▶ La stratégie de contrôle utilisée est plus performante qu'une approche basée sur des règles car elle s'adapte aux **variations** aléatoires de grandeurs externes (production PV, consommation).
- ▶ Un algorithme de transfert de politique s'appuyant sur le **Batch-Constrained Q-learning** [2] est développé afin de gagner du temps de calcul pour le contrôle des micro-réseaux.

## Parties prenantes

## Auteurs

Valentin Père  
Fabien Baillon  
Mathieu Milhé  
Jean-Louis Dirion

## Mots-clés

Apprentissage par renforcement

Contrôle

Dimensionnement

Micro-réseau

Optimisation

## Références

[1] R. S. Sutton, A. G. Barto, 1995. Reinforcement Learning: An Introduction. The MIT Press

[2] S. Fujimoto et al., 2019. Off-Policy Deep Reinforcement Learning without Exploration