Les Cahiers du GERAD ISSN: 0711–2440

Optimal energy trading in residential prosumer clusters via graphon mean field games

M. Aziz, H. Suprême H. Dagdougui, A. Lesage-Landry

G–2025–80

November 2025

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2025
– Bibliothèque et Archives Canada, 2025

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2025
– Library and Archives Canada, 2025

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Optimal energy trading in residential prosumer clusters via graphon mean field games

Mohamad Aziz [a, b, c]

Hussein Suprême [a, c]

Hanane Dagdougui [a, b, d]

Antoine Lesage-Landry [a, b, d]

[a] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[b] *Polytechnique Montréal, Montréal (Qc), Canada, H3T 1J4*

[c] *Gestion des Réseaux, Institut de Recherche d'Hydro-Québec (IREQ), Varennes, Québec, Canada*

[d] *MILA Québec Artificial Intelligence Institute, Montréal, Québec, Canada*

mohamad.aziz@polymtl.ca
supreme.hussein2@hydroquebec.com
hanane.dagdougui@polymtl.ca
antoine.lesage-landry@polymtl.ca

November 2025
Les Cahiers du GERAD
G–2025–80

Abstract : In this work, we tackle the optimal energy trading (OET) problem in distribution grids with a large number of prosumer households. We first introduce a clustering architecture that partitions the grid into residential prosumer clusters (RPCs), each managed by an aggregator responsible for internal energy coordination and external interactions with neighbouring clusters and the distribution system operator. To optimize energy exchanges within and across clusters, we develop a novel decentralized control framework based on graphon mean field game theory. This framework models the OET as a decentralized dynamic game, deriving optimal control strategies that minimize each prosumer's individual energy cost. Household dynamics and cost functions are influenced by both local aggregate effects within their RPC and global interactions across interconnected cluster. Numerical experiments conducted on a dense energy network with 100 clusters, each containing 200 uniform households, validate the effectiveness of the proposed method in achieving scalable and cost-efficient energy coordination.

Keywords: Decentralized energy management; graphon mean field game; optimal energy exchange; optimal pricing mechanisms; residential prosumer cluster; home energy management system

Résumé : Dans ce travail, nous nous intéressons au problème de l'Optimal Energy Trading (OET) dans les réseaux de distribution intégrant un grand nombre de foyers prosommateurs. Nous proposons d'abord une architecture de regroupement qui partitionne le réseau en des groupes de prosommateurs résidentiels (RPC), chacun étant supervisé par un agrégateur chargé de la coordination énergétique interne et des interactions externes avec les groupes voisins ainsi qu'avec l'opérateur du système de distribution. Pour optimiser les échanges d'énergie inter et intra groupes, nous avons développé un cadre de contrôle décentralisé innovant fondé sur la théorie des graphon mean field games. Ce dernier modélise l'OET comme un jeu dynamique décentralisé et permet de dériver des stratégies de contrôle optimales visant à minimiser le coût énergétique individuel de chaque prosommateur. La dynamique des foyers et leurs fonctions de coût sont influencées à la fois par les effets agrégés locaux au sein de leur RPC et par les interactions globales entre groupes interconnectés. Des simulations réalisées sur un réseau dense comprenant 100 clusters, chacun composé de 200 foyers homogènes, démontrent l'efficacité de la méthode proposée pour atteindre une coordination énergétique évolutive et économiquement avantageuse.

Mots clés : Gestion décentralisée de l'énergie; graphon mean field game; optimal energy exchange; mécanismes de tarification optimale; cluster de prosommateurs résidentiels; système de gestion énergétique domestique

Acknowledgements: This work is supported by the NSERC Alliance-Mitacs Accelerate grant ALLRP 571311–21 ("Optimization of future energy systems") in collaboration with Hydro-Québec.

1 Introduction

The widespread adoption of demand-side management technologies and distributed energy resources (DERs), such as rooftop solar panels, home energy storage systems (ESSs), and electric vehicles (EVs), has transformed households into active residential prosumers. These prosumers now play a dynamic role in enhancing grid flexibility by modulating their electricity consumption and generation in response to system conditions. However, the individual impact of a single household on the broader demand-supply balance remains limited, reinforcing the need for coordinated mechanisms that aggregate prosumer actions to improve grid reliability and resilience. To address this challenge, we propose a novel cluster-of-clusters (CoC) architecture for residential energy networks. In this framework, multiple prosumer households, equipped with DERs are organized into residential prosumer clusters (RPCs). Each RPC is managed by an aggregator that coordinates internal energy exchanges and interfaces with neighbouring clusters and the distribution system operator (DSO). The CoC architecture enables bidirectional energy trading between clusters, facilitated by an energy management system (EMS) that supports real-time energy exchange.

In this paper, we adopt the definition of demand response (DR) from [14], which characterizes it as voluntary, active, and temporary adjustments in electricity consumption or production by customer-sited energy resources in response to external signals such as prices, commands, or measurements. Moreover, each RPC consists of a large number of uniformly equipped prosumer households, identical in terms of appliances and generation technologies, and a central aggregator. Figure 1 illustrates the proposed hierarchical CoC architecture, where each RPC connects to a central grid provider, e.g., a utility or the DSO. Through its aggregator, each RPC engages in bidirectional energy exchange with neighbouring clusters, forming the basis of the proposed coordination mechanism.

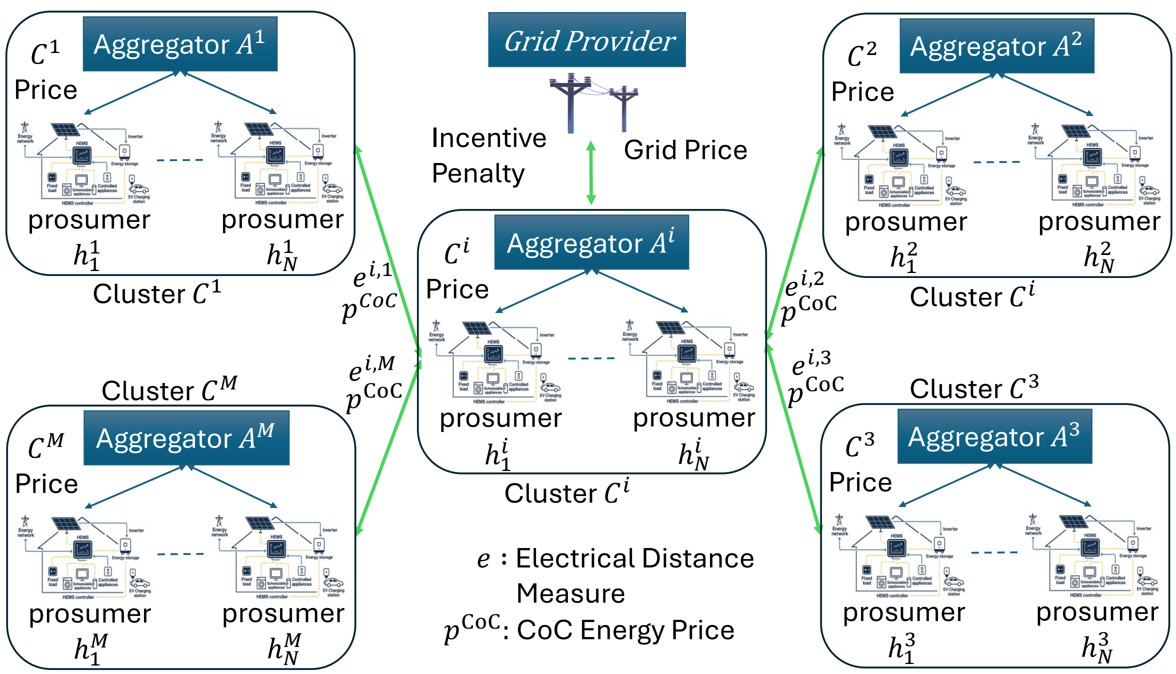

Figure 1: Cluster of clusters architecture energy grid diagram

Recent research has explored decentralized energy trading and coordination mechanisms for residential microgrids and community energy systems. Peer-to-peer (P2P) and community-based markets have been widely studied to enhance prosumer autonomy and local welfare, with game-theoretic formulations such as Nash [18], Stackelberg [19], and cooperative games [13] proposed to capture strategic interactions under uncertainty. Network-aware P2P designs have introduced electrical-distance-based pricing and distribution-level locational marginal prices to mitigate congestion and losses in low-voltage networks [9]. On the control side, multi-agent reinforcement learning (MARL) has been applied to demand response and integrated energy systems, but broad stability and convergence guarantees remain limited for safety-critical grid operations [5, 20]. To address scalability and provide theoretical guarantees, mean field game theory has been employed for large populations of flexible resources [15], such as EVs [12], using coupled Hamilton–Jacobi–Bellman and Fokker–Planck–Kolmogorov equations. However, existing MFG-based approaches typically assume a single global mean field, abstracting away heterogeneity induced by network topology, while network-aware P2P mechanisms lack equilibrium-based control laws at household resolution.

This paper fills the gap by introducing a graphon mean field game (GMFG) framework tailored to the CoC residential architecture. Unlike prior models, our approach captures both intra-cluster and inter-cluster interactions through an electrical-distance-weighted graphon kernel, enabling network-aware coordination in addition to promoting privacy and allowing for scalability. This work's key contributions include:

1. GMFG formulation: The optimal energy trading (OET) problem is modelled as a decentralized GMFG, where each node represents a prosumer household. The framework captures both local (RPC-level) and global (network-level) mean field interactions.
2. Theoretical guarantees: Under standard monotonicity and convexity assumptions, the existence and uniqueness of the GMFG equilibrium are established using coupled GMFG-Hamilton–Jacobi Bellman and GMFG-Fokker–Planck–Kolmogorov equations.
3. Dual-layer pricing: Decentralized control strategies are derived, along with dual-layer equilibrium prices, local (within-cluster) and global (inter-cluster), to support scalable and fair energy exchange.
4. Numerical validation: Dedicated algorithms are developed and validated for solving the decentralized nonlinear optimization problem in the context of a dense distribution grid structure consisting of 100 RPCs each with 200 prosumer households.

The remainder of this paper is structured as follows. Section 2 introduces the system dynamics and control variables for a generic household, along with the formulation of the decentralized OET problem. Section 3 presents the GMFG approach, detailing the coupled GMFG-HJB and GMFG-FPK equations and establishing the existence and uniqueness of the fixed-point solution. Section 4 outlines the computational investigation of the proposed method. Finally, Section 5 sums up the paper and presents potential extensions to the proposed OET problem via the GMFG approach.

2 Generic household system dynamics and OET problem formulation

As illustrated in Figure 1, each RPC comprises a cluster of prosumer households connected to a single feeder. These households interact with their cluster aggregator through local trading price p_j^{RPC} and with other clusters via a global price p^{CoC}, forming the basis of the proposed coordination mechanism. Each household is equipped with a home energy management system (HEMS), depicted in Figure 2, which monitors and optimizes energy flows in real-time.

In this framework, the HEMS operates the controllable appliances, sets the shiftable appliance schedule, manages charging/discharging of local ESS and EV batteries, and communicates bidirectionally with the aggregator and grid provider. It estimates local and global mean field effects and computes optimal control strategies using GMFG theory. In this paper, we are considering a CoC with uniform households. Uniformity in the sense of load appliances, and DER technologies (standardized batteries for EVs, ESSs, and solar panels). The GMFG approach presented here can be easily generalized to non-uniform households by incorporating the statistical distribution functions that characterize the load and the DER technical operating parameters.

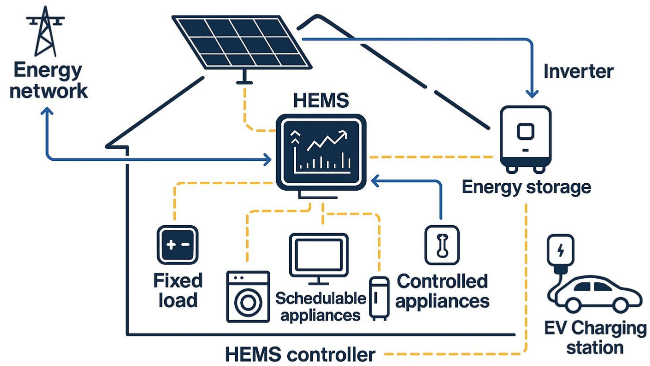

Figure 2: Generic household HEMS architecture

2.1 HEMS System variables

This subsection presents the system variable dynamics for a generic household h_i^j HEMS in a generic RPC C^j. The household's energy profile includes:

2.1.1 Generation

modelled as a stochastic process influenced by environmental variability, represented by a Brownian motion W_θ^i. The generation dynamics for a generic household h_i^j is as follows:

$$\partial\theta^i = P_{\mathrm{PV}}^i \partial t + \epsilon_\theta^i \partial W_\theta^i, \tag{1}$$

where $\epsilon_\theta > 0$ represents the intensity of stochastic fluctuations in solar generation, and P_{PV}^i is the generated solar power.

2.1.2 Load

composed of base load $y_{\mathrm{b}}^i(t)$, controllable $y_{\mathrm{c}}^i(t)$, and shiftable $y_{\mathrm{s}}^i(t)$ components. Each is modelled using independent Brownian W_z^i, W_{c}^i, W_{s}^i, with corresponding constants ϵ_z^i, ϵ_{c}^i, ϵ_{s}^i representing the variability intensity of each load type. The total load evolves as:

$$\partial y^i = z^i \partial t + u_{\mathrm{c}}^i P_{\mathrm{c}} \partial t + u_{\mathrm{s}}^i P_{\mathrm{s}} \partial t + \epsilon_y^i \partial W_y^i, \tag{2}$$

where ϵ_y^i aggregates the stochastic effects from all load components, W_y is a Brownian motion represent the commutative sum of the independent Brownian motions (W_z^i, W_{c}^i, W_{s}^i), u_{s}^i, and u_{c}^i are the control variables for the shiftable and controllable loads, respectively, and where $z^i := \frac{\mathrm{d}y_{\mathrm{b}}^i}{\mathrm{d}t}$ assuming the base load is differential. Lastly, P_{c} and P_{s} represent the rated power of the controllable appliances and shiftable appliances, respectively.

2.1.3 Storage dynamics

Each HEMS prioritizes energy from local PV generation to meet the local load. Surplus energy may be stored in the ESS or sold to the cluster aggregator. In case of a deficit, the HEMS can draw energy from the ESS, EV battery, or the CoC grid through its cluster aggregator. The HEMS is aware of the EV's schedule, distinguishing between grid-connected and on-road states. The states of energy (SoE) for ESS $b_{\mathrm{s}}^i(t)$ and EV $b_{\mathrm{e}}^i(t)$, when connected to the HEMS, are governed by bidirectional charging/discharging equations:

$$\partial b_{\mathrm{s}}^i = -r_{\mathrm{s}}^i(t)\partial t + \gamma_{\mathrm{s}}^i(t)\partial t + \epsilon_{\mathrm{s}}^i \partial W_{\mathrm{s}}^i \tag{3}$$

$$\partial b_{\mathrm{e}}^i = -r_{\mathrm{e}}^i(t)\partial t + \gamma_{\mathrm{e}}^i(t)\partial t + \epsilon_{\mathrm{e}}^i \partial W_{\mathrm{e}}^i, \tag{4}$$

where $r(t)$ and $\gamma(t)$ are discharging and charging rates, and the ϵ terms represent stochastic variability. Control actions ensure non-simultaneous operation and respect battery constraints. Denote by $\delta^i(t) = \theta^i(t) - y^i(t)$ the net energy after meeting local demand from DER generation at time t, and denote by $u_{\gamma_\delta}^i$ the control action for charging the ESS using $\delta^i(t)$ and $u_{\gamma_g,s}^i$ the control action for charging the ESS from the grid. First, consider the case $\delta^i(t) \geq 0$, then the charging and discharging values are given by:

$$\gamma_{\mathrm{s}}^i(t) = \begin{cases} \eta^{\mathrm{s}} \left(u_{\gamma_\delta}^i \delta^i(t) + u_{\gamma_{g,s}}^i P_{\mathrm{g2s}} \right), & \text{if} \quad b_{\mathrm{s}}^i(t) < b_{\max}^{\mathrm{s}} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$$r_{\mathrm{s}}^i(t) = \begin{cases} \frac{1}{\eta^{\mathrm{s}}} u_{r_{\mathrm{s}}}^i P_{\mathrm{s2g}}, & \text{if} \quad b_{\mathrm{s}}^i(t) > b_{\min}^{\mathrm{s}} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where $\eta^{\mathrm{s}} \in (0, 1]$ is the charging/discharging efficiency of ESS battery, P_{g2s} and P_{s2g} represent the rated power for charging and discharging of the ESS battery, and $b_{\min}^{\mathrm{s}}$, and $b_{\max}^{\mathrm{s}}$ are the b_{s} SoE capacity limits. If the EV is connected to the HEMS, the charging and discharging energy of the EV battery dynamics are as follows:

$$\gamma_{\mathrm{e}}^i(t) = \begin{cases} \eta^{\mathrm{e}} u_{\gamma_{\mathrm{e}}}^i(t) P_{\mathrm{g2e}}, & \text{if} \quad b_{\mathrm{e}}^i(t) < b_{\max}^{\mathrm{e}} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$r_{\mathrm{e}}^i(t) = \begin{cases} \frac{1}{\eta^{\mathrm{e}}} u_{r_{\mathrm{e}}}^i(t) P_{\mathrm{e2g}}, & \text{if} \quad b_{\mathrm{e}}^i(t) > b_{\min}^{\mathrm{e}} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where $\eta^{\mathrm{e}} \in (0, 1]$ is the charging/discharging efficiency of EV, $u_{\gamma_{\mathrm{e}}}$ and $u_{r_{\mathrm{e}}}$ represent the charging and discharging controls of the EV battery, respectively, P_{g2e}, and P_{e2g} are the rated power for EV charging and discharging, and $b_{\min}^{\mathrm{e}}$, and $b_{\max}^{\mathrm{e}}$ are the b_{e} SoE capacity limits.

Second, consider the case $\delta^i(t) \leq 0$, i.e., where the DER-generated energy is not sufficient to meet the total load. At this time, the HEMS decides to discharge the batteries or purchase energy from its cluster aggregator to meet the shortage. The control actions and the dynamics are described by:

$$\gamma_{\mathrm{s}}^i(t) = \begin{cases} \eta^{\mathrm{s}} u_{\gamma_{\mathrm{g,s}}}^i P_{\mathrm{g2s}}, & \text{if} \quad b_{\mathrm{s}}^i(t) < b_{\max}^{\mathrm{s}} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

$$r_{\mathrm{s}}^i(t) = \begin{cases} \frac{1}{\eta^{\mathrm{s}}} u_{r_{\mathrm{s}}}^i P_{\mathrm{s2g}}, & \text{if} \quad b_{\mathrm{s}}^i(t) > b_{\min}^{\mathrm{s}} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

$$\gamma_{\mathrm{e}}^i = \begin{cases} \eta^{\mathrm{e}} u_{\gamma_{\mathrm{e}}}^i P_{\mathrm{g2e}}, & \text{if} \quad b_{\mathrm{e}}^i(t) < b_{\max}^{\mathrm{e}} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

$$r_{\mathrm{e}}^i = \begin{cases} \frac{1}{\eta^{\mathrm{e}}} u_{r_{\mathrm{e}}}^i P_{\mathrm{e2g}}, & \text{if} \quad b_{\mathrm{e}}^i(t) > b_{\min}^{\mathrm{e}} \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

2.2 HEMS objective function

Based on the energy flow model in Section 2.1, the HEMS minimizes the household's energy cost by optimizing the net energy exchange $\xi^i(t) = d^i(t) - s^i(t)$, where $d^i(t)$ and $s^i(t)$ are demand and supply.

The cost function is:

$$
\begin{aligned}
l_j^i(t) =& \xi^i(t)\big(w_1(t)p_j^{\mathrm{RPC}}(t) + w_2(t)p^{\mathrm{CoC}}(t) + w_3(t)p^G\big) \\
&+ \alpha_{C_j}(t) - \beta_{C_j}(t) + O(\theta, t),
\end{aligned}
\tag{13}
$$

where:

- $w_1(t)$, $w_2(t)$, and $w_3(t)$, are time variant weights for local, inter-cluster (CoC), and grid-level energy exchanges, respectively, such that $w_k(t) \geq 0$ for all $k \in [1, 3]$ and $t \in [0, T]$ and $\sum_{k=1}^{3} w_k = 1$,
- Moreover, $w_3(t)$ ensures the balance in import/export is met through the grid provider.
- $p_j^{\mathrm{RPC}}(t)$ and $p^{\mathrm{CoC}}(t)$ are the local and global market prices, and p^{G} is the main grid set energy price,
- α_{C_j} and β_{C_j} are the penalty and reward terms,
- $O(\theta, t)$ captures DERs operating cost.

The aggregator A^j ensures fairness in pricing and incentives across the RPC. The HEMS computes optimal control actions using the GMFG framework, solving for equilibrium strategies based on local and global aggregate behaviours.

3 HEMS optimal controls via graphon mean field game theory

We now present our decentralized game-theoretical approach based on graphon mean-field game theory to address optimal energy trading in the proposed cluster-of-clusters energy grid. Within the proposed architecture, households in RPC C^j exchange energy with each other, with the aggregators of neighbouring clusters, and with the main grid.

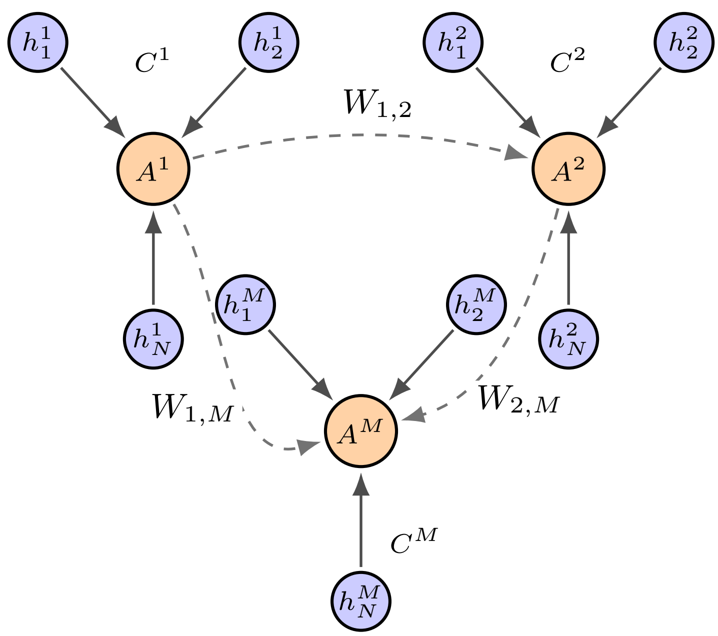

Figure 3: Cluster-of-clusters graph illustration

3.1 Cluster-of-clusters graph representation

Figure 3 depicts the cluster-of-clusters residential energy network as a weighted, undirected graph which is modelled as: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V} \subset \mathbb{N}$ is the set of nodes, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ the set of edges, and $W : \mathcal{V} \times \mathcal{V} \in [0, 1]$ a symmetric weight function representing the electrical distance between the nodes.

Nodes are partitioned into clusters $\{C^p\}_{p=1}^{M}$, $M \in \mathcal{N}$, each with aggregator A^p where M represents the number of clusters.

Denote by $e(x, y)$ the electrical distance between node $x \in \mathcal{V}$ and node $y \in \mathcal{V}$. In this framework, the electrical distance quantifies how strongly two nodes in a distribution network influence each other, and is derived from a sensitivity matrix based on the voltage–admittance relationship. The combined sensitivities ensure uniqueness and positive correlation, and the resulting distances can be used as edge weights (e.g., in the Louvain algorithm) to form partitions with high internal cohesion and low inter-region coupling. The electrical distance between the nodes is calculated using the methodology presented in [7], and is bounded such that $0 \leq e^{i,j} \leq 1$.

3.2 Graphon coefficient using the electrical distance

The interaction between agents x and y is modelled via a graphon derived from their electrical distance $e(x, y)$. In this work, we assume lossless intra-cluster exchange, modelled at the cluster level by an electrical distance matrix $e \in [0,1]^{M\times M}$ with $e_{pp} = 0$ and $e_{pq} = e_{qp} \in (0, 1]$ for $p \neq q$ and for all $p, q \in [1, M]$. Define the graphon weights $W_{pq} = \exp(-\alpha e_{pq}) \in [0, 1]$ and consider a measurable partition $[0, 1] = \dot{\cup}_{p=1}^{M} I_p$ with the cluster map $\kappa(x) = p$ if $x \in I_p$.

$$W(x, y) = W_{\kappa(x),\,\kappa(y)} = \exp(-\alpha\, e_{\kappa(x),\,\kappa(y)}), \tag{14}$$

such that $W : [0, 1]^2 \to [0, 1]$. Hence, $W(x, y)$, is symmetric, bounded, and measurable.

3.3 Local and global aggregates and price formulation

Let $\mathbf{X}(t, x) = (t, \theta(t, x), y_{\mathrm{c}}(t, x), y_{\mathrm{s}}(t, x), b_{\mathrm{s}}(t, x), b_{\mathrm{e}}(t, x)) \in \mathbb{R}^5$ be the state of a generic household $x \in [0, 1]$ at time t. Let $u(t, x) \in u$ be the control vector (load controls, ESS/EV charging/discharging decisions). Based on (1)–(4), the stochastic differential equation (SDE) for $\mathbf{X}(t, x)$ can be written as follows:

$$\mathrm{d}\mathbf{X}(t, x) = f\big(\mathbf{X}(t, x), u(t, x)\big)\, \mathrm{d}t + \Sigma\big(\mathbf{X}(t, x)\big)\, \mathrm{d}B_t(x) \tag{15}$$

where $B_t(x)$, a Brownian motion, f, the drift of the system dynamics, and Σ, the diffusion matrix (e.g., due to stochastic solar generation and demand), are Lipschitz in X, and $\mathcal{U}$ is convex and compact. Let $\xi(t, x) = d(t, x) - s(t, x)$ denote the agent's net exchange with the CoC. Denote by $\mu(t, \cdot)$ the population density of $X(t, \cdot)$ on $\mathbb{R}^4$ indexed by $x \in [0, 1]$. Define the cluster p aggregate by:

$$m_{\mathrm{loc}}^{(p)}(t) = \int_{I_p} \xi(t, z)\, \mu(t, z)\, \mathrm{d}z. \tag{16}$$

Let the CoC-aggregate as seen at node x be defined as:

$$A(x, t) = \int_0^1 W(x, y)\, \xi(t, y)\, \mu(t, y)\, \mathrm{d}y \tag{17}$$

and the scalar aggregate as:

$$\begin{aligned} M_{\mathrm{g}}(t) &= \int_0^1 A(x, t)\mathrm{d}x \\ &= \int_0^1 \int_0^1 W(x, y)\xi(t, y)\, \mu(t, y)\, \mathrm{d}y\mathrm{d}x. \end{aligned} \tag{18}$$

Local and global prices are generated by the Lipschitz pricing functions $\mathcal{P}_{\mathrm{loc}}, \mathcal{P}_{\mathrm{glob}} : \mathbb{R} \to \mathbb{R}_+$:

$$p_{(p)}^{\mathrm{RPC}}(t) = \mathcal{P}_{\mathrm{loc}}\big(m_{\mathrm{loc}}^{(p)}(t)\big) \tag{19}$$

$$p^{\mathrm{CoC}}(t) = \mathcal{P}_{\mathrm{glob}}\big(M_{\mathrm{g}}(t)\big) \tag{20}$$

3.4 GMFG-HJB and GMFG-FPK equations

The total cost-to-go, $V(t,x)$ for agent x is given by:

$$\min_{u(\cdot)} \mathbb{E}\Big[\int_t^T \Big(\ell(\mathbf{X}(s,x),u(s,x)) + \xi(s,x)\,[w_1 p^{\mathrm{RPC}}_{(\kappa(x))}(s) + w_2 p^{\mathrm{CoC}}(s) + w_3 p^{G}(s)]\Big) ds + \Phi(\mathbf{X}(T,x))\Big] \tag{21}$$

where Φ is the terminal cost function.

3.4.1 Hamilton–Jacobi–Bellman (HJB) equation

The HJB equation governs the value function $V(t,x)$, i.e., the minimal expected cost-to-go for agent x at time t:

$$\begin{aligned}-\partial_t V(t,x) = \min_{u\in u}\Big\{&\nabla_{\mathbf{X}} V(t,x)(f(\mathbf{X}(t,x),u(t,x)) \\ &+ \tfrac{1}{2}[\Sigma\Sigma^\top \nabla^2_{\mathbf{X}} V(t,x)] + \ell(\mathbf{X}(t,x),u(t,x)) \\ &+ \xi(t,x)\big[w_1 p^{\mathrm{RPC}}_{\kappa(x)}(t) + w_2 p^{\mathrm{CoC}}(t) + w_3 p^{G}(t)\big]\Big\}\end{aligned} \tag{22}$$

Let $\nabla_{\mathbf{X}} V^\top = (V_\theta, V_{y_\mathrm{s}}, V_{y_\mathrm{c}}, V_{b_\mathrm{s}}, V_{b_\mathrm{e}})$. Then, the switching scalars of each control in $\nabla_X V^\top$ are:

$$\begin{aligned}&\text{Loads:} && S_\mathrm{c} = P_\mathrm{c} V_{y_\mathrm{c}}, && S_\mathrm{s} = P_\mathrm{s} V_{y_\mathrm{s}} \\ &\text{ESS:} && S_{\gamma,\mathrm{s}} = (\eta_\mathrm{s} P_{\mathrm{g2s}})\, V_{b_\mathrm{s}}, && S_{\mathrm{r,s}} = -\left(\frac{P_{\mathrm{s2g}}}{\eta_\mathrm{s}}\right) V_{b_\mathrm{s}} \\ &\text{EV:} && S_{\gamma,\mathrm{e}} = (\eta_\mathrm{e} P_{\mathrm{g2e}})\, V_{b_\mathrm{e}}, && S_{\mathrm{r,e}} = -\left(\frac{P_{\mathrm{e2g}}}{\eta_\mathrm{e}}\right) V_{b_\mathrm{e}}.\end{aligned} \tag{23}$$

The optimal control decisions are then provided by:

$$u^*_\mathrm{c}(t,x,X) = \begin{cases} -1, & S_\mathrm{c} > 0 \\ 1, & S_\mathrm{c} < 0 \\ 0, & S_\mathrm{c} = 0 \end{cases} \tag{24}$$

$$u^*_\mathrm{s}(t,x,X) = \begin{cases} 1, & S_\mathrm{s} < 0 \\ 0, & S_\mathrm{s} \geq 0 \end{cases} \tag{25}$$

$$(u^*_{\gamma,\mathrm{s}}, u^*_{\mathrm{r,s}}) = \begin{cases} (0,0), & S_{\gamma,\mathrm{s}} \geq 0 \text{ and } S_{\mathrm{r,s}} \geq 0 \\ (1,0), & S_{\gamma,\mathrm{s}} < \min\{0, S_{\mathrm{r,s}}\} \\ (0,1), & S_{\mathrm{r,s}} < \min\{0, S_{\gamma,\mathrm{s}}\} \end{cases} \tag{26}$$

$$(u^*_{\gamma,\mathrm{e}}, u^*_{\mathrm{r,e}}) = \begin{cases} (0,0), & S_{\gamma,\mathrm{e}} \geq 0 \text{ and } S_{\mathrm{r,e}} \geq 0 \\ (1,0), & S_{\gamma,\mathrm{e}} < \min\{0, S_{\mathrm{r,e}}\} \\ (0,1), & S_{\mathrm{r,e}} < \min\{0, S_{\gamma,\mathrm{e}}\}. \end{cases} \tag{27}$$

3.4.2 Fokker–Planck–Kolmogorov equation

The propagation of the state probability density function $\mu(t,x)$ is calculated using the following FPK equation:

$$\begin{aligned}\partial_t \mu(t,x) = &-\nabla_{\mathbf{X}}\left(f\left(\mathbf{X}(t,x), u^*(t,x)\right)\mu(t,x)\right) \\ &+ \frac{1}{2}\nabla^2_{\mathbf{X}}\big(\Sigma\left(\mathbf{X}(t,x)\right)\Sigma^\top\left(\mathbf{X}(t,x)\right)\mu(t,x)\big).\end{aligned} \tag{28}$$

3.5 GMFG-loop and fixed-point solution

Let $\mu(t,x)$ denote the probability density function of the state vector $\mathbf{X}(t,x)$ for a generic household $x \in [0,1]$ at time t. Let $u^*(t,x)$ be the optimal control strategy derived from the GMFG-HJB equation. Define the GMFG operator $\mathcal{T}$ as:

$$\mathcal{T} : \mu \mapsto u^* \mapsto \mu', \tag{29}$$

where μ' is the updated distribution obtained by solving the GMFG-FPK equation using the control u^*. A fixed-point μ^* satisfies $\mathcal{T}(\mu^*) = \mu^*$. The operator $\mathcal{T}$ in (29) represents the GMFG-loop defined by the coupled system of the Graphon Mean Field Hamilton–Jacobi–Bellman (GMFG-HJB) and Graphon Mean Field Fokker–Planck–Kolmogorov (GMFG-FPK) equations. The goal is to establish the existence and uniqueness of a fixed-point solution (μ^*, u^*), where μ^* is the density function and u^* are the HEMS optimal control strategies. The following are the necessary assumptions for the existence and uniqueness of the fixed-point solution:

1. The drift $f(\mathbf{X}, u)$ and cost function $\ell(\mathbf{X}, u)$ are Lipschitz continuous in both $\mathbf{X}$ and u.
2. The control set $\mathcal{U}$ is convex and compact.
3. The graphon $W(x,y) = \exp(-\alpha e(x,y))$ is symmetric, bounded, and measurable with $W(x,y) \in [0,1]$.
4. The diffusion matrix $\Sigma(\mathbf{X})$ is uniformly bounded and positive definite.
5. The pricing functions p^{RPC} and p^{CoC} in (19) and (20) are Lipschitz continuous.
6. Graphon–Monotonicity hold: for any μ_1, μ_2 with net exchanges ξ_1, ξ_2,

$$\int_0^T \int_{[0,1]} \Big(\mathcal{C}(x,\mu_1) - \mathcal{C}(x,\mu_2)\Big)\, \mathrm{d}(\mu_1 - \mu_2)(x)\, \mathrm{d}t \;\geq 0$$

where $\mathcal{C}(x,\mu) = w_1\, \mathcal{P}_{\mathrm{loc}}(m_{\mathrm{loc}}^{(\kappa(x))}) + w_2\, \mathcal{P}_{\mathrm{glob}}(M_g) + w_3\, p^{\mathrm{G}}$.

Building on the preceding assumption, we now present theorems that establish the existence and uniqueness solution for the formulated OET problem.

Theorem 1 (Existence of GMFG equilibrium). *Based on the above assumptions, the GMFG operator* $\mathcal{T} : \mu \mapsto u^* \mapsto \mu'$ *has at least one fixed point* $\mu^\star \in \mathcal{C}([0,T]; \mathcal{P}_2(\mathbb{R}^4))$. *The pair* $(\mu^\star, u^\star)$ *solves the coupled GMFG–HJB/FPK system.*

Proof. Fix μ. Viscosity theory yields a solution to the HJB and a measurable selector $u^\star$; stability under Lipschitz perturbations gives continuity in the data [2]. Given $u^\star$, the FPK has a unique weak solution; tightness and relative compactness follow from standard a priori estimates or the gradient-flow viewpoint refer to [1]. The maps $\mu \mapsto (m_{\mathrm{loc}}, A, M_{\mathrm{glob}}) \mapsto (p^{\mathrm{RPC}}, p^{\mathrm{CoC}})$ are continuous, and the cut-norm continuity of graphon integrals [10]. Hence $\mathcal{T}$ is continuous with relatively compact range and the Schauder theorem [17] holds. □

Theorem 2 (Uniqueness under graphon–monotonicity). *Suppose the above assumptions holds and that* ℓ *is strictly convex in* u *and the coupling satisfies the Lasry–Lions monotonicity adapted to* (17)–(18). *Then, the GMFG equilibrium* $(\mu^\star, u^\star)$ *is unique.*

Proof. Let (μ_1, u_1) and (μ_2, u_2) be two equilibria. Consider the difference of HJB equations tested along optimal trajectories and add the FPK identities; strict convexity in u and monotonicity of the graphon-aggregated coupling force $\mu_1 = \mu_2$ and $u_1 = u_2$. See Lasry–Lions for the classical argument and recent generalizations [8, 11]. □

Theorem 3 (ε-Nash for large clustered grids). *The finite game conversion based on the infinite game if an* ε*-Nash Equilibrium exists in the finite game. Let* $\{G_N\}$ *be a sequence of household networks (cluster-of-clusters) whose step-graphons* W_{G_N} *converge in the cut metric to* W. *Let* $(\mu^\star, u^\star)$ *be the* *unique GMFG equilibrium for W. Then there exists $\varepsilon_N \to 0$ as $N \to \infty$ such that the controls $u^\star$ of G_N are an ε_N-Nash equilibrium for the N-agent game.*

Proof. Using the results of [16] graphon-game approximation, the equilibrium responses and payoffs are continuous in the cut metric. GMFG ε-Nash theorems yield $\varepsilon_N \to 0$ as $N \to \infty$ for dense graph sequences. For the full details underlying the N-player/GMFG convergence proof, we refer the reader to [3, 6]. □

Theorem 1 and theorem 2 present the existence and uniqueness of the fixed-point solution to the GMFG-loop in the infinite graph game, and for a finite CoC graph theorem 3 proves the existence of an ε_N Nash-equilibrium solution for the CoC OET problem.

4 Computational experiments and results

In essence, the HEMS observes the state variables of its household, i.e., generation, load, and SoE in the batteries. Using the GMFG algorithm, the HEMS calculates the equilibrium prices for energy exchange within its cluster and cluster of clusters. Subsequently, the HEMS finds the optimal control strategies.

4.1 Fixed-point algorithm for the GMFG OET problem

Figure 4 represents the algorithmic sketch for finding the unique solution for the GMFG-loop given a tolerance $\epsilon_{\text{con}} > 0$. The output of the algorithm is the fixed-point solution for the control actions and the probability density functions (PDFs).

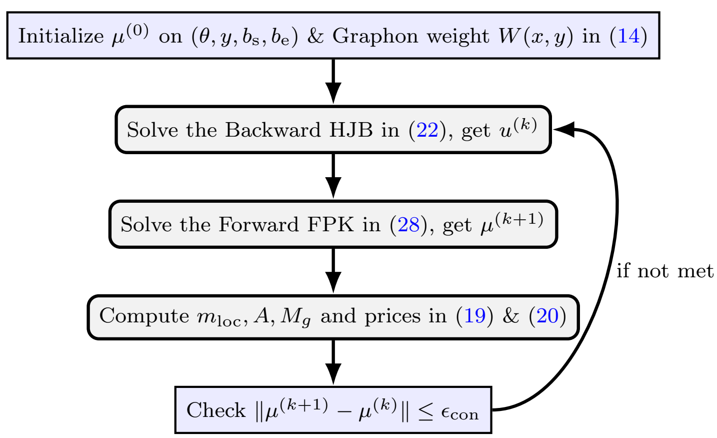

Figure 4: GMFG-loop fixed-point solution algorithmic steps

4.2 Numerical result

This subsection presents the numerical results for the OET problem via GMFG. The proposed approach is validated on a dense CoC energy grid consisting of 100 RPCs each containing 200 prosumers. The electrical distances are computed based on [7] and converted into a graphon weights based on equation (14). The system variables parameters considered for the simulations are: $P_c = 2.5\,\text{kW}$, $P_s = 2.0\,\text{kW}$, ESS $P_{g2s} = P_{s2g} = 3.0\,\text{kW}$ with $\eta_s = 0.95$, EV battery specs: $P_{g2e} = P_{e2g} = 7.0\,\text{kW}$ with $\eta_e = 0.93$, main grid tariff is flat (i.e., $p^G = \$0.07/kWh$), and $\epsilon_{\text{con}} = 10^{-3}$.

Figure 5 and Figure 6 present the converging PDFs and equilibrium prices obtained through the iterative process described in Figure 4. Household load is concentrated at lower demand levels, while

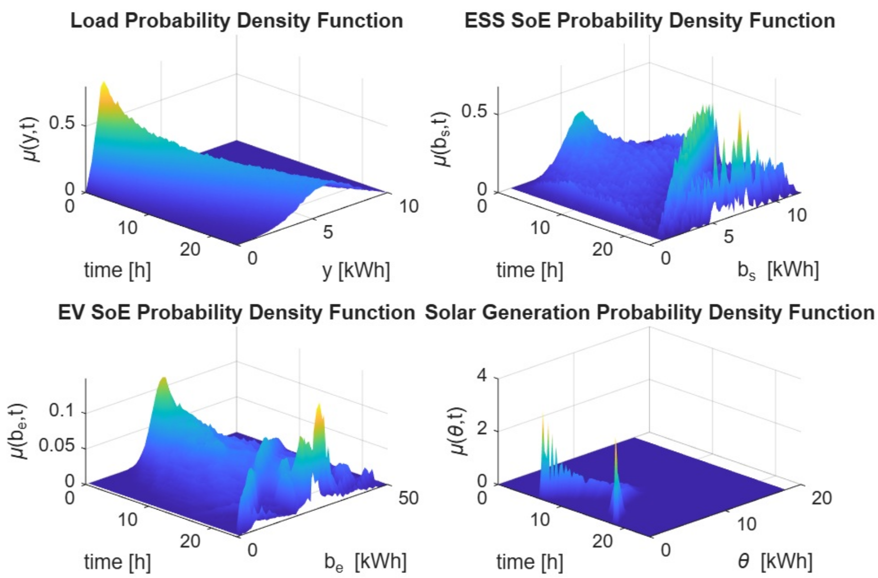

Figure 5: Probability density function solutions

solar generation remains skewed toward low output, reflecting limited PV capacity. ESS SoE distributions exhibit distinct peaks, indicating preferred operating ranges driven by optimal charge/discharge strategies, and EV battery PDFs reveal multiple peaks consistent with heterogeneous charging patterns such as overnight and daytime usage and EV availability.

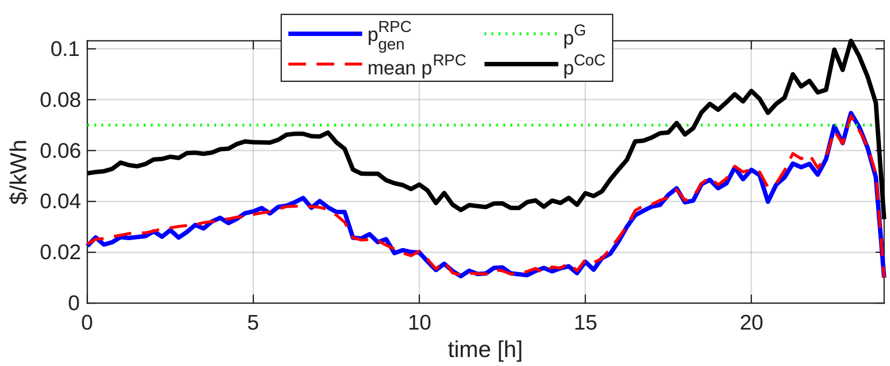

Figure 6: Equilibrium energy exchange prices $p^{\textbf{Coc}}(t), \quad p^{\textbf{RPC}}(t)$

The convergence of the PDFs and equilibrium price trajectories validates that the GMFG algorithm achieves a unique fixed-point solution to the GMFG-loop with an error ϵ_{con}, in our simulation $\epsilon_{\text{con}} = 10^{-3}$.

Figure 7 depicts the generic agent energy profile showing the state variables of its HEMS. In essence, the generic agent HEMS observes the state variables, and finds the optimal control strategies that minimize its energy bill by solving its HJB using the fixed-point solution of the GMFG-algorithm presented in Figure 5, and Figure 6.

4.3 CoC validation: Economic objective and scalability

The validation of the proposed CoC architecture addresses two key claims: *economic efficiency* and *scalability*.

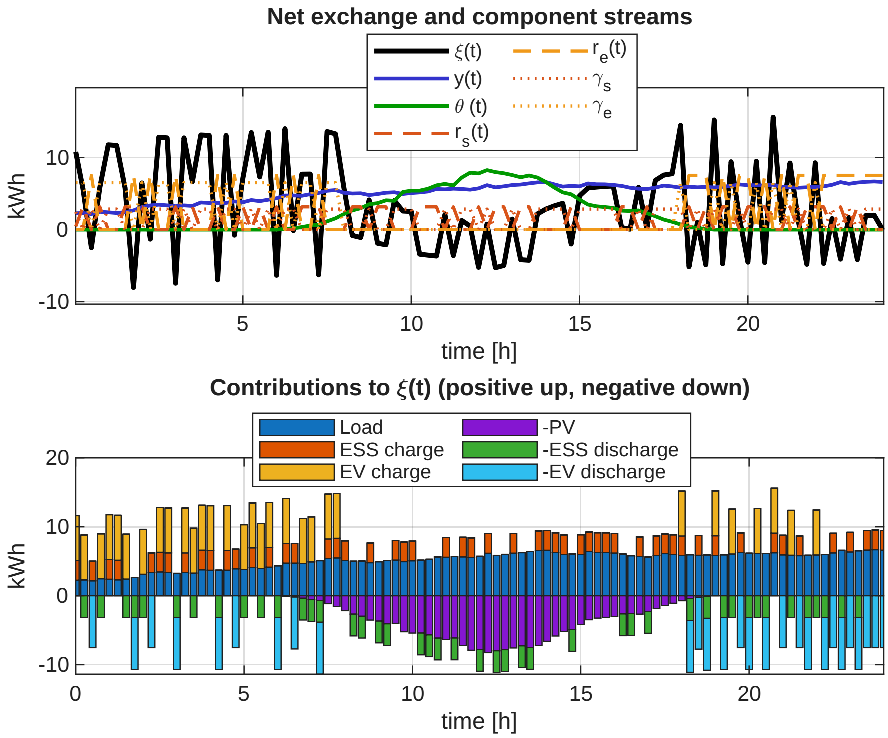

Figure 7: Energy profile for a generic agent

4.3.1 Economic validity

The objective of the HEMS-GMFG approach for the OET problem is to determine household-level optimal controls that minimize energy bills while accounting for both local and global interactions specifically, energy availability and corresponding equilibrium prices. To validate this claim, we compare the energy cost incurred by a generic agent operating under three configurations:

(i) Stand-alone mode: the agent interacts only with the main grid;
(ii) RPC-only mode: the agent interacts with its RPC and the main grid, no inter-RPC energy exchange;
(iii) CoC mode: the proposed architecture enabling inter-RPC energy exchange.

Numerical results for the generic agent (see Figure 7) show a significant cost reduction: 16% in RPC-only mode ($5.11) 29% in CoC mode ($4.33) compared to stand-alone operation ($6.09). These findings illustrate that the CoC architecture, combined with the GMFG-based optimization, achieves better economic performance by balancing the local RPC and CoC flexibility to better serve the participating prosumers.

4.3.2 Scalability and security

The GMFG algorithm computes the fixed-point solution offline, ensuring convergence before real-time operations. Once the equilibrium is established, each HEMS can determine optimal controls independently in real-time with negligible computational and communication overhead. The complexity of the GMFG fixed-point loop depends on the discretization of the random variables state space rather than the number of agents, due to the fact that the algorithm operates on probability distributions and estimates aggregate effects instead of simulating each prosumer. Consequently, runtime scales with grid resolution and the dimensionality of the random variables, not with network size. This design guarantees scalability for large networks, making the approach suitable for high-density residential energy systems.

5 Conclusion

This paper presents a novel decentralized optimization framework for optimal energy trading in large-scale residential prosumer networks using Graphon Mean Field Game theory. The proposed architecture models both intra- and inter-cluster interactions through an electrical-distance-weighted graphon kernel, enabling network-aware coordination while maintaining scalability and privacy. Theoretical analysis establishes the existence and uniqueness of GMFG equilibrium under standard monotonicity and convexity conditions and proves convergence of the fixed-point algorithm for the coupled GMFG–HJB and GMFG–FPK equations.

Numerically, we have tested the proposed framework on a dense cluster-of-clusters topology with 100 RPCs of 200 households each, showing up to a 29% reduction in energy costs relative to stand-alone operation. The simulations also demonstrate efficient convergence to equilibrium strategies with promising real-time performance with minimal agent-to-agent communication, confirming the method's scalability for large-scale residential prosumer grid.

Future research will extend the presented GMFG framework to incorporate the distribution grid power flow constraints (voltage and apparent power limits) and time-varying tariffs, generalizing the OET problem toward a decentralized optimal power flow formulation. Additionally, we aim to develop graphexon-based mean field game [4] models for sparse topologies, such as the IEEE 8500-bus feeder, to broaden applicability to realistic distribution networks.

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich. Birkhäuser, 2. ed edition, 2008.

[2] Martino Bardi and Italo Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhäuser Boston Inc., 1997.

[3] Peter E. Caines and Minyi Huang. Graphon mean field games and their equations. SIAM Journal on Control and Optimization, 59(6):4373–4399, 2021.

[4] Peter E. Caines and Minyi Huang. Mean field games on dense and sparse networks: The graphexon mfg equations. In 2024 American Control Conference (ACC), pages 4230–4235, 2024.

[5] Yicheng Chen, Wei Wei, Feng Liu, and Shengwei Mei. Reinforcement learning for integrated energy systems operation: A review. IEEE Transactions on Smart Grid, 12(4):3163–3176, July 2021.

[6] Markus Fischer. On the connection between symmetric n-player games and mean field games. The Annals of Applied Probability, 27, 05 2014.

[7] Jiacheng Ge, Zaijun Wu, Junjun Xu, and Qinran Hu. A two-stage flow-based partition framework for unbalanced distribution networks. CSEE Journal of Power and Energy Systems, 8:1–11, 2023.

[8] P. Jameson Graber and Alpár R. Mészáros. On monotonicity conditions for mean field games. Journal of Functional Analysis, 285(9):Article 110095, 2023.

[9] Jaysson Guerrero, Archie C. Chapman, and Gregor Verbič. Decentralized p2p energy trading under network constraints in a low-voltage network. IEEE Transactions on Smart Grid, 10(5):5163–5173, 2019.

[10] Svante Janson. Graphons, cut norm and distance, couplings and rearrangements. New York Journal of Mathematics, 4:77–143, 2013.

[11] Jean-Michel Lasry, Pierre-Louis Lions, and Olivier Guéant. Mean Field Games and Applications. Paris-Princeton Lectures on Mathematical Finance. Springer, 2011.

[12] Caroline Le Floch, Emre Can Kara, and Scott Moura. Pde modeling and control of electric vehicle fleets for ancillary services: A discrete charging case. IEEE Transactions on Smart Grid, 9(2):573–581, 2018.

[13] Chao Long, Jianzhong Wu, Yue Zhou, and Nick Jenkins. Peer-to-peer energy sharing through a two-stage aggregated battery control in a community microgrid. Applied Energy, 226:261–276, 2018.

[14] Johanna L. Mathieu, Gregor Verbic, Thomas Morstyn, Mads Almassalkhi, Kyri Baker, Julio Braslavsky, Kenneth Bruninx, Yury Dvorkin, Gregory S. Ledva, Nariman Mahdavi, Hrvoje Pandžic, Alessandra Parisio, and Vedran Peric. A new definition and research agenda for demand response in the distributed energy resource era. IEEE Transactions on Energy Markets, Policy and Regulation, 3:324–339, 2025.

[15] Francesca Parise, Marcello Colombino, Sergio Grammatico, and John Lygeros. Mean field constrained charging policy for large populations of plug-in electric vehicles. 53rd IEEE Conference on Decision and Control, pages 5101–5106, 2014.

[16] Francesca Parise and Asuman Ozdaglar. Graphon games: A statistical framework for network games and interventions. Econometrica, 91(1):191–225, 2023.

[17] Juliusz Schauder. Der fixpunktsatz in funktionalräumen. Studia Mathematica, 2:171–180, 1930.

[18] Tiago Sousa, Tiago Soares, Pierre Pinson, Fabio Moret, Thomas Baroche, and Etienne Sorin. Peer-to-peer and community-based markets: A comprehensive review. Renewable and Sustainable Energy Reviews, 104:367–378, 2019.

[19] Wayes Tushar, Chau Yuen, Tapan K. Saha, Thomas Morstyn, Archie C. Chapman, M. Jan E. Alam, Sarmad Hanif, and H. Vincent Poor. Peer-to-peer energy systems for connected communities: A review of recent advances and emerging challenges. Applied Energy, 282:116–131, 2021.

[20] José R. Vázquez-Canteli and Zoltán Nagy. Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied Energy, 235:1072–1089, 2019.