POSTPRINT VERSION. The final version is published here: Khalatbarisoltani, A., Boulon, L. et Hu, X. (2023) Integrating Model Predictive Control With Federated Reinforcement Learning for Decentralized Energy Management of Fuel Cell Vehicles. IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2023.3303991

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Integrating Model Predictive Control with Federated Reinforcement Learning for Decentralized Energy Management of Fuel Cell Vehicles

Arash Khalatbarisoltani, *Member, IEEE,* Loïc Boulon, *Senior Member*, *IEEE*, Xiaosong Hu, *Senior Member*, *IEEE*

1

***Abstract*— The optimization-based energy management strategy (EMS) enables expertise to improve the performance of fuel cell vehicles (FCVs). Ongoing efforts are mostly focused on optimizing a centralized EMS using a variety of high-computing technologies without offering appropriate scalability and modularity for the onboard powertrain components. In real-time applications, the time-accomplishment capability of EMSs is crucial; hence, decentralized EMSs with low-cost components and limited processing capability are necessary. Local units handle the computation load on a modular platform. In addition, the decentralized system's plug-and-play functionality minimizes the total cost. This paper presents a decentralized model predictive control (D-MPC) based on the consensus-based alternating direction method of multipliers (C-ADMM) that explicitly considers the coordination of the dynamic reactions of powertrain components and future driving profiles. In addition, a decentralized learning method is proposed to seek the optimal policy for the moving horizon dimensions in the D-MPC using the federated reinforcement learning (FRL) algorithm in order to improve processing time. Due to the deployment of a fully modular system in the proposed learning technique, agents are restricted from sharing their trajectories. Using a highly dynamic module-to-module communication layer in a fully decentralized arrangement, the powertrain components utilize the multi-step method to attain the global optimum. The performance of the proposed framework is evaluated with regards to its precision, convergence speed, and scalability. The results of numerical simulation and implementation demonstrated that the proposed method is superior to the centralized and fixed-horizon MPC approaches.**

Index Terms— **Alternating direction method of multipliers (ADMM), federated reinforcement learning (FRL), proton exchange membrane fuel cell (PEMFC), distributed optimization algorithms, energy management strategy (EMS), fuel cell vehicle (FCV), model predictive control (MPC), multi-agent system.**

Abbreviation

Abbreviation	Description
ADMM	Alternating direction method of multipliers
C-MPC	Centralized MPC
C-ADMM	Consensus-based alternating direction method of multipliers
D-MPC	Decentralized MPC
DOD	Depth of discharge
DCO	Distributed convex optimization
EMS	Energy management strategy
FRL	Federated Reinforcement Learning
FL	Federated learning
FC	Fuel cell
FCS	Fuel cell system
FCV	Fuel cell vehicle
ICE	Internal combustion engine
MDP	Markov decision process
MPC	Model predictive control
MES	Modular energy system
PEMFC	Proton exchange membrane fuel cell
RL	Reinforcement learning
SoC	State of charge
SoH	State of health

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chairs program. The work of X. Hu was supported by the National Key Research and Development Program (No.2021YFE0193800) and the project of basic research funds for central universities (2023CDJXY-021).
(Corresponding authors: Loïc Boulon and Xiaosong Hu.)

A. Khalatbarisoltani and X. Hu are with Department of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044, China (e-mails: arash.khalatbarisoltani@cqu.edu.cn; xiaosonghu@ieee.org). L. Boulon is with the Hydrogen Research Institute, Department of Electrical and Computer Engineering, Université du Québec à Trois-Rivières, QC G8Z 4M3, Canada (loic.boulon@uqtr.ca).

Nomenclature

Symbol	Description
A_n and C	Powertrain and coupling constraints.
$\mathcal{A}$	Action set.
A_f	Front area (m^2).
$a_{n,j}$	Action.
a_t	Vehicle acceleration (m/s^2).
$c_{n,j}$	Total cost.
c_r	Coefficient of rolling friction.
c_d	Coefficient of drag.
C_c	*Polarization capacitor* (F).
$c_{h_{n,t+r}}$	Hydrogen cost.
C_{H_2}	Hydrogen price (\$/Kg).
$c_{d_{n,t}}$	Total degradation costs.
C_{FC_n}	FCS price (\$/kW).
C_B	Battery price (\$/kWh).
c_{SoC_t}	Punishment cost to consider the SoC error.
$c_{B,t+r}$	Battery degradation cost.
$D_{n,t}$	Control signal of the boost converters.
d_n^l	Low load degradation.
d_n^h	High load degradation.
E	Number of local learning updates.
$F_{w,t}$	Wheel force (N).
$F_{n,t}$	Power of FCSs (W).
g	Gravitational constant (m/s^2).
i	Iterations number of D-MPC.
I_B	Battery pack current (A).
I_n	Current of FC_n (A).
j	Iteration number of FRL.
L_n	Smoothing inductor inductance (mH).
$m_{h,t}$	Modulation ratio of the converters.
m_{tot}	Total weight with payload (kg).
m_r	Mass of all rotating parts (kg).
m_{FC}	Mass of the primary power supply (kg).
m_{con}	Weight of the converter (kg).
m_{Bat}	Mass of the FCV's secondary power supply (kg).
m_0	Weight of initial vehicle components (kg).
$m_{h,t}$	Modulation ratio of the converters.
$\mathcal{M}_l$	Imagined environment.
$\mathcal{M}_n$	Environments for $n = 1, \dots, N$
n_B	Total number of cycles during the lifetime of the battery unit.
$n \in N$	Each PEMFC module.
n_{FC}	Cell number of PEMFC.
P_n^{nom}	Nominal power of the modules (W).
$P_{loss,t}$	Auxiliary power of FCSs (W).
P_n^{min} and P_n^{max}	Minimum and maximum values of $P_{n,t}$ (W).
$P_{n,t}$	Power of FC_n (W).
$P_{B,t}$	Power of battery (W).
$P_{L,t}$	Requested power from the propulsion system (W).
P_B^{min} and P_B^{max}	Minimum and maximum limits of $P_{B,t}$ (W).
$P_{w,t}$	Wheel power (W).
$P_{n,t}^{ref}$	Reference power for the modules.
P_{avg}	Average requested power for the maximum horizon.
$\{\mathcal{P}_n\}_{n=1}^N$	Probability of state transitions.
Q_B	Battery capacity (Ah).
r	Length of the moving horizon.
R	Maximum horizon length
$\mathcal{R}$	Reward function
r	
r_n	Smoothing inductor resistance ($m\Omega$).
R_c	Polarization resistance *(Ω)*.
R_s	Series ohmic resistance (Ω).
R_n^{down} and R_n^{up}	Boundaries of the slew rate.
$R_B{}^{down}$ and $R_B{}^{up}$	Slew rate boundaries of $P_{B,t}$.
$\mathcal{S}$	State space.
SoC	State of charge.
SoC^{min}	Minimum SoC.
SoC^{max}	Maximum SoC.
SoC_0	Initial SoC level.
SoH^{min}	Minimum value.
$SoH_{t=0}$	Initial SoH level.
SoC_j	Current SoC
$t_{n,j}$	Computational time of module n
t	Current optimization time step (s).
t_h	Duration that the FC works at the high load condition (s).
t_l	Duration that the FC works at the low load (s).
T	Computational burden
v_t	Vehicle speed (m/s).
V_0	Open-circuit voltage (V).
V_B	Terminal voltage (V).
V_n	Voltage of FC_n (V).
$V_{nom,n}$	Nominal FC_nvoltage drop (V).
$W_{n,t}$	Global power
$\alpha_{n,1}$, $\alpha_{n,2}$, and $\alpha_{n,3}$	Coefficients of hydrogen curve.
$\beta_{n,1}$, $\beta_{n,2}$, $\beta_{n,3}$	Coefficients of degradation curve.
θ	road grad.
ρ_a	Air density (kg/m^3).
ε_o	Start-stop coefficient ($\frac{\mu V}{cycle}$).
n_o	Number of start-stop cycles.
ε_l	Low load coefficient ($\frac{\mu V}{h}$).
ε_t	load transition coefficient ($\frac{\mu V}{h}$).
η_i	Efficiency of the inverter (%).
η_e	Efficiency of the electric motor (%).
η_d	Efficiency of the final gear (%).
η_n	Efficiency of the boost converters (%).
η_h	Average efficiency (%).
ρ	Tuning value.
λ_n	Lagrangian multipliers
φ	Positive coefficient.
γ	Discounting factor
ω_1	Weighting variable of the reward function
Δt	Time step.
ΔV_{cell}	Voltage drops of the FCS

I. Introduction

Due to concerns over fossil fuel usage, transportation is shifting from internal combustion engine (ICE) cars to fuel cell vehicles (FCVs) [1]. Proton exchange membrane fuel cells (PEMFCs) are used in FCVs due to their power density, low-temperature working range, low noise, and high-efficiency [2]. FCVs have a fuel cell system (FCS) and a battery pack. Energy management strategies (EMSs) improve the performance and service life of FCV powertrain components [3, 4]. Since heavy-duty FCVs require high power, a large FCS is needed. Multilayer fuel cell (FC) stacks reduce powertrain safety and reliability. These drawbacks motivate a shift to the modular energy systems (MESs) [5]. [6] provides an overview of modular FCSs with different fluidic and power conditioning topologies.

EMSs have been designed using power point tracking [7], optimization [8, 9], hierarchical [10], hysteresis [11], and droop control [12]. When integrating dynamic component responses, multiple-time steps must be minimized. Standard [13-17], nonlinear [18, 19], hierarchical [20], mixed-integer [21], and multi-mode [22] model predictive control (MPC) approaches are suggested. In an MPC scheme, the moving window size determines how far into the future the optimization scheme evaluates control actions. Centralized MPC (C-MPC) implementation is limited by computation. C-MPC hyperparameters require tuning and lack plug-and-play (modularity) and scalability. In rule-based and optimization-based EMSs, the time-varying behavior of powertrain components and their performance in diverse driving profiles are not considered, resulting in erroneous optimal solutions.

To address these flaws, different reinforcement learning (RL) algorithms [23] are recommended, such as standard [24-26], hierarchical [27, 28], and recursive [29] in FCV. RL has also been suggested for the gaming [30], robotics [31], and healthcare [32]. Poor sample efficiency hinders RL's adoption in real-world applications despite its simulation performance [33, 34]. FC module data can prepare advanced EMSs. Massive amounts of data can cause computational overhead and divergent optimization results in data-driven EMSs. Attackers may access centralized data. These centralized learning-based EMSs need durability, modularity, and computational burden improvements.

Decentralized EMS with cooperative subproblems makes sense. Local subproblems have fewer share variables, constraints, and control variables than centralized EMS. Decentralized EMS handles anomalies and offers plug-and-play. This plug-and-play feature enables manufacturers to include a reconfigurable controller, reducing the cost of installing new components. Decomposing MPC and using multi-agent RL to build a decentralized EMS method optimizes each. Distributed convex optimization (DCO) algorithms are new [35, 36]. The alternating direction method of multipliers (ADMM) [37] is a popular DCO that combines the convergence of multipliers with Lagrangian relaxation decomposability. In a multi-agent system, it's logical to encourage RL agents to share their trajectories to make better decisions collectively than individually. Multi-agent RL has many benefits [38-40]. Parallel processing allows multi-agent RL to learn quickly. Second, local analysis reduces the amount of data sent. This reduces network costs, synchronization, and latency issues. Third, if one agent stops working due to an electrical failure or cyberattack, the others can compensate. This structure makes adding new agents easily, giving it flexibility and scalability.

Numerous obstacles must be considered when implementing RL in real-world situations. For instance, when action and state space are abundant, agents' performance is susceptible to collected samples. Another RL issue is overcoming the divide between simulation and the real world. Moreover, multiple RL algorithms need more sample efficiency. Through the exchange of information between agents, learning can be significantly accelerated. However, one of the most important concerns is that specific duties must secure agent privacy. Federated learning (FL) preserves data privacy by training models using decentralized, multi-agent data [41]. Not only can FL complete information exchange without compromising privacy, but it can also adapt various agents to their unique environments [42]. Therefore, switching to FL may be the more reasonable option. For instance, BrainTorrent [43] is a serverless peer-to-peer method where all agents interact with each other. Federated Reinforcement Learning (FRL) [44, 45] tries to federatively construct a better policy from several agents without forcing them to exchange raw trajectories. Unlike FL, FRL needs every agent to interact with the environment while learning. FRL is attractive for resolving RL's real-world inefficiencies. Traditional RL assumes a fixed state transition. In real-world applications, modules with various behaviors may be seen as environments with separate state transitions, meaning they may respond differently to the same operation [46]. Moreover, real data is usually inadequate for a powertrain application to design a reliable policy. Thus, module policy training works in heterogeneous FRL environments. The unsharring architecture in FRL for the MES application can give a fully decentralized system with less communication and more modularity.

To achieve optimal results for the FCV control problem, it is essential to use a long fixed-horizon optimization window. However, this may lead to longer computational times and accurate future-driving predictions may not always be possible. Conversely, using a short fixed-horizon optimization window may result in less efficient outcomes but requires fewer computational resources. This paper contributes to a decentralized MPC (D-MPC) EMS framework and investigates how the FRL algorithm helps decide the prediction moving horizon in an FCV. Our article presents a learning method that selects the best optimization window by minimizing a reward function, considering both computational time and total cost.

In contrast to [47] and [48], the previously developed single-step DCO-based techniques are extended into a multi-step decentralized approach with parallel moving horizons. The proposed strategy reduces total system costs by identifying preventive control responses for future driving profiles. Integrating the receding prediction horizon into the decomposition technique improves initial points, optimization convergence time, modularity, and fault robustness. Since the prediction horizon length affects D-MPC's performance, a learning technique integrating FRL with D-MPC is proposed. The FRL structure learns a shared horizon control policy without a central unit. FC modules trained locally to find the control policy on their data for a few epochs, and then they shared their policies to make an aggregated and fine-tuned model. To the best of our knowledge, no study has been performed to build a D-MPC approach with fully decentralized learning capabilities for a modular FCV.

The rest of this paper proceeds as follows: Section II describes FCV powertrain modeling, and Section III formulates the multi-step EMS optimization problem. Section IV derives the suggested D-MPC framework using the consensus-based alternating direction method of multipliers (C-ADMM). Section V provides an FRL framework for learning the optimal length of the D-MPC's receding prediction horizon. Sections VI and VII present comprehensive numerical simulations and experimental results. Section VIII summarises the proposed decentralized approach and its prospects.

II. CONFIGURATION AND MODELING OF THE MODULAR FCV PROPULSION SYSTEM

As depicted in Fig.1, a modular small-scale test bench was built at the Hydrogen Research Institute of the University of Quebec in Trois-Rivières (UQTR) to investigate transforming the centralized power sharing problem into D-MPC and integrating it with FRL. The test bench was constructed on the foundation of a low-speed electric vehicle [49]. This setup includes two FC modules, a battery pack, a programmable DC electronic load (8514 BK Precision), and a programmable DC power supply. Each FC module has a 500-W open-cathode PEMFC (Horizon, model: H-500), a smoothing inductor, and an adjustable boost DC-DC converter (Zahn ElectronicsTM, model: DC5036-SU). Six 12-V 18-Ah batteries power the DC bus. In a semi-active design, the battery pack is directly connected to the DC bus, and the modules are coupled using converters. Each FC module has a controller in an NI CompactRIO (NI9022) by LabVIEW program. CompactRIO consists of an FPGA, a real-time controller, a programmable I/O, and an Ethernet chassis. Microprocessors are used to control real-time control units, and FPGAs measure and process data efficiently. Each FC module's control signal is set at 10 Hz. The programable load and power supply generate the desired FCV power.

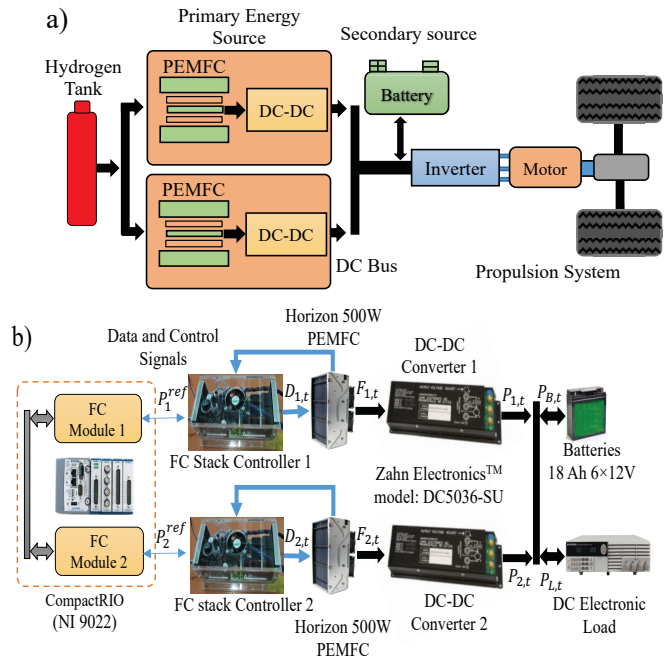

Fig. 1. (a) the schematic of the powertrain system involving two PEMFC modules and one battery unit, (b) the established small-scale modular test bench.

A. Longitudinal Fuel Cell Vehicle Dynamics Formulation

The modular FCV is a rigid body with a constant mass. Throughout the article, n refers to each PEMFC module, t is the current optimization time step, i denotes the iteration number of D-MPC and j is the iteration number of FRL. Using Newton's equation of motion, the required wheel force, $F_{w,t}$, can be expressed as

$$F_{w,t} = (m_{tot} + m_r)a_t + c_r m_{tot}\, g\, cost\, \theta + m_{tot}\, g\, sin\, \theta + 0.5\, \rho_a c_d\, A_f\, {v_t}^2, \quad (1)$$

where m_{tot} is the total weight with payload, m_r represents the mass of all rotating parts, c_r, g, θ, ρ_a, c_d, and A_f represent the coefficient of rolling friction, the gravitational constant, road grad, air density, coefficient of drag, and front area, respectively, v_t is the car speed and a_t is the vehicle acceleration. The FCV is on a horizontal surface, with constant friction and drag resistance. The overall modular FCV mass is calculated by

$$m_{tot} = 2\,(m_{FC} + m_{con}) + m_{Bat} + m_0, \quad (2)$$

where m_{FC} is the mass of the primary power supply, m_{con} is the weight of the converter, m_{Bat} and m_0 represent, respectively, the mass of the FCV's secondary power supply

and the remaining components. Table I presents complete information of the electric vehicle parameters [49].

Table I
Technical details of the low-speed electric vehicle [49]

$c_r = 0.015$	$m_{tot} = 896$ kg	$A_f = 4\ m^2$
$\theta = 0°$	$\rho_a = 1.2\ kg/m^3$	$c_d = 0.42$

The demanded power $P_{L,t} = F_{w,t} v_t$ can be calculated using modular FCV longitudinal dynamics, where $P_{L,t}$ is the required wheel power, $F_{w,t}$ is the required wheel force, and v_t is the preferred modular FCV speed. A standard driving cycle determines car speed. An electric machine powers the wheels via a differential. The electric machine can act as a generator to recover regenerative braking energy. The PEMFC modules and battery pack power the vehicle's electric machine for the desired driving cycle. A DC-to-AC inverter is used between the converters and the electric machine. Modeling the modular FCV powertrain is formulated by considering the efficiency of the electric motor, boost converter, inverter, and final gear. When $P_{L,t}$ is negative, $\eta_i\ \eta_e \eta_d$ multiplies $P_{L,t}$ and divides when $P_{w,t}$ is positive. η_i, η_e, and η_d denote the efficiency of the inverter, electric motor, and final gear, respectively. Figure 2 displays an efficiency map for a 24-kW electric motor provided by ADVISOR software, scaled in this study [50].

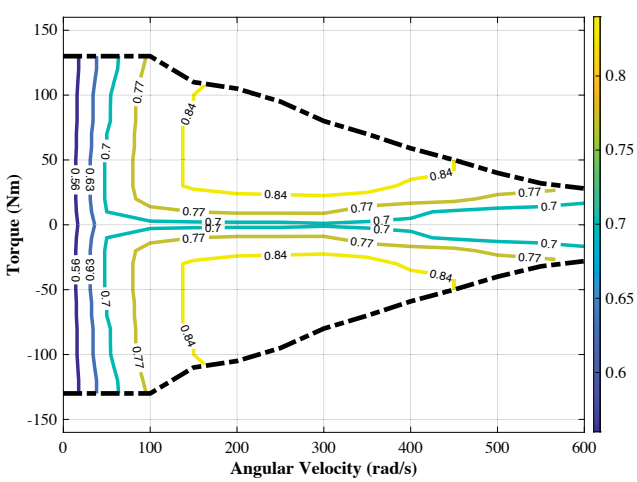

Fig. 2. Electric motor efficiency map.

The power balance equation of the modules and the battery unit on the DC bus at each step of the optimization window is formulated in (3)

$$\sum_n P_{n,t} + P_{B,t} = P_{L,t}, \tag{3.a}$$

$$P_{n,t} = \eta_n \left(F_{n,t} D_{n,t} - P_{loss,t}\right), \tag{3.b}$$

where $P_{n,t}$ denotes the power by each one of the modules FC_n, $P_{B,t}$ denotes the power provided by the battery unit, $P_{L,t}$ is the requested power from the propulsion system, $F_{n,t}$ indicates the generated power of each of the 500-W FCSs, η_n and $D_{n,t}$ are the efficiency and the control signal of the boost converters, respectively, $P_{loss,t}$ denotes the consumed power by the auxiliary of FCSs.

B. *Fuel Cell System Modeling*

PEMFC stands for polymer electrolyte membrane fuel cell and is an electrochemical device that directly converts chemical energy into electrical energy.

1) *Polarization curves*

The 500-W FCSs, FC_n, are modelled as voltage sources. Table II lists PEMFC stack technical specifications. Figure 3 shows the polarization curves and generated powers versus requested currents. The CompactRIO transmits each Horizon 500-W PEMFCS's voltage, current, and hydrogen consumption to the PC via Ethernet to extract these curves. To make the PEMFC profiles, the programable load is used.

Table II
Technical specification of the open-cathode PEMFC stacks (H-500)

Number of cells= 24	Hydrogen pressure= 0.45-0.55 bar
Rated power=500 W	External temperature= 5-30 °C
Rated performance= 14.4 V, 35 A	Max stack temperature= 65 °C
Output voltage range= 12 V-24 V	Humidification= Self-humidified
Reactants= Hydrogen and Air	Flow rate at max power= 6.5 L/min

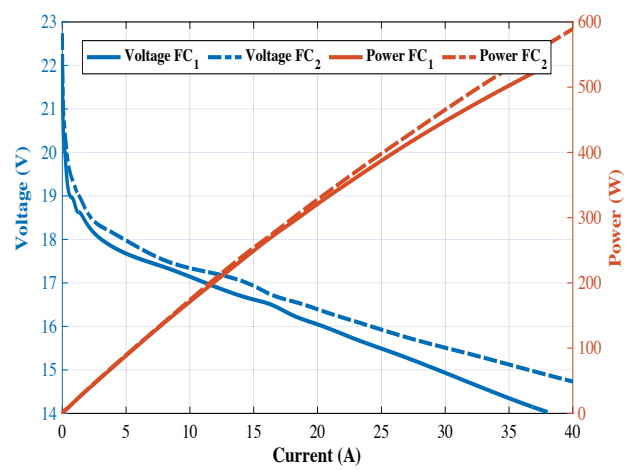

Fig. 3. The characteristic curves of PEMFCs: the polarization curves (blue) and the power curves (red).

2) *Hydrogen Consumption Modeling and Constraints*

The hydrogen consumption curves, $h_{n,t}$, are approximated by means of a second-order polynomial technique, $h_{n,t} = \alpha_{n,1} {P_{n,t}}^2 + \alpha_{n,2} P_{n,t} + \alpha_{n,3}$. The coefficients of the FC_1 hydrogen curve are $\alpha_{11} = 0.1862$ e − 08, $\alpha_{12} = 0.1292$ e-04, and $\alpha_{13} = 0.7613$ e-05 and the coefficients of FC_2 hydrogen

curve are $\alpha_{21} = 0.1128$ e-08, $\alpha_{22} = 0.1506$ e-04, and $\alpha_{23} = 0.6249$ e-04. A comparison between the measured hydrogen consumption flows and the convex models is presented in Fig. 4, where Data FC_1 and Data FC_2 represent the hydrogen consumption measurements of FC_1 and FC_2, respectively. And Model FC_1 and Model FC_2 are convex models that were used to approximate their behaviours. The graphs shown in Fig.3 and Fig.4 illustrate that there are differences in the polarisation curves and hydrogen consumption between the two FCs, and this is because FC_1 is older than FC_2.

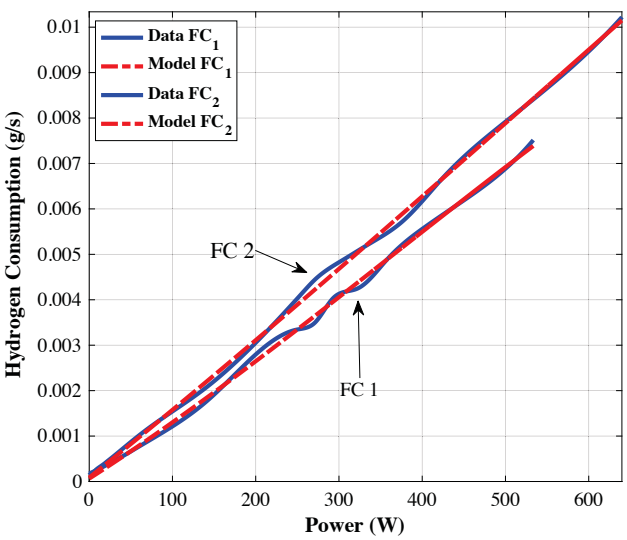

Fig. 4. The measured hydrogen flow curves and the convex approximation models.

Equation (4) demonstrates the upper, lower, and slew rate limits on instantaneous powers $P_{n,t}$

$$P_n^{min} \le P_{n,t} \le P_n^{max}, \tag{4.a}$$

$$R_n^{down}\Delta t \le P_{n,t} - P_{n,t-1} \le R_n^{up}\Delta t, \tag{4.b}$$

where P_n^{min} and P_n^{max} are the minimum and maximum values of $P_{n,t}$, R_n^{down} and R_n^{up} are boundaries of the slew rate, and Δt indicates the time step.

3) Degradation mechanism Modeling

Throughout typical PEMFC operation, degradation occurs. The maximum power of an FCS drops by 10% during its lifespan, and this value is not attained with the same current at the beginning and end of its life [51]. Several physical parameters, including, temperature, pressure, and membrane water content, affect the process of PEMFC degradation. Due to the fact that each PEMFC component degrades differently and has an effect on other components, it is difficult to quantify component-level degradation [52]. Consequently, the stack level is typically considered when predicting PEMFC degradation [52]. Membrane electrode assembly degradation, which is mostly attributable to the degradation of the catalyst layer [53-55], proton exchange membrane [56], and gas diffusion layer [57, 58], have been identified as the most important factor affecting performance and longevity. The degradation of a PEMFC is highly related to the power demand and operating conditions, which are categorized into four groups, including start-stop, extra high, extra low, and transition [59-62]. Degradation of the PEMFC can result in a voltage decrease, ΔV_{cell}, and the level of PEMFC degradation is defined as the summation of the entire voltage degradation rate (μV) of a cell under each one of the four load profile conditions. The degradation voltage droop rate for start-stop cycle is calculated by $\varepsilon_o\, n_o$, where ε_o represents the start-stop coefficient ($13.79\frac{\mu V}{cycle}$) and n_o denotes the number of start-stop cycles. The rate of voltage droop for the extra-high load is calculated by $\varepsilon_h \frac{t_h}{3600}$, where ε_h represents the extra high load coefficient ($10.00\ \frac{\mu V}{h}$) and t_h denotes the duration that the PEMFC works at the extra-high load condition. The voltage droop rate for the extra low load is calculated by $\varepsilon_l \frac{t_l}{3600}$, where ε_l represents the extra low coefficient ($8.662\,\frac{\mu V}{h}$) and t_l denotes the duration that the PEMFC works at the extra low load. The PEMFC degradation voltage droop due to the transition load is calculated by $\varepsilon_t\, n_t$, where ε_t denotes the transition load coefficient ($0.04185\,\frac{\mu V}{h}$) and ε_t is the amount of the load change [63]. In this work, since the main influence on the voltage degradation of PEMFC is due to the start-stop condition, the PEMFC modules are always operational, and the related degradation rate is disregarded. The extra low load and extra high-power constraints are defined as 20% and 80% of the maximum power of the PEMFC modules, respectively. To avoid PEMFC degradation because of inadequate humidification caused by gas starvation or water, PEMFC output power variations are restricted to 10% of maximum output when rising and 30% of maximum output when decreasing. The extra low load degradation, d_n^l, and the extra high load degradation, d_n^h, terms for a stack can be calculated by

$$d_n^l = \frac{n_{fc}\varepsilon_l\mu_{l,n}}{3600\ 0.1V_{nom,n}}, d_n^h = \frac{n_{fc}\varepsilon_h\mu_{h,n}}{3600\ 0.1V_{nom,n}}, \tag{5.a}$$

where n_{fc} represents cell numbers, $\mu_{l,n}$ and $\mu_{h,n}$ are equal to

$$\mu_{l,n} = \begin{cases} 1 & \text{if } P_n^{min} \le P_{n,t} \le 0.2P_n^{nom} \\ 0 & \text{otherwise.} \end{cases}, \tag{5.c}$$

$$\mu_{h,n} = \begin{cases} 1 & \text{if } 0.8P_n^{nom} \le P_{n,t} \le P_n^{max} \\ 0 & \text{otherwise.} \end{cases}, \tag{5.d}$$

where $V_{nom,n}$ is the nominal FC_nvoltage drop and P_n^{nom} is the nominal power of the modules. Usually, it is 80 % of the maximum power of the FCS. Due to the nonconvex characteristics of the extra low-power and extra high-power degradation rates, these terms are approximated with the convex second-order polynomial approach, $d_{n,t} = \beta_{n,1}{P_{n,t}}^2 + \beta_{n,2}P_{n,t} + \beta_{n,3}$. The coefficients of FC_1 are $\beta_{11} = 0.3861$ e − 09 , $\beta_{12} =$ -0.1942 e-06, and $\beta_{13} =$ 0.2217 e-04 and the coefficients of FC_2 are $\beta_{21} =$ 0.2415 e-09, $\beta_{22} =$ -0.1480 e-06, and $\beta_{23} =$ 0.2058 e-04.

C. Battery modeling and constraints

The battery pack primarily supplies peak power and stores energy in the event of regenerative braking.

1) Standard Modeling and Constraints

The first-order RC model of the battery pack is formulated by

$$I_{B,t} = \frac{V_{0,t} - R_s I_{B,t} - V_{B,t}}{R_c} + C_c \frac{d}{dt}\left(V_{0,t} - R_s I_{B,t} - V_{B,t}\right), \tag{6}$$

where I_B is the battery pack current, V_0 is the open-circuit voltage, R_s is the series ohmic resistance, V_B is the terminal voltage, R_c denotes the polarization resistance, and C_c is the polarization capacitor. A liner approximation of the battery model was used for the optimisation purpose. Equation (7) presents the state of charge (SoC) calculation formula along with the constraints on the battery SoC level

$$SoC_{t+1} = SoC_t - \frac{P_{B,t}\Delta t}{Q_B V_{B,t} 3600}, \tag{7.a}$$

$$SoC^{min} \le SoC_t \le SoC^{max}, \tag{7.b}$$

where SoC^{min} and SoC^{max} denote the minimum and maximum limits of SoC, respectively, the initial SoC level $SoC_{t=0}$ is SoC_0 , and Q_B represent the battery capacity. Equation (8) imposes power and slew rate limits for the battery unit $P_{B,t}$

$$P_B^{min} \le P_{B,t} \le P_B^{max}, \tag{8.a}$$

$${R_B}^{down}\Delta t < P_{B,t} - P_{B,t-1} \le {R_B}^{up}\Delta t, \tag{8.b}$$

where P_B^{min} and P_B^{max} are the minimum and maximum limits of $P_{B,t}$, respectively, ${R_B}^{down}$ and ${R_B}^{up}$ are the slew rate boundaries of $P_{B,t}$, and Δt indicates the time step.

2) Degradation Modeling

The battery lifetime is affected by the depth of discharge (DOD) and is defined as an initial capacity drop (reaching 80% of the initial battery capacity). The state of health (SoH) is calculated by

$$SoH_{t+1} = SoH_t - \frac{|P_{B,t}|\Delta t}{2 n_B Q_B V_{B,t} 3600}, \tag{9.a}$$

$$SoH^{min} \le SoH_t, \tag{9.b}$$

where SoH^{min} is the minimum value and $SoH_{t=0}$ is the initial SoH level, n_B denotes the total number of cycles during the whole lifetime of the battery unit, and Δt indicates the time step. The parameters of the battery unit obtained from experimental tests are listed in Table III, where V_{BP} is the open-circuit voltage of the battery pack.

Table III
The approximated battery unit parameters.

$V_0 = 12.21\ V$	$R_s = 0.014\ \Omega$	$V_{BP} = 73.26$	$R_c = 0.017\ \Omega$
Cc =1792 F	$Q_B = 18.2\ Ah$	$SoC^{min} = 0.65$	$SoC^{max} = 0.75$

D. Boost converter modeling and characteristics

The two converters are modeled as follows:

$$L_n \frac{d}{dt} I_{n,t} = V_{n,t} - V_{h_{n,t}} - r_n I_{n,t}, \tag{10.a}$$

$$V_{h,t} = m_{h,t} V_{B,t}, I_{h_{n,t}} = m_{h_{n,t}} I_{n,t} \eta_h, \tag{10.b}$$

where I_n and V_n are the current and voltage of FC_n , respectively, $L_n = 1.1\ mH$ presents the smoothing inductor inductance, $r_n = 23.9\ m\Omega$ is the smoothing inductor resistance, $\eta_h = 96.21\%$ is the average efficiency, and $m_{h,t}$ is the modulation ratio of the converters.

E. Driving Cycle Specifications

Five driving cycles, including one real diving cycle [64], and four standard driving cycles [65] are used to compare and evaluate the proposed EMS. Figure 5 depicts the driving cycles, including the probability distribution of their speed profiles. Table IV displays detailed information for the four-standard light-duty cycles. Table IV provides specific information regarding the driving cycles.

III. THE GENERAL MATHEMATICAL FORMULA OF THE PRIMARY POWER SHARE OPTIMIZATION PROBLEM

In this section, the mathematical equation for the primary centralized EMS optimization problem for the multi-stack FCV is presented. Since the goal of this study isn't to predict future driving speeds, the proposed technique assumes that the power profile is already known. Figure 6 provides a general

description of combining the D-MPC approach with the FRL algorithm. The cost function takes into account several factors, including the penalty for battery SoC, the costs of two FC modules' hydrogen and degradation, and the cost of battery degradation. In this form, all the expenses are considered with the same gain. This way of describing the cost function is widespread in the literature. Since the optimisation used is an MPC framework, r is applied to the whole set.

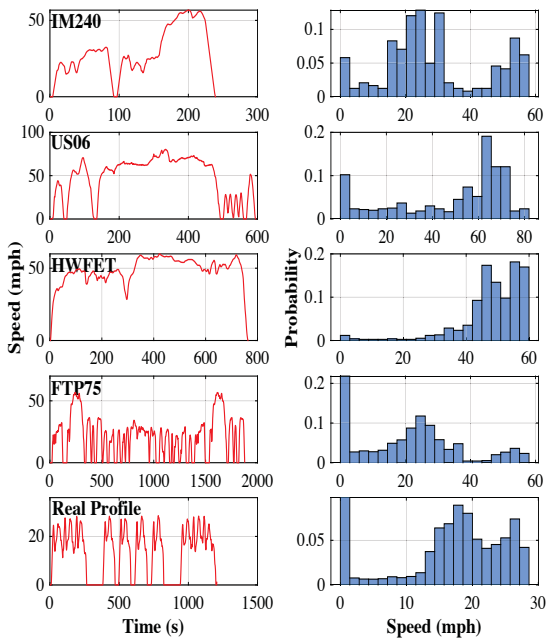

Fig. 5. The driving cycle speed profiles and their distributions [64, 65].

Table IV
The specifications related to the utilized driving cycles [65].

Driving cycle	Time (min)	Distance (mile)	Maximum Speed (mph)	Average Speed (mph)	Average Driving Speed (mph)	Stops
Real Profile	20.10	4.57	28.5	13.7	14.8	6
IM-240	4	1.96	56.7	29.3	30.8	2
US06	9.93	8.01	80.3	48.3	51.9	5
HWFET	12.75	10.26	59.9	48.2	48.6	1
FTP75	31.28	11.04	56.7	21.2	26.2	22

The centralized multi-objective problem can be formulated in the following sequence.

$$\underset{P_{n,t}}{argmin}\ c_{SoC_{t+R}} + \sum_{r=0}^{R-1}\sum_{n=1}^{N} c_{n,t+r} + c_{B,t+r} \tag{11.a}$$

$$s.t\ \sum_{n=1}^{N} A_n P_{n,t+r} = C,\ r = 0, \cdots, R-1$$

(3), (4),(6)-(9)

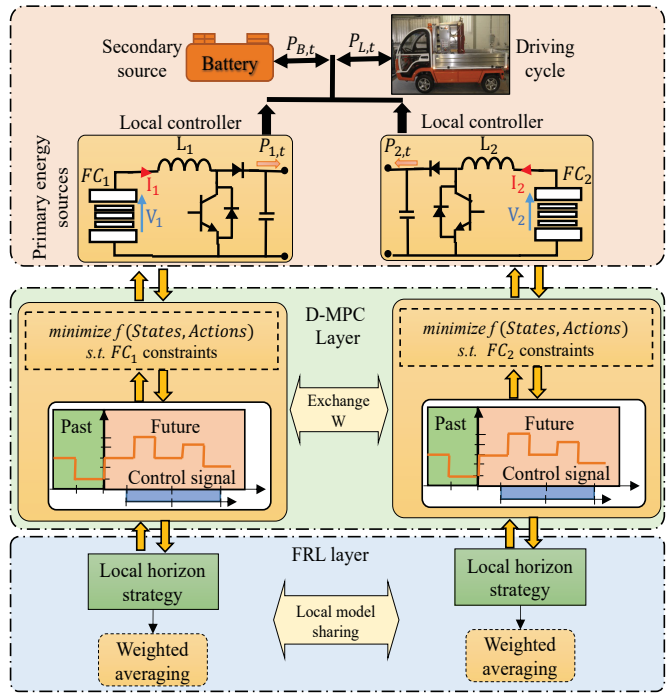

Fig. 6. The comprehensive overview of the fully decentralized design achieved by integrating the D-MPC framework with the FRL strategy in the modular powertrain system.

where $n \in N$ refers to each module, t is the current optimization step time, $r \in R$ denotes the length of the moving optimization horizon, c_{SoC} is the SoC penalty cost, $c_{n,t+r} = c_{h_{n,t+r}} + c_{d_{n,t+r}}$ denotes the sum of hydrogen and degradation costs of FC_n, $c_{B,t+r}$ is the battery unit degradation cost, $P_{n,t}$ and $P_{B,t}$ stand for the power of FC_n and the battery pack. A_n and c apply the powertrain and the coupling constraints. The hydrogen cost $c_{h_{n,t+r}}$ is computed by $h_{n,t+r} C_{H_2} \Delta t$, where $h_{n,t+r} = \alpha_{n,1} {P_{n,t+r}}^2 + \alpha_{n,2} P_{n,t+r} + \alpha_{n,3}$ is a convex function, $C_{H_2} = 3.92$ \$/Kg is the hydrogen price [66], and Δt indicates the time step. $c_{d_{n,t}}$ incorporates both extra low-power and extra high-power degradation costs, calculated by $d_{n,t+r} C_{FC_n} \Delta t$, where $d_{n,t+r} = \beta_{n,1} {P_{n,t+r}}^2 + \beta_{n,2} P_{n,t+r} + \beta_{n,3}$ is a convex function and $C_{FC_n} = 93$ \$/kW is the FCS price [67]. As stated previously, the start-stop and transitional degradation costs are disregarded in this analysis by assuming that all modules are constantly active and restricting the module power dynamic. The battery degradation cost, $c_{B,t+r}$, is determined by

$$c_{B,t+r} = C_B (SoH_{B,t+r} - SoH_{B,0}), \tag{11.b}$$

where $C_B = 189$ \$/kWh is the battery price [68], c_{SoC_t} is a punishment item to consider the SoC level error, which is defined by

$$c_{SoC_{t+R}} = \varphi (SoC_{t+R} - SoC_0)^2, \tag{11.c}$$

where SoC_0 is the initial SoC, and φ is a significant positive coefficient. The constrained finite-time optimization problem (11) is solved at each sampling time t to establish a strategy $P_{n,\{t+1,\dots,t+R\}}$ for the fixed time horizon R. The first optimal input is applied to the powertrain system. At the next optimization step $t+1$, the optimization problem is solved again over the shifted horizon based on new state measurements [69, 70].

A. Transforming Centralised Model Predictive Control into Decentralised Model Predictive Control Using C-ADMM

This section explains how to reformulate C-MPC (11) using C-ADMM. Each subproblem is assigned to an FC module control unit. Since the modules are coupled through (3) on the DC bus, the general problem (11) is not inherently decomposable. To tackle this issue, as demonstrated in Fig. 7, $P^i_{n,\{t+1,\dots,t+R\}}$ is duplicated into its neighboring module value as a virtual power and coupled with a global power vector, $w^i_{n,\{t+1,\dots,t+R\}}$, where i denotes the number of optimization iterations. Equation (12) is added to guarantee that the duplicated variables for FC_1 are equal and the modified optimization problem converges to the same optimal result [71]. Similarly, these equations can be formulated for FC_2.

$$P^i_{1,\{t+1,\dots,t+R\}} - w^i_{1,\{t+1,\dots,t+R\}} = 0 \tag{12.a}$$

$$P^i_{12,\{t+1,\dots,t+R\}} - w^i_{1,\{t+1,\dots,t+R\}} = 0, \tag{12.b}$$

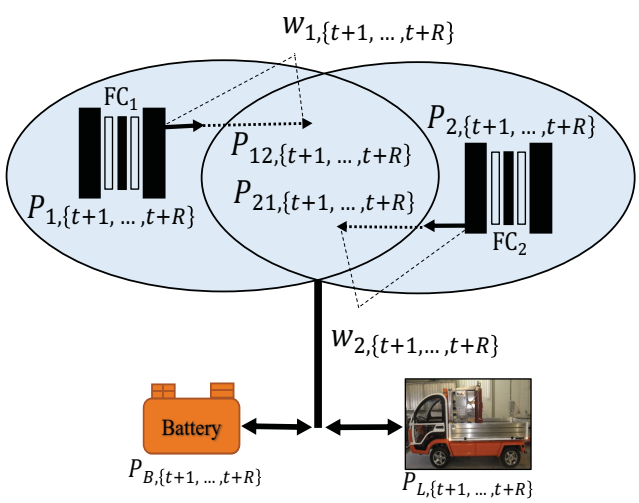

Fig. 7. Visual representation of the employed decomposition technique.

To improve the convergence performance and ease the communication burden, the number of the optimization variables is reduced by assuming that the virtual variables are equal to the previous global variables. For instance, $P^{i+1}_{12,\{t+1,\dots,t+R\}}$ is equal to $w^i_{1,\{t+1,\dots,t+R\}}$. In this way, the main EMS problem can be transformed into two decentralized subproblems functions of $P^i_{n,\{t+1,\dots,t+R\}}$ and $w^i_{n,\{t+1,\dots,t+R\}}$.

During the decentralized optimization, the two modules exchange candidate output powers through module-to-module communication. This iterative process continues until local control units agree on the stopping condition. The optimization procedure is then repeated and shifted, as seen in Fig. 8.

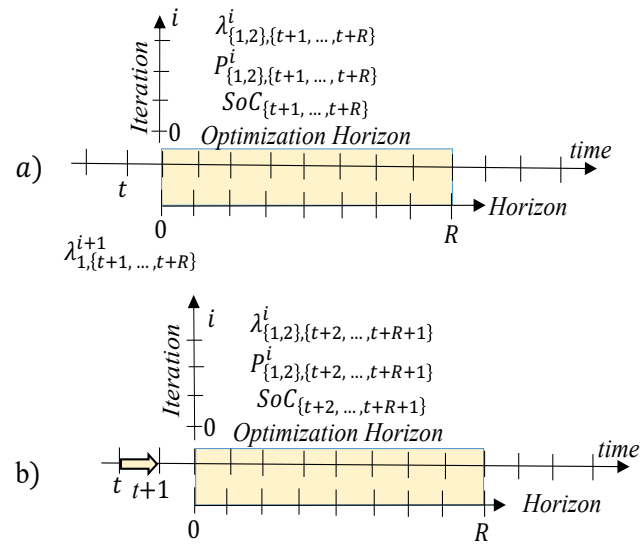

Fig. 8. The sequence operation of the D-MPC framework.

The decentralized optimization process of (11) and (12) includes a three-step procedure, where ρ is a positive tuning value, i denotes the number of iterations, and λ_n are Lagrangian multipliers [71]. The equations related to FC_1 are given in (13). The equations for FC_2 can be expressed in the same way.

$$\underset{P^{i+1}_{1,\{t+1,\dots,t+R\}}}{argmin} \; c_{SoC_{t+R}} + \sum_{r=0}^{R-1}\sum_{n=1}^{N} c_{n,t+r} + c_{B,t+r} \; \lambda^i_{1,\{t+1,\dots,t+R\}} P_{1,\{t+1,\dots,t+R\}} + \frac{\rho}{2}\left(P_{1,\{t+1,\dots,t+R\}} - w^i_{1,\{t+1,\dots,t+R\}}\right)^2\Big\}, \tag{13.a}$$

$$w^{i+1}_{1,\{t+1,\dots,t+R\}} = \frac{P^{i+1}_{1,\{t+1,\dots,t+R\}} + w^i_{1,\{t+1,\dots,t+R\}}}{2}, \tag{13.b}$$

$$\lambda^{i+1}_{1,\{t+1,\dots,t+R\}} = \lambda^i_{1,\{t+1,\dots,t+R\}} + \rho\left(P^{i+1}_{1,\{t+1,\dots,t+R\}} - w^{i+1}_{1,\{t+1,\dots,t+R\}}\right), \tag{13.c}$$

The detailed procedure of D-MPC is presented in Fig. 9. First, a cold-start initialization of $\lambda_{n,\{t+1,\dots,t+R\}}$, $P_{n,\{t+1,\dots,t+R\}}$, and $W_{n,\{t+1,\dots,t+R\}}$ are required for each module n. The local D-MPC problem are solved to determine $P_{n,\{t+1,\dots,t+R\}}$. Then, $W_{n,\{t+1,\dots,t+R\}}$ are calculated and sent to the neighbor module in parallel. After that, $\lambda_{n,\{t+1,\dots,t+R\}}$ updates. If $\left\|\lambda^{i+1}_n - \lambda^i_n\right\|^2_2 \le \mu_1$ and $\rho\left\|P^{i+1}_{n,t} - P^i_{n,t}\right\|^2_2 \le \mu_2$, where μ_1 and μ_2 are the limiting

values, as the stopping criteria are fulfilled, the optimization processes stop and $P_{n,t}^{i+1}$ send to the converters as $P_{n,t}^{ref}$. After that, the optimization windows shift for one time step. If not, the optimization process goes back to Step 3.

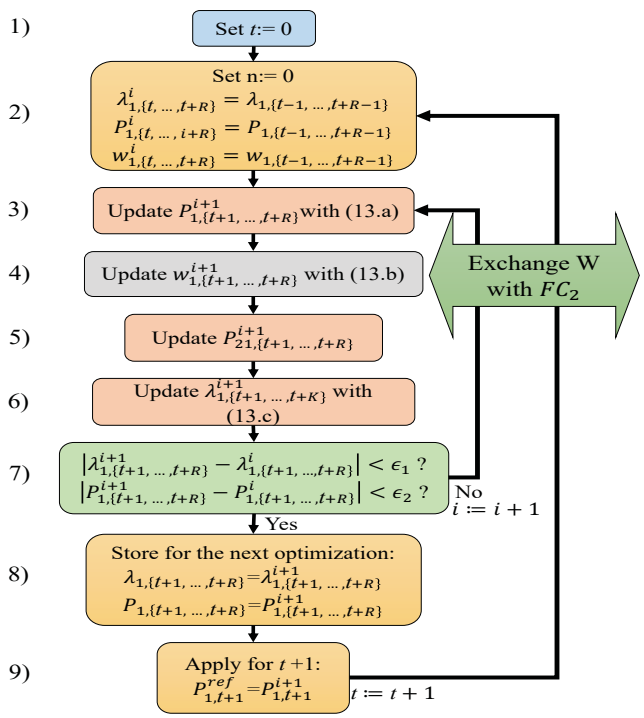

Fig. 9. The step-by-step flowchart of the D-MPC algorithm for the FC_1.

IV. General description of the hyper-parameters tuning algorithm based on Federated reinforcement learning algorithm

In the FRL algorithm, agents are part of an FRL model that can enhance their behaviour through interaction with an external environment. The agents follow policies to perform a series of actions in the environment to achieve a particular objective. The state reflects the agents' observations of the environment. Agents evaluate the environment, make decisions, and perform specific environmental actions. The environment provides the agents with their current state and the rewards for their previous actions. The rewards represent the agents' evaluations of their actions. Unlike RL algorithms, FRL agents do not share their actions with others but instead share their models after internal learning processes [42].

The length of the moving window significantly impacts the optimization accuracy and speed. Enabling agents to share their models, the model-free FRL method QAvg [46], which is a federated variant of Q learning [72], is used to develop a uniformly great strategy for determining the optimization horizon length. Through repeated local updates and iterative global aggregation, this tabular FRL method obtains a value function. It is prohibited for the n agents to share their trajectories due to the architecture of a fully modular system.

Assume that n agents communicate with environments that are distinct from one another. These environments are described using Markov decision processes (MDPs), $\mathcal{M}_n = \langle \mathcal{S}, \mathcal{A}, \mathcal{R}, \gamma, \mathcal{P}_n \rangle$, for $n = 1, \dots, N$ where $\mathcal{S}$ denotes the set of states $s \in \mathcal{S} = \{P_{avg,j}, SoC_j\}$, P_{avg} and SoC are the average requested power for the maximum horizon length and the current SoC value, respectively, $\mathcal{A}$ represents the action set $a_{n,j} \in \mathcal{A} = \{1, \dots, R\}$, where R is the maximum prediction horizon length, $\vartheta_{n,j} \in \mathcal{R}$, $\vartheta_{n,j} = -\omega_1 t_{n,j} - (1-\omega_1) c_{n,j}$ is the normalized reward value obtained when an action $a_{n,j} \in \mathcal{A}$ is taken, ω_1 is the weighting variable, $t_{n,j}$ denotes the D-MPC computational time of module n and $c_{n,j}$ is the sum of the cost of module n and the battery unit, $\gamma \in [0,1]$ presents a discounting factor, $\{\mathcal{P}_n\}_{n=1}^{N}$ are the probabilities of transitioning into $s_{j+1} \in \mathcal{S}$ at step $j+1$ when the D-MPC agents take action $a_j \in \mathcal{A}$, in the state $s_j \in \mathcal{S}$ at step j. The environments share the same state space $\mathcal{S}$, action space $\mathcal{A}$, and reward function $\mathcal{R}$, but their state transitions are different. QAvg acquires knowledge of an $|\mathcal{S}| \times |\mathcal{A}|$ table by oscillating between local updates and global aggregations [46].

In this work, the action space chooses different optimisation horizons for the MPC approach. We are using the state transition function to determine how likely agents are to move from one state to another. This function is considered deterministic, meaning that the agents will always move to the same next state if they are in the same current state and taking the same control action. The process of decentralized learning based on FRL is indicated in Algorithm 1.

Algorithm 1: FRL-based updating procedure for optimum horizon allocation for decentralized controllers.

1. Initialize local Q_n, aggregated $\bar{Q}_n$, iteration number, learning parameters, state and action lists, and probability of actions of each agent

%% Learning an optimal horizon policy for each D-MPC's agent

2. for n= $1, \dots, N$ training data in parallel for N agents **do**

3. **for** local learning episode E= 1, *Max* local training episode **do**

4. **for** j=1, Maximum value for the load profile **do**

5. ⊳ Select an action $a_{n,j}$ with $\mathcal{P}_n$ for each state $s_{n,j}$

6. ⊳ Compute the chosen action and observe $R_{n,j}$

7. ⊳ update Q_n by

$$Q_{n,j+1}(s_{n,j}, a_{n,j}) \leftarrow (1-\eta_j) \times Q_{n,j}(s_{n,j}, a_{n,j}) + \eta_j \times \left[\vartheta(s_{n,j}, a_{n,j}) + \gamma \sum_{s'_{n,j}} \mathcal{P}_n(s'_{n,j} | s_{n,j}, a_{n,t}) max_{a' \in \mathcal{A}} Q_{n,j}(s'_{n,j}, a'_{n,j})\right].$$

8. **end**

9. **end**

10. ⊳ After local updates, each agent shares its partially trained Q_n

$$\bar{Q}_n \leftarrow \frac{1}{N} \sum_{n=1}^{N} Q_n, \ \forall s, a$$

11. end

Figure 10 depicts the basic functioning of the FRL-based method for determining the appropriate hyperparameter. The upper layer is depicted in detail in Figure 9, while the bottom part is based on Algorithm 1. In a discrete action space, FRL

agents efficiently distribute FC module output powers. First, each control module sends the optimal action based on the aggregated model and receives a reward. Second, each FC module conducts multiple local control policy adjustments using the reward value. The agents trade their learned models and average the aggregated model to generate a consistently approved policy. Repeat training until all module control policies converge. Each agent has a highly tuned aggregated local control strategy for unseen states after learning.

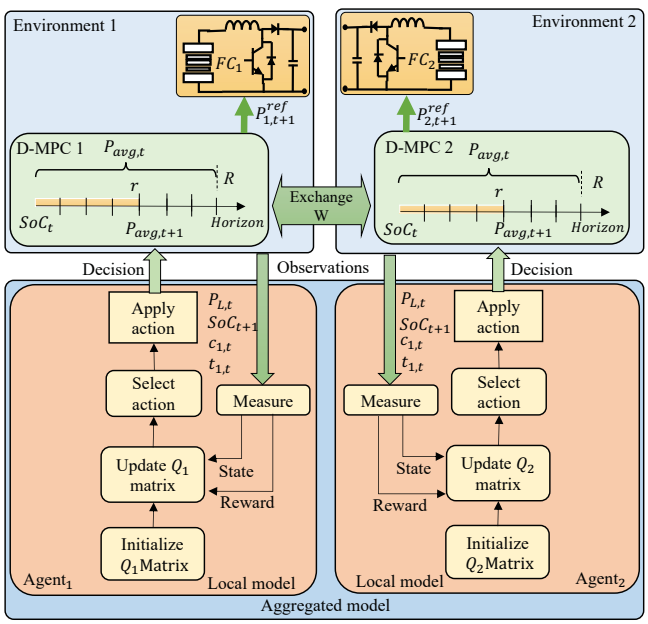

Fig. 10. The general operation of the FRL-based approach to seeking the optimal hyperparameter based on the current powertrain states.

Based on the convergence proof presented in [46], the QAvg algorithm with $E \geq 1$ not only converges but also approaches the Q function of π_I^* in an imagined environment $\mathcal{M}_I$ when considering the number of local updates as E, $\eta_j = \frac{2}{(1-\gamma)(j+E)}$, and j is the iteration number. Authors in [46] have demonstrated that

$$\left\| \bar{Q}_j - Q_I^{\pi_I^*} \right\|_\infty \leq \frac{16\gamma E}{(1-\gamma)^3 (j+E)}. \tag{14}$$

when $E = 1$, the agents exchange information after each local update of their Q functions. However, this process can be time-consuming and impractical. On the other hand, when $E = \infty$, the algorithm never involves communication [46].

V. Results and Discussions on Numerical Case Studies

This section explains the D-MPC and FD-MPC optimization results. DP and MPC both solve the same EMS problem. DP is the optimal global solution. The MPC answer is used to assess the accuracy, convergence speed, and communication performance of DCO-based MPCs. To quantitatively test proposed methods, processing times are compared on the same system (processor: Apple M1 Pro, RAM: 32 GB, MATLAB: R2021b). The learning parameters are based on learning sequences and system requirements [73, 74].

A. Regular Operation Optimized Results

Figure 11 shows the MPC, D-MPC, and FD-MPC optimum output powers by means of the real driving profile. The 10-step prediction horizon is chosen. The battery pack charges during deceleration and discharges during acceleration. Also, the battery pack has a strong power fluctuation propensity that follows load changes. The FC modules' minimal power fluctuation rates make PEMFC models' powers very steady. PEMFC modules' output power is minimal. D-MPC and FD-MPC power profiles and SoC curves match MPC with minimal errors. D-MPC and FD-MPC SoC curves track MPC with 0.0571 and 0.0597% errors, respectively.

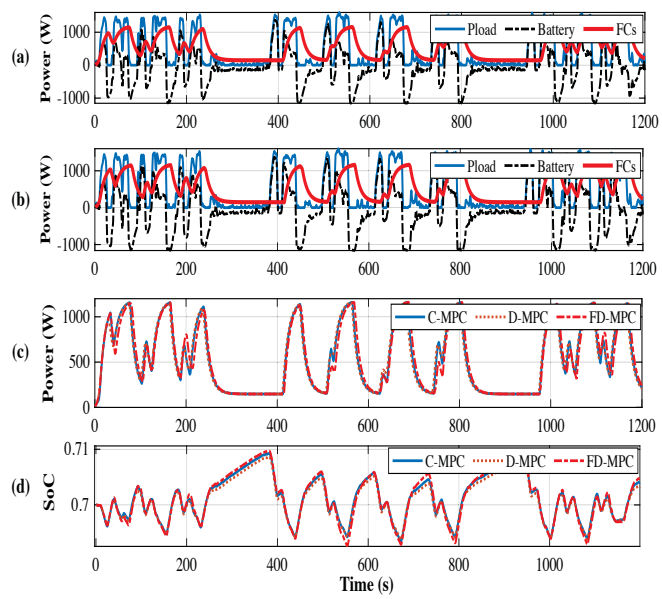

Fig. 11. Optimized results of the MPC-based approaches: (a) the powers based on D-MPC (Pload: the power profile, Pfc: the total power provided by the modules, and Pb: the battery power), (b) the powers based on FD-MPC, (c) the comparison between the total output modules powers of the FC modules, and (d) the comparison between SoC levels.

Table V compares the computational burden, T, and final operational costs, c_n, of decentralized algorithms to DP and MPC. The suggested approaches' final prices are very close to DP's. D-MPC and FD-MPC final costs are \$0.0617 and \$0.0628, 5.8997% and 7.7528% higher than DP and 1.7486% and 3.5291% higher than MPC. D-MPC and FD-MPC hydrogen consumptions contribute \$0.0173 and \$0.0109 under D-MPC and \$0.0174 and \$0.010 under FD-MPC, respectively. D-MPC and FD-MPC's second-largest expenses are module degradations, at \$0.0096 and \$0.0098, respectively. D-MPC and FD-MPC battery degradation costs are \$0.0048 and \$0.005, or 11.3% and 11.6% of final costs.

Figure 12 (a) presents the computational times of C-MPC, D-MPC, and FD-MPC under the real driving profile. The computational times of D-MPC and FD-MPC take an average of 0.0139 s and 0.0096 s, respectively. The computational times of D-MPC and FD-MPC are reduced by about 40.4040 % and 58.9062 % compared to MPC (0.0233 s). As shown in Fig. 12 (b), the FD-MPC iteration is reduced by 11.8384 %. In this study, a high-speed solution was chosen for FD-MPC convergence. FD-MPC can improve convergence speed if output power errors can be tolerated. If accuracy is the design criterion, more iterations are needed.

Table V
The detailed comparison of computational burden and final price

	DP	MPC	D-MPC	FD-MPC
T	-	25.5914	14.6237	9.2473
c_n	0.0583	0.0609	0.617	0.0628

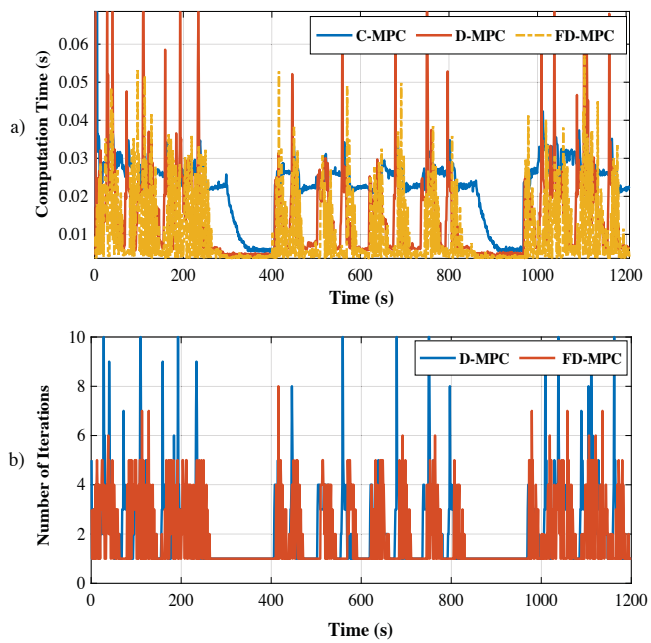

Fig. 12. (a) The computational burden of the developed MPC-based approaches and (b) The number of iterations based on the DCO-based MPCs.

B. *Impact of driving cycles on the optimal final cost of modular powertrain*

In addition to the actual driving profile, four additional light-duty cycles are taken into account here. In Fig. 5 and Table IV, all information relating to these profiles is shown. This demonstrates how the driving pattern influences the ideal cost. Figure 13 depicts the overall costs in $/mile for the four driving cycles using FD-MPC for a ten-step prediction horizon. The most severe HWFET is the priciest, costing around $0.055 per mile. In HWFET, the most significant expenses are the hydrogen costs of the two PEMFCs with the largest loads. The smoothest US06 results in the least expensive battery and two PEMFC degradations of around $0.034 per mile (38.18 percent less than HWFET). In FTP75, the PEMFCS and the loss of battery life have the highest costs of the four driving cycles.

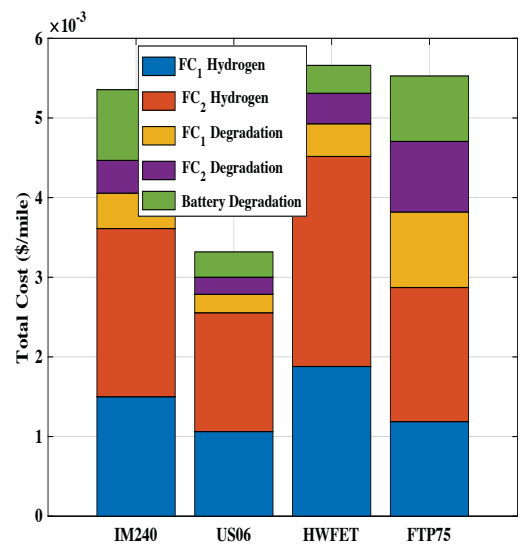

Fig. 13. Comparison of the hydrogen and degradation costs for the IM240, US06, HWFET, and FTP75 driving cycles.

C. *Influence of different driving cycles on the Horizon policy*

The moving horizon distribution for each driving cycle is compared in Fig. 14. The appropriate horizon range is significantly affected by driving characteristics. FD-MPC recommends utilizing a distinct prediction horizon for each zone of state space and determining whether to employ long and short horizons. On the short horizon, the MPC is unable to achieve the lowest cost, preferring to use more hydrogen at a high degradation cost. However, the learning-based policy mitigates this problem to some extent by simultaneously evaluating the long and short horizons. This demonstrates that the FRL approach can determine the range of the optimization horizon for various driving cycles in an adaptable manner.

D. *Impact of prediction horizon length*

Figure 15 compares final costs and computational times for different prediction horizon lengths under the real driving profile. As the moving optimization window length increases, final costs and computational burdens decrease. If the prediction moving window length is too short, the optimized power values produce unsatisfactory finite horizon approximations. MPC's execution time increases linearly with increasing moving window length when compared to others, but significantly when the optimization horizon dimension exceeds 12s. FD-MPC has 61.54% less computational burden

and 76.95% fewer sensitivities than D-MPC. It can be concluded that, as the length of the optimization horizon increases, the FD-MPC technique incurs a smaller increase in calculation time than the D-MPC and C-MPC methods. Figure 15 shows that the provided technique costs more than C-MPC. It must be stressed, however, that this modest increase in the final price can be accompanied by a substantial reduction in computational effort.

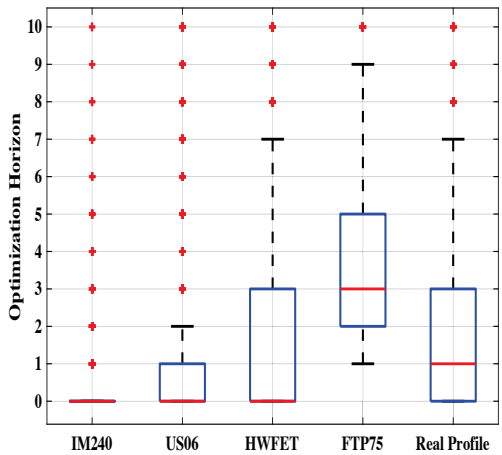

Fig. 14. The distribution of the horizon selected by FRL during the driving cycles.

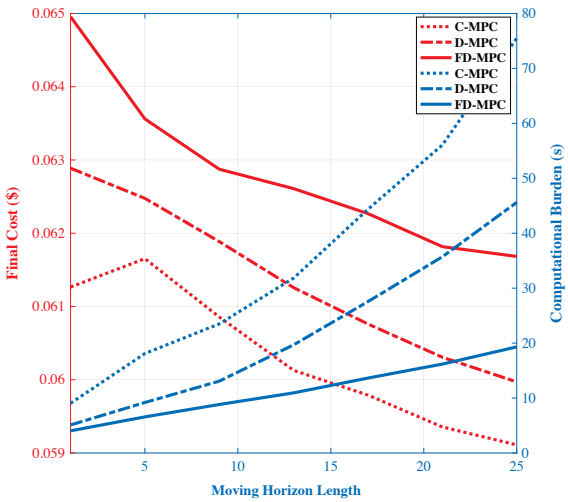

Fig. 15. Optimal final cost and computational time of Cen-MPC, Dec-MPC, and adjustable Dec-MPC as functions of optimization window size.

E. *Investigating multi-objective reward functions*

The role of rewards is crucial to the success of the designed learning technique. For the purpose of analyzing the influence of the reward coefficient, an analysis is performed to demonstrate the variation of the final cost and the total time under the five driving cycles, as shown in Fig. 16. The optimal function gain, based on the normalized reward function, is between 0.3 and 0.4.

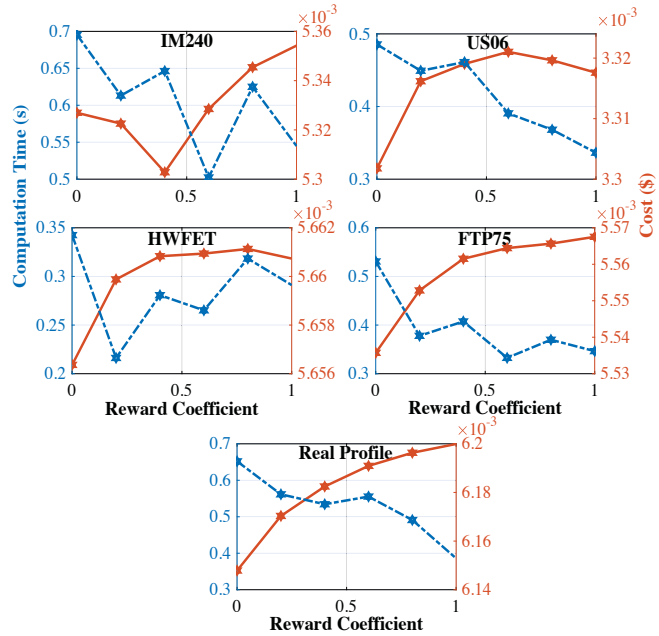

Fig. 16: The analysis of the computational burden and total cost fluctuations of the developed approach, from considering only the final cost (Reward coefficient = 0) to considering only the computational time (Reward coefficient = 1).

F. *Examination of Module Numbers*

To highlight the scalability of the presented design, the calculation times for four and eight agents are compared to those of the preceding two-agent system. Figure 17 contrasts the durations of learning for these two different cases. For FD-MPC, each case was evaluated ten times, and the mean results are shown below. D-MPC requires 51.74 percent and 39.22 percent less calculation time with four and eight agents compared to two agents, respectively. Comparing the four-agent and eight-agent scenarios to the two-agent scenario, DF-MPC reveals reductions of 73.26 and 72.16 percent, respectively. It is mostly due to the fact that the learned Q tables are utilized to extend the model to new circumstances. Horizon rules converge faster and attain the same level of precision as policies developed by integrating all FC module data.

VI. EXPERIMENTAL IMPLEMENTATION

To evaluate D-EMSs, a small-scale experiment is conducted on a modular test bench using the real driving profile to verify earlier numerical studies. Fig. 1 displays the developed modular powertrain test bench and the National Instruments (NI) Compact RIO device with the model number NI9022, which were used to implement the proposed strategy. The control strategy for both modules was programmed in MATLAB software and then transmitted to NI9022. The optimisation

methods have been meticulously developed to ensure the efficient performance of the decentralised optimisation approach.

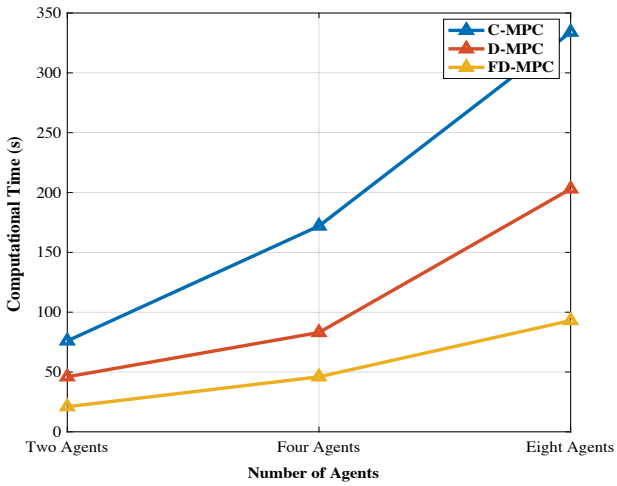

Fig. 17. The evolution of the computational times of C-MPC, D-MPC, and FD-MPC under two-agent, four-agent, and eight-agent cases

Figure 18 depicts the FD-MPC's output powers and SoC level. Figure 19 shows FC module power distributions. The results indicate that modules run near efficiency zones to reduce the hydrogen economy. Module power ramping fulfilled all requirements. The battery pack charges in low-power zones and discharges in high-power zones. The battery compensates for fast-dynamic load variations. The battery SOC oscillation level shows that the FD-MPC helps its two modules obtain the lowest departures from the original SOC levels. FD-MPC adds 6.32 percent to the simulation cost ($0.0668). FD-MPC is 50.07 percent faster than C-MPC.

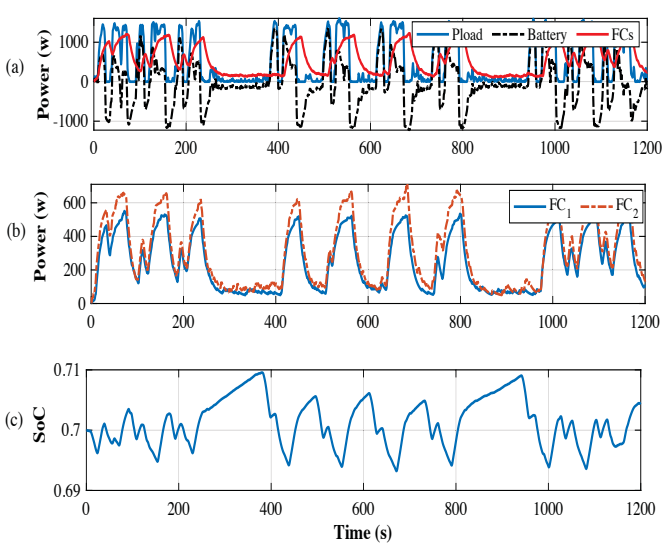

Fig.18. The experimental results of the suggested FD-MPC: (a) the output power profiles, (b) the FC modules' power profiles, and (c) the SoC level fluctuations.

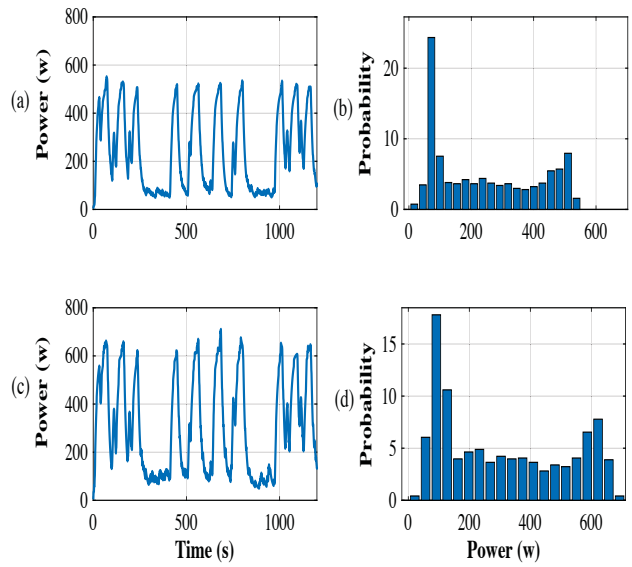

Fig.19. The experimental results of the FD-MPC approach: (a) the power profile of FC_1, (b) the distribution of FC_1, (c) the power profile of FC_2, and (d) the distribution of FC_2.

VII. Conclusion

This study proposes a fully decentralized energy management architecture to provide scalability and adaptability for the onboard powertrain components in FCV applications. C-ADMM is utilized to decompose the multiple objective minimization power-sharing problem, which incorporates hydrogen consumption and degradation costs, into its constituent parts. To enhance D-MPC's computing time, federated reinforcement learning is proposed for hyperparameter adjustment. Each module interacts with its neighbor on future power responses in order to agree on the best solutions, thereby providing a cost-effective and adaptable power splitting system. Several numerical and experimental investigations examine optimal solution precisions, data processing time efficiency, convergence performances, and scalability. Simulation and implementation outcomes revealed that the suggested strategy outperformed centralized approaches. A future study may concentrate on the following: It is possible to construct a general-purpose, non-linear, decentralized method for FCV decision-making. The reliability of decentralized approaches with delays and data loss might be evaluated. For several FCV applications, it is possible to specify the size and number of parallel modules. The recommended method may be simplified to enhance the real-time optimization of the FCV problem. Traffic data can be analyzed to increase system performance and decrease hydrogen use.

References

[1] Z. P. Cano *et al.*, "Batteries and fuel cells for emerging electric vehicle markets," *Nature Energy,* vol. 3, no. 4, pp. 279-289, 2018/04/01 2018, doi: 10.1038/s41560-018-0108-1.

[2] U. Eberle, B. Müller, and R. Helmolt, "Fuel cell electric vehicles and hydrogen infrastructure: Status 2012," *Energy & Environmental Science,* vol. 5, pp. 8790-8798, 07/30 2012, doi: 10.1039/C2EE22596D.

[3] N. Sulaiman, M. A. Hannan, A. Mohamed, E. H. Majlan, and W. R. Wan Daud, "A review on energy management system for fuel cell hybrid electric vehicle: Issues and challenges," *Renewable and Sustainable Energy Reviews,* vol. 52, pp. 802-814, 2015/12/01/ 2015, doi: https://doi.org/10.1016/j.rser.2015.07.132.

[4] H. S. Das, C. W. Tan, and A. H. M. Yatim, "Fuel cell hybrid electric vehicles: A review on power conditioning units and topologies," *Renewable and Sustainable Energy Reviews,* vol. 76, pp. 268-291, 2017/09/01/ 2017, doi: https://doi.org/10.1016/j.rser.2017.03.056.

[5] A. K. Soltani, M. Kandidayeni, L. Boulon, and D. L. St-Pierre, "Modular Energy Systems in Vehicular Applications," *Energy Procedia,* vol. 162, pp. 14-23, 2019/04/01/ 2019, doi: https://doi.org/10.1016/j.egypro.2019.04.003.

[6] N. Marx, L. Boulon, F. Gustin, D. Hissel, and K. Agbossou, "A review of multi-stack and modular fuel cell systems: Interests, application areas and on-going research activities," *International Journal of Hydrogen Energy,* vol. 39, no. 23, pp. 12101-12111, 2014.

[7] B. Somaiah and V. Agarwal, "Distributed maximum power extraction from fuel cell stack arrays using dedicated power converters in series and parallel configuration," *IEEE Transactions on Energy Conversion,* vol. 31, no. 4, pp. 1442-1451, 2016.

[8] X. Han, F. Li, T. Zhang, T. Zhang, and K. Song, "Economic energy management strategy design and simulation for a dual-stack fuel cell electric vehicle," *International Journal of Hydrogen Energy,* vol. 42, no. 16, pp. 11584-11595, 2017/04/20/ 2017, doi: https://doi.org/10.1016/j.ijhydene.2017.01.085.

[9] N. Herr *et al.*, "Decision process to manage useful life of multi-stacks fuel cell systems under service constraint," *Renewable Energy,* vol. 105, pp. 590-600, 2017/05/01/ 2017, doi: https://doi.org/10.1016/j.renene.2017.01.001.

[10] Y. Yan, Q. Li, W. Chen, W. Huang, and J. Liu, "Hierarchical Management Control Based on Equivalent Fitting Circle and Equivalent Energy Consumption Method for Multiple Fuel Cells Hybrid Power System," *IEEE Transactions on Industrial Electronics,* vol. 67, no. 4, pp. 2786-2797, 2019.

[11] H. Zhang, X. Li, X. Liu, and J. Yan, "Enhancing fuel cell durability for fuel cell plug-in hybrid electric vehicles through strategic power management," *Applied Energy,* vol. 241, pp. 483-490, 2019/05/01/ 2019, doi: https://doi.org/10.1016/j.apenergy.2019.02.040.

[12] T. Wang, Q. Li, X. Wang, W. Chen, E. Breaz, and F. Gao, "A Power Allocation Method for Multistack PEMFC System Considering Fuel Cell Performance Consistency," *IEEE Transactions on Industry Applications,* vol. 56, no. 5, pp. 5340-5351, 2020, doi: 10.1109/TIA.2020.3001254.

[13] W. Greenwell and A. Vahidi, "Predictive Control of Voltage and Current in a Fuel Cell–Ultracapacitor Hybrid," *IEEE Transactions on Industrial Electronics,* vol. 57, no. 6, pp. 1954-1963, 2010, doi: 10.1109/TIE.2009.2031663.

[14] B. Geng, J. K. Mills, and D. Sun, "Two-Stage Energy Management Control of Fuel Cell Plug-In Hybrid Electric Vehicles Considering Fuel Cell Longevity," *IEEE Transactions on Vehicular Technology,* vol. 61, no. 2, pp. 498-508, 2012, doi: 10.1109/TVT.2011.2177483.

[15] Amin, R. T. Bambang, A. S. Rohman, C. J. Dronkers, R. Ortega, and A. Sasongko, "Energy Management of Fuel Cell/Battery/Supercapacitor Hybrid Power Sources Using Model Predictive Control," *IEEE Transactions on Industrial Informatics,* vol. 10, no. 4, pp. 1992-2002, 2014, doi: 10.1109/TII.2014.2333873.

[16] X. Hu, C. Zou, X. Tang, T. Liu, and L. Hu, "Cost-Optimal Energy Management of Hybrid Electric Vehicles Using Fuel Cell/Battery Health-Aware Predictive Control," *IEEE Transactions on Power Electronics,* vol. 35, no. 1, pp. 382-392, 2020, doi: 10.1109/TPEL.2019.2915675.

[17] H. He, S. Quan, F. Sun, and Y.-X. J. I. T. o. I. E. Wang, "Model Predictive Control with Lifetime Constraints Based Energy Management Strategy for Proton Exchange Membrane Fuel Cell Hybrid Power Systems," 2020.

[18] J. Luna, E. Usai, A. Husar, and M. Serra, "Enhancing the efficiency and lifetime of a proton exchange membrane fuel cell using nonlinear model-predictive control with nonlinear observation," *IEEE transactions on Industrial Electronics,* vol. 64, no. 8, pp. 6649-6659, 2017.

[19] D. F. Pereira, F. D. C. Lopes, and E. H. Watanabe, "Nonlinear Model Predictive Control for the Energy Management of Fuel Cell Hybrid Electric Vehicles in Real-Time," *IEEE Transactions on Industrial Electronics,* pp. 1-1, 2020, doi: 10.1109/TIE.2020.2979528.

[20] S. Liu, Y. Bin, Y. Li, and B. Scheppat, "Hierarchical MPC Control Scheme for Fuel Cell Hybrid Electric Vehicles," *IFAC-PapersOnLine,* vol. 51, no. 31, pp. 646-652, 2018/01/01/ 2018, doi: https://doi.org/10.1016/j.ifacol.2018.10.151.

[21] H. Zheng, J. Wu, W. Wu, and Y. Wang, "Integrated Motion and Powertrain Predictive Control of Intelligent Fuel Cell/Battery Hybrid Vehicles," *IEEE Transactions on Industrial Informatics,* vol. 16, no. 5, pp. 3397-3406, 2020, doi: 10.1109/TII.2019.2956209.

[22] Y. Zhou, A. Ravey, and M.-C. Péra, "Multi-mode predictive energy management for fuel cell hybrid electric vehicles using Markov driving pattern recognizer," *Applied Energy,* vol. 258, p. 114057, 2020/01/15/ 2020, doi: https://doi.org/10.1016/j.apenergy.2019.114057.

[23] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[24] R. C. Hsu, S. Chen, W. Chen, and C. Liu, "A Reinforcement Learning Based Dynamic Power Management for Fuel Cell Hybrid Electric Vehicle," in *2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS)*, 25-28 Aug. 2016 2016, pp. 460-464, doi: 10.1109/SCIS-ISIS.2016.0104.

[25] N. P. Reddy, D. Pasdeloup, M. K. Zadeh, and R. Skjetne, "An intelligent power and energy management system for fuel cell/battery hybrid electric vehicle using reinforcement learning," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019: IEEE, pp. 1-6.

[26] W.-S. Lin and C.-H. Zheng, "Energy management of a fuel cell/ultracapacitor hybrid power system using an adaptive optimal-control method," *Journal of Power Sources,* vol. 196, no. 6, pp. 3280-3289, 2011.

[27] J. Yuan, L. Yang, and Q. Chen, "Intelligent energy management strategy based on hierarchical approximate global optimization for plug-in fuel cell hybrid electric vehicles," *International Journal of Hydrogen Energy,* vol. 43, no. 16, pp. 8063-8078, 2018.

[28] H. Sun, Z. Fu, F. Tao, L. Zhu, and P. J. J. o. P. S. Si, "Data-driven reinforcement-learning-based hierarchical energy management strategy for fuel cell/battery/ultracapacitor hybrid electric vehicles," vol. 455, p. 227964, 2020.

[29] X. Lin, B. Zhou, and Y. Xia, "Online Recursive Power Management Strategy based on the Reinforcement Learning Algorithm with Cosine Similarity and a Forgetting Factor," *IEEE Transactions on Industrial Electronics,* pp. 1-1, 2020, doi: 10.1109/TIE.2020.2988189.

[30] V. Mnih *et al.*, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602,* 2013.

[31] S. Levine, N. Wagener, and P. Abbeel, "Learning contact-rich manipulation skills with guided policy search (2015)," *arXiv preprint arXiv:1501.05611,* 2015.

[32] S. Liu, K. C. See, K. Y. Ngiam, L. A. Celi, X. Sun, and M. Feng, "Reinforcement learning for clinical decision support in critical care: comprehensive review," *Journal of medical Internet research,* vol. 22, no. 7, p. e18477, 2020.

[33] G. Dulac-Arnold *et al.*, "Challenges of real-world reinforcement learning: definitions, benchmarks and analysis," *Machine Learning,* vol. 110, no. 9, pp. 2419-2468, 2021.

[34] S. Levine, A. Kumar, G. Tucker, and J. Fu, "Offline reinforcement learning: Tutorial, review, and perspectives on open problems," *arXiv preprint arXiv:2005.01643,* 2020.

[35] A. J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition techniques in mathematical programming: engineering and science applications*. Springer Science & Business Media, 2006.

[36] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE,* vol. 95, no. 1, pp. 215-233, 2007, doi: 10.1109/JPROC.2006.887293.

[37] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.
[38] T. Logenthiran, D. Srinivasan, and A. M. Khambadkone, "Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system," *Electric Power Systems Research,* vol. 81, no. 1, pp. 138-148, 2011.
[39] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 38, no. 2, pp. 156-172, 2008.
[40] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," *arXiv preprint arXiv:1911.10635,* 2019.
[41] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*, 2017: PMLR, pp. 1273-1282.
[42] J. Qi, Q. Zhou, L. Lei, and K. Zheng, "Federated reinforcement learning: Techniques, applications, and open challenges," *arXiv preprint arXiv:2108.11887,* 2021.
[43] A. G. Roy, S. Siddiqui, S. Pölsterl, N. Navab, and C. Wachinger, "Braintorrent: A peer-to-peer environment for decentralized federated learning," *arXiv preprint arXiv:1905.06731,* 2019.
[44] H. H. Zhuo, W. Feng, Y. Lin, Q. Xu, and Q. Yang, "Federated deep reinforcement learning," *arXiv preprint arXiv:1901.08277,* 2019.
[45] X. Fan, Y. Ma, Z. Dai, W. Jing, C. Tan, and B. K. H. Low, "Fault-tolerant federated reinforcement learning with theoretical guarantee," *Advances in Neural Information Processing Systems,* vol. 34, pp. 1007-1021, 2021.
[46] H. Jin, Y. Peng, W. Yang, S. Wang, and Z. Zhang, "Federated Reinforcement Learning with Environment Heterogeneity," in *AISTATS*, 2022.
[47] A. Khalatbarisoltani, M. Kandidayeni, L. Boulon, and X. Hu, "Power Allocation Strategy based on Decentralized Convex Optimization in Modular Fuel Cell Systems for Vehicular Applications," *IEEE Transactions on Vehicular Technology,* pp. 1-1, 2020, doi: 10.1109/TVT.2020.3028089.
[48] A. Khalatbarisoltani, M. Kandidayeni, L. Boulon, and X. Hu, "Comparison of decentralized ADMM optimization algorithms for power allocation in modular fuel cell vehicles," *IEEE/ASME Transactions on Mechatronics,* 2021.
[49] F. Martel, S. Kelouwani, Y. Dubé, and K. Agbossou, "Optimal economy-based battery degradation management dynamics for fuel-cell plug-in hybrid electric vehicles," *Journal of Power Sources,* vol. 274, pp. 367-381, 2015/01/15/ 2015, doi: https://doi.org/10.1016/j.jpowsour.2014.10.011.
[50] O. Sundström, L. Guzzella, and P. Soltic, "Optimal hybridization in two parallel hybrid electric vehicles using dynamic programming," *IFAC Proceedings Volumes,* vol. 41, no. 2, pp. 4642-4647, 2008.
[51] U. DOE, "The fuel cell technologies office multi-year research, development, and demonstration plan," Technical report: US Department of Energy, 2016.
[52] M. Yue, S. Jemei, R. Gouriveau, and N. Zerhouni, "Review on health-conscious energy management strategies for fuel cell hybrid electric vehicles: Degradation models and strategies," *International Journal of Hydrogen Energy,* vol. 44, no. 13, pp. 6844-6861, 2019.
[53] P. Schneider, C. Sadeler, A.-C. Scherzer, N. Zamel, and D. Gerteisen, "Fast and reliable state-of-health model of a PEM cathode catalyst layer," *Journal of The Electrochemical Society,* vol. 166, no. 4, p. F322, 2019.
[54] S. Zhang, X.-Z. Yuan, J. N. C. Hin, H. Wang, K. A. Friedrich, and M. Schulze, "A review of platinum-based catalyst layer degradation in proton exchange membrane fuel cells," *Journal of Power Sources,* vol. 194, no. 2, pp. 588-600, 2009.
[55] B. Sompalli, B. A. Litteer, W. Gu, and H. A. Gasteiger, "Membrane degradation at catalyst layer edges in PEMFC MEAs," *Journal of The Electrochemical Society,* vol. 154, no. 12, p. B1349, 2007.
[56] P. Ren, P. Pei, Y. Li, Z. Wu, D. Chen, and S. Huang, "Degradation mechanisms of proton exchange membrane fuel cell under typical automotive operating conditions," *Progress in Energy and Combustion Science,* vol. 80, p. 100859, 2020.
[57] J. Park, H. Oh, T. Ha, Y. I. Lee, and K. Min, "A review of the gas diffusion layer in proton exchange membrane fuel cells: durability and degradation," *Applied Energy,* vol. 155, pp. 866-880, 2015.
[58] Y. Yang, X. Zhou, B. Li, and C. Zhang, "Recent progress of the gas diffusion layer in proton exchange membrane fuel cells: Material and structure designs of microporous layer," *International Journal of Hydrogen Energy,* vol. 46, no. 5, pp. 4259-4282, 2021.
[59] Y. Wang, S. J. Moura, S. G. Advani, and A. K. Prasad, "Power management system for a fuel cell/battery hybrid vehicle incorporating fuel cell and battery degradation," *International Journal of Hydrogen Energy,* vol. 44, no. 16, pp. 8479-8492, 2019.
[60] T. Fletcher, R. Thring, and M. Watkinson, "An Energy Management Strategy to concurrently optimise fuel consumption & PEM fuel cell lifetime in a hybrid vehicle," *international journal of hydrogen energy,* vol. 41, no. 46, pp. 21503-21515, 2016.
[61] Y. Zhou, A. Ravey, and M.-C. Péra, "Real-time cost-minimization power-allocating strategy via model predictive control for fuel cell hybrid electric vehicles," *Energy Conversion and Management,* vol. 229, p. 113721, 2021.
[62] X. Hu, C. Zou, X. Tang, T. Liu, and L. Hu, "Cost-optimal energy management of hybrid electric vehicles using fuel cell/battery health-aware predictive control," *ieee transactions on power electronics,* vol. 35, no. 1, pp. 382-392, 2019.
[63] H. Chen, P. Pei, and M. Song, "Lifetime prediction and the economic lifetime of Proton Exchange Membrane fuel cells," *Applied Energy,* vol. 142, pp. 154-163, 2015/03/15/ 2015, doi: https://doi.org/10.1016/j.apenergy.2014.12.062.
[64] J. Solano, S. Jemei, L. Boulon, L. Silva, D. Hissel, and M.-C. Pera, "IEEE VTS Motor Vehicles Challenge 2020-Energy Management of a Fuel Cell/Ultracapacitor/Lead-Acid Battery Hybrid Electric Vehicle," in *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2019: IEEE, pp. 1-6.
[65] NREL DriveCAT - Chassis Dynamometer Drive Cycles [Online] Available: www.nrel.gov/transportation/drive-cycle-tool
[66] S. Satyapal, "US Department of energy hydrogen and fuel cell technology overview," *Presented at The 14th International Hydrogen and Fuel Cell Expo (FC EXPO 2018),* 2018.
[67] T. Benjamin, R. Borup, N. Garland, C. Gittleman, B. Habibzadeh, and S. Hirano, "Fuel Cell Technical Team Roadmap," *Energy. gov (Office of Energy Efficiency & Renewable Energy),* 2017.
[68] K. Mongird *et al.*, "Energy Storage Technology and Cost Characterization Report," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2019.
[69] E. F. Camacho and C. B. Alba, *Model predictive control*. Springer science & business media, 2013.
[70] J. Mattingley, Y. Wang, and S. Boyd, "Receding horizon control," *IEEE Control Systems Magazine,* vol. 31, no. 3, pp. 52-65, 2011.
[71] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning,* vol. 3, no. 1, pp. 1-122, 2011.
[72] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning,* vol. 8, no. 3-4, pp. 279-292, 1992.
[73] R. Liessner, J. Schmitt, A. Dietermann, and B. Bäker, "Hyperparameter Optimization for Deep Reinforcement Learning in Vehicle Energy Management," in *ICAART (2)*, 2019, pp. 134-144.
[74] F. C. Fernandez and W. Caarls, "Parameters tuning and optimization for reinforcement learning algorithms using evolutionary computing," in *2018 International Conference on Information Systems and Computer Science (INCISCOS)*, 2018: IEEE, pp. 301-305.

Arash Khalatbarisoltani (S'18) received the Ph.D. degree in electrical engineering from the Université du Québec à Trois-Rivières (UQTR), Canada, in 2022. He is a Postdoctoral Researcher with the College of Mechanical and Vehicle Engineering, Chongqing University, China. His research interests include fuel cell systems, energy management, renewable energy, and intelligent transport systems.

Loïc Boulon (M'10, SM'15) received the master's degree in electrical and automatic control engineering from the University of Lille (France) in 2006. Then, he obtained a Ph.D. in electrical engineering from the University of Franche-Comté (France). Since 2010, he has been a professor at UQTR (Full Professor since 2016) and works at the Hydrogen Research Institute (Deputy director since 2019). His work deals with modeling, control, and energy management of Multiphysics systems. His research interests include hybrid electric vehicles, energy, and power sources (fuel cell systems, batteries, and ultracapacitors). He has published more than 120 scientific papers in peer-reviewed international journals and international conferences and has given over 35 invited conferences worldwide. Since 2019, he is the world's most-cited author of the topic "Proton exchange membrane fuel cells (PEMFC); Fuel cells; Cell stack" in Elsevier SciVal. In 2015, Loïc Boulon was general chair of the IEEE-Vehicular Power and Propulsion Conference in Montréal (QC, Canada). Prof. Loïc Boulon is now VP-Motor Vehicles of the IEEE Vehicular Technology Society, and he founded the "International Summer School on Energetic Efficiency of Connected Vehicles" and the "IEEE VTS Motor Vehicle Challenge". He is the holder of the Canada Research Chair in Energy Sources for the Vehicles of the Future.

Xiaosong Hu (SM'16) received the Ph.D. degree in automotive engineering from the Beijing Institute of Technology, Beijing, China, in 2012. He did scientific research and completed the Ph.D. dissertation in Automotive Research Center at the University of Michigan, Ann Arbor, MI, USA, between 2010 and 2012.

He is currently a Professor with the Department of Mechanical and Vehicle Engineering, Chongqing University, Chongqing, China. He was a Postdoctoral Researcher with the Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA, between 2014 and 2015, as well as at the Swedish Hybrid Vehicle Center and the Department of Signals and Systems at Chalmers University of Technology, Gothenburg, Sweden, between 2012 and 2014. He was also a Visiting Postdoctoral Researcher with the Institute for Dynamic Systems and Control at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 2014. His research interests include modeling and control of alternative powertrains and energy storage systems.

Dr. Hu has been the recipient of numerous prestigious awards/honors, including Web of Science Highly Cited Researcher by Clarivate Analytics, SAE Environmental Excellence in Transportation Award, IEEE ITSS Young Researcher Award, SAE Ralph Teetor Educational Award, Emerging Sustainability Leaders Award, EU Marie Currie Fellowship, ASME DSCD Energy Systems Best Paper Award, and Beijing Best Ph.D. Dissertation Award. He is an IET Fellow.