A Scalable Cooperative Game Theoretic Framework for Coalitional Energy Management

Liyang Han

Mansfield College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2020

Abstract

The advances in distributed energy resources (DER) propelled power networks worldwide to transition from a unidirectional model to a bidirectional one, under which uncertain and intermittent renewable generation cause load balancing issues. Energy storage as flexible resources have the potential to mitigate the stress on the network, and local energy markets emerged in recent years to encourage efficient management of these resources. However, the formulation of a market scheme that incentivizes such operation by offering profit for all the participants remains an unresolved challenge. This thesis proposes the novel application of cooperative game theory to coalitional energy management, which extends existing cooperative theoretic frameworks to the energy storage operation. The proposed scheme is then proven to have a non-empty core, ensuring the existence of a payoff allocation that incentivizes participation in the grand coalition of all participants.

This thesis focuses on two main payoff allocations: the Shapley value and the nucleolus.

The Shapley value is a commonly used payoff allocation in cooperative game theory, but is highly computationally complex. This thesis proposes a novel multi-step sampling method based on optimal stratified sampling for estimating the Shapley value, and demonstrates its effectiveness to improve the scalability of the proposed scheme. It is shown, however, that the Shapley value is not necessarily stabilizing, meaning that given the payoffs based on the Shapley value, some participants may choose to defect from the grand coalition and form smaller coalitions for higher profits.

The nucleolus, albeit a less commonly used payoff allocation in cooperative game theory, is proven to be stabilizing, preventing defection from the grand coalition. Similarly, the computation time of the nucleolus increases exponentially with the number of participants, strictly limiting the size of this scheme. This thesis proposes a new strategy which incorporates clustering techniques to estimate the nucleolus at reduced computation times, where a novel marginal contribution profile is constructed as the clustering features. While maintaining a high estimation accuracy, it is shown to easily scale up the scheme from less than 15 participants to over 100, a well applicable number of load buses in a real distribution feeder.

A Scalable Cooperative Game Theoretic Framework for Coalitional Energy Management



Liyang Han

Mansfield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2020

Feeling gratitude and not expressing it is like wrapping a present and not giving it.

— William Arthur Ward

Acknowledgements

What a journey!

I am grateful to have this little space in my thesis where I can strip away all my academic facade, become a human, and not worry about what is the correct tense to use or which words are too colloquial for scientific writing. Even though it is placed in the front of the thesis, this is the last thing I am writing. As I was writing the other parts of my thesis, something felt missing but I could not put my finger on it. Now I know - it was the idea that I had to use rigid formality to summarize my entire PhD purely based on the 'successful bits' of my research, when it was so much more than that.

And what is success anyway? I am sure everyone has their own definitions, but Mama taught me to follow Og Mandino's advice:

"I will greet this day with love in my heart. For this is the greatest secret of success in all ventures."

Sure, working out mathematical proofs or getting a complicated model to yield results that I wanted gave me a great sense of accomplishment, but when I think of the triumphant moments as a PhD student, none of them were entirely dependent on my research. It has always been and will hopefully always be about the love I give and the love I receive.

Academic Acknowledgements

I would like to express my sincere gratitude towards Malcolm McCulloch and Thomas Morstyn, for one could not ask for better supervisors.

In January 2016, it was pure luck when I stumbled upon the website of the Energy and Power Group three days before the application deadline. Upon sending in my application, I emailed Malcolm to state my case. Being a person that is perpetually busy, he did not reply, but being a person that likes to take his chances, he gave me an offer. "What do you do when people take a leap of faith by giving you a chance?" My Mama asked me once. "I will do my best to prove them right" was my answer. I will be forever grateful for Malcolm's trust because my life path would have been very different otherwise, and I hope I indeed proved to Malcolm that he made the right decision.

If Malcolm was the one that helped me envision the big picture of my research, then Thomas was the one that handed me the paintbrush and guided my every stroke in painting it. Thomas's research ability is unparalleled in my field among our peers, and for every question I had, he offered three possible solutions, and of course ten more

questions. He was an incredibly reliable source of guidance, and his comments with lightning quick turnovers always helped elevate my work, including this thesis, to a higher level. Although Thomas set a high bar for academic achievements, he was the most understanding academic mentor. Having recently been through the PhD journey himself, he was able to empathize with my struggles as a PhD student and offered me valuable advice that saved me from many frustrated moments.

I am very proud to be in a very diverse research group, with a 50/50 gender balance and people of various backgrounds from six different continents. I am grateful to everyone for making it an inclusive and fun research environment. I am particularly grateful to Matt and Constance for their company at conferences and local pubs, and for indulging me with my scattered mind and occasional political rants, and to Miriam and Alycia, for being great listeners and allies in addressing equality issues in and outside of our research community.

I would also like to thank the Zhejiang University Yongping foundation for giving me a student loan, as well as the China Oxford Scholarship Fund for offering me a scholarship. I would not have been able to complete my overseas studies without their financial support.

I am immensely grateful to Shaz for helping to proofread my thesis.

Personal Acknowledgements

It would be far from the truth to attribute my research achievements only to the ones acknowledged above, and it is very important to me to express my gratitude to some other people near and dear to my heart, without whom this thesis would have never materialized.

To my longest friend in Oxford, Mara, I am in awe of your strength despite the physical and mental challenges you go through day in and day out. Thank you for inspiring me to be humble and to learn, to be brave and to fight for social equality and justice, and at the same time to be self-loving and to take breaks. Thank you for teaching me so much about disabilities, feminism, and queer culture.

To my fellow peer supporters and my partners in crime, Shaz, Jess, Anna, and Sam, it is no exaggeration to say that you transformed my Oxford experience, and because of you all, I have so many positive memories about Oxford I can carry with me when I leave. Shaz, thank you for constantly going out of your way to care for people, and for being a loud and proud advocate for a better world. Jess, thank you for always listening to me without an ounce of judgement, and for sharing your always so reflective and critical thoughts. Anna, thank you for your sisterly love, and for showing me the strength of vulnerability by wearing your heart on your sleeve. Sam, thank you for always being a source of joy even on the most gloomy days, and for never mincing your words to call out social injustice.

I am grateful for Mansfield College for giving me a place I call home for the past four years. Helen, thank you for being an incredible principal who puts the students first and stands for equality and diversity. John, thank you for instilling music to my

life and bearing with my singing in the College Choir for four years. Alan, Lynne and Chris, thank you for trusting me with the Junior Dean job, giving me a chance to foster a closer relationship with the College.

Filip, thank you for being the best flatmate for 2 good Oxford years and for your generous gifts of books, and more importantly, your time. Bailey and Leonie, thank you for always being so close even after you moved so far away and for your words of encouragement that pushed me through many mental barriers. Dave and Ross, thank you for being my comrades through the ups and downs on this DPhil journey. Yuan and Venky, thank you for the many conversations in our native language about problems that only we understand. Benedetta and Enrike, thank you for forcing me to look outside of my comfortable social bubble so I could see the various shades of humanity. Elizabeth, thank you for being the kindest friend to support me through the last stretch of this DPhil journey. Marianna, thank you for being the best colleague in our MCR and with the organization of the Diversity and Equality Conference (OUDEC 2019). Azamati, thank you for being the most humorous, optimistic and hard-working friend, and for using your story to inspire so many others.

Among the various student initiatives I was involved in, OU Coffee Ambassadors and the BAME Drama Society were the closest to my heart. Leonie, thank you for coming up with the brilliant idea to foster human connection within the Oxford community, and for always being more of a friend than just a co-founder. Fran and Taiwo, thank you for putting your trust in me to help your effort in diversifying Oxford drama even before I was involved in any production.

Last but not least, I would like to thank my family and friends both in China and in the US. Mama and Papa, thank you for teaching me how to love and forgive, which has laid the core foundation of who I am as a person whatever job title I am wearing. Mother, thank you for giving me life and generously providing the resources I needed to get to where I am today. Carishma and Dimitry, thank you for being the most supportive and understanding friends, for always being in my corner comforting me and celebrating me, and for recently making me the proudest uncle. Christa and Tim, thank you for keeping our wonderful community together and for reminding me to put down work to enjoy life from time to time. Hara and Roy, thank you for always challenging me to reflect and work hard. Wouter, thank you for teaching me what true love is about. Trish, thank you for being my companion on our journey to understand our past and emotions. Xie, Ronghui, Yuyan, and Shize, thank you for your 15 years of friendship and for being just a message away. Peng and Haoshu, thank you for your silliness and for never failing to cheer me up no matter the circumstances.

Thank you, Oxford, for the incredible four years. I vow to carry the lessons you have taught me for the rest of my life.

Those who stand for nothing fall for everything.

— Alexander Hamilton

Abstract

The advances in distributed energy resources (DER) propelled power networks worldwide to transition from a unidirectional model to a bidirectional one, under which uncertain and intermittent renewable generation cause load balancing issues. Energy storage as flexible resources have the potential to mitigate the stress on the network, and local energy markets emerged in recent years to encourage efficient management of these resources. However, the formulation of a market scheme that incentivizes such operation by offering profit for all the participants remains an unresolved challenge. This thesis proposes the novel application of cooperative game theory to coalitional energy management, which extends existing cooperative theoretic frameworks to the energy storage operation. The proposed scheme is then proven to have a non-empty core, ensuring the existence of a payoff allocation that incentivizes participation in the grand coalition of all participants.

This thesis focuses on two main payoff allocations: the Shapley value and the nucleolus.

The Shapley value is a commonly used payoff allocation in cooperative game theory, but is highly computationally complex. This thesis proposes a novel multi-step sampling method based on optimal stratified sampling for estimating the Shapley value, and demonstrates its effectiveness to improve the scalability of the proposed scheme. It is shown, however, that the Shapley value is not necessarily stabilizing, meaning that given the payoffs based on the Shapley value, some participants may choose to defect from the grand coalition and form smaller coalitions for higher profits.

The nucleolus, albeit a less commonly used payoff allocation in cooperative game theory, is proven to be stabilizing, preventing defection from the grand coalition. Similarly, the computation time of the nucleolus increases exponentially with the number of participants, strictly limiting the size of this scheme. This thesis proposes a new strategy which incorporates clustering techniques to estimate the nucleolus at reduced computation times, where a novel marginal contribution profile is constructed as the clustering features. While maintaining a high estimation accuracy, it is shown to easily scale up the scheme from less than 15 participants to over 100, a well applicable number of load buses in a real distribution feeder.

Contents

List of Figures **ix**

List of Tables **xii**

List of Abbreviations **xiii**

Nomenclature **xiv**

1 Introduction **1**
1.1 Context and Motivation 2
1.1.1 Distributed Renewable Energy Generation 4
1.1.2 Energy Network Flexibility and Storage 7
1.1.3 Energy Market Transition 8
1.1.4 Game Theory 10
1.1.5 Discussion 11
1.2 Original Contributions 12
1.2.1 Contributing Publications 13
1.2.2 Thesis Scope 13
1.3 Thesis Structure 16

2 Literature Review **19**
2.1 Prosumer-Centric Local Energy Markets 21
2.1.1 System-Focused Market Schemes 23
2.1.2 Individual-Focused Market Schemes 24
2.2 Game Theory for Prosumer-Centric Energy Market Design 28
2.2.1 Non-cooperative Game Theoretic Frameworks 28
2.2.2 Cooperative Game Theoretic Frameworks 30
2.2.2.1 Computational Complexity 32
2.3 Customer Load Profile Clustering 33
2.3.1 Clustering Features 34
2.3.2 Clustering Techniques 35
2.4 Gap Analysis 37

3 Cooperative Game Theoretic Energy Market **40**
3.1 Cooperative Game Framework 42
3.2 Coalitional Energy Management using Cooperative Game Theory . . . 45
3.2.1 Coalitional Energy Cost 46
3.2.2 Optimal Energy Storage Operation 47
3.2.3 Value of Prosumer Coalitions 48
3.2.4 Payoff Allocations and Imputations 49
3.3 Balancedness of the Coalitional Energy Management Scheme 50
3.3.1 Coalition Matrix Operation Definitions 50
3.3.2 Coalitional Energy Management in Matrix Format 51
3.3.2.1 Proof of a Nonempty Core 54
3.4 Model Implementation 56
3.4.1 Energy Management Optimization 57
3.4.2 Computation of Shapley Value and Nucleolus 58
3.5 Case Studies 60
3.5.1 Load Balancing and Financial Reward 60
3.5.2 Nucleolus vs. Shapley Value 63
3.5.3 The Impact of Uncertainty 65
3.6 Concluding remarks 68

4 Sampling for Estimating the Shapley Value - a Fair Payoff Allocation **70**
4.1 Shapley Value of the Coalitional Energy Management Scheme 71
4.1.1 Shapley Value Computation and Properties 72
4.1.2 Shapley Value Computational Complexity 74
4.2 Shapley Value Estimation with Coalition Sampling 75
4.2.1 Stratified Random Sampling 76
4.2.2 Coalitional Stratified Sampling with Optimal Sample Allocation 77
4.2.3 Multi-Step Sampling 80
4.3 Case Studies 82
4.3.1 Validation of Sampling-Based Shapley Estimation 82
4.3.2 Single Prosumer Payoff Analysis 85
4.3.3 Sampling-Based Shapley Estimation for Large Schemes 88
4.4 Concluding remarks 90

5 Clustering for Estimating the Nucleolus - a Stabilizing Payoff Allocation **92**
5.1 Nucleolus of the Coalitional Energy Management Scheme 94
5.1.1 Nucleolus: A Stabilizing Imputation of the Coalitional Energy Management Scheme 94
5.1.2 Nucleolus Computational Complexity 98
5.2 Prosumer Clustering Techniques 99

5.2.1 K-Means Clustering 99
5.2.2 Hierarchical Clustering 102
5.2.3 Gaussian Mixture Model 103
5.2.4 Prosumer Clustering Features and Principal Component Analysis 105
5.3 Nucleolus Estimation with Prosumer Clustering 111
5.3.1 Simple K-Means for Computing the Nucleolus for Clustered Players 111
5.3.2 Full Nucleolus Estimation with Clustering 115
5.3.3 Evaluating Nucleolus Estimation with Stratified Random Sampling 117
5.3.3.1 Stratified Random Sampling for Direct Nucleolus Estimation 118
5.3.3.2 Stratified Random Sampling for Evaluating Nucleolus Estimation 121
5.4 Case Studies 122
5.4.1 Validation of Nucleolus Estimation 123
5.4.2 Trade-Off between Estimation Accuracy and Computation Time 125
5.4.3 Comparison of Clustering Techniques 128
5.4.4 Scaling up with Relaxed Computation Time 130
5.5 Concluding remarks 131

6 Conclusions **133**
6.1 Concluding Discussion 133
6.2 Future Work 139

Appendices

A Additional Proofs **143**
A.1 Proof of Lemma 3.1 143
A.2 Proof of Lemma 3.2 144
A.3 Proof of Lemma 3.3 145
A.3.1 Proof of Positivity Homogeneity 145
A.3.2 Proof of Subadditivity 146
A.4 Proof of Theorem 3.4 147

B Case Studies Inputs **149**

Bibliography **151**

List of Figures

1.1 Renewable electricity generation by source (non-combustible), United Kingdom 1990-2018. 6
1.2 Energy networks with and without P2P trading. 9

2.1 Three main components of a prosumer-centric energy market platform. . 22
2.2 Categorization of prosumer-centric local energy markets. 23

3.1 10-prosumer grand coalition 24-hour load profile. 61
3.2 10-prosumer coalitional energy costs and cost savings over 24 hours by end user. 62
3.3 Comparison of the individual payoff allocations (nucleolus vs. Shapley value) for a 10-prosumer coalitional energy management scheme. . . . 63
3.4 Largest excesses by number of prosumers – nucleolus vs. Shapley. . . . 64
3.5 Comparison of 24-hour nucleolus excesses and Shapley excesses for all coalitions in a 10-prosumer coalitional energy management scheme . . 65
3.6 Comparison of predicted energy profiles and the actual energy profiles with ES operation. 66
3.7 Comparison of the predicted energy cost and the actual energy cost for the grand coalition. 67
3.8 Changes in energy cost savings for the grand coalition as percentages of the predicted values due to prediction errors. 68

4.1 16 Prosumer 24-hour load profiles. 73
4.2 Standard error comparison using the multi-step sampling technique: random sampling vs. stratified sampling (N=14). 83
4.3 Comparison of the true Shapley value with the estimated Shapley values using multi-step stratified sampling (N=14). 84
4.4 Energy profiles of a prosumer with both PV and ES in a 14-prosumer coalitional energy management scheme. 86
4.5 The impact of stochastic errors on the grand coalition energy cost savings. 87
4.6 Estimated Shapley value for a 50-prosumer scheme by DER ownership type with two different DER adoption rates (20% vs. 50%). 88

4.7 Estimated Shapley value for a 50-prosumer scheme with varying PV and ES adoption rates. 89

5.1 20 Prosumer 24-hour load profiles 96
5.2 100 Prosumer 24-hour load profile. 101
5.3 Centroid cooperative net load profiles of 6 clusters through hierarchical clustering of 100 prosumers with different linkage criteria 103
5.4 Centroid cooperative net load profiles of 6 clusters through GMM clustering of 100 prosumers with different covariance matrices 105
5.5 GMM clustering of 100 prosumers' marginal contribution profiles . . . 107
5.6 Visualization of 3 principal components of the K-means clustering results of 100 prosumers' cooperative energy profiles ($K = 6$) 108
5.7 Visualization of 2 principal components of the hierarchical clustering results of 100 prosumers' cooperative energy profiles ($K = 6$). 109
5.8 Visualization of 2 principal components of the GMM clustering results of 100 prosumers' cooperative energy profiles ($K = 6$). 109
5.9 Visualization of 2 principal components of the GMM clustering results of 100 prosumers' marginal contribution profiles ($K = 6$). 110
5.10 Total Euclidean distances for 1000 K-means (K=8) clustering runs on 50 prosumers' cooperative energy profiles. 111
5.11 Energy costs (EC) and imputation for clustered players. 114
5.12 Comparison of the estimated nucleolus with simple K-mean clustering and the true nucleolus. 115
5.13 Estimation performance verification of the stratified random sampling method for a 14-prosumer coalitional energy management scheme using the full nucleolus as a benchmark: varying stratum sample sizes ($m = 1, 2, 3, 4, 5$). 119
5.14 Estimation performance verification of the stratified random sampling method with smaller sample sizes ($m = 1, 2$) using $m = 3$ as a benchmark: varying the scheme size ($N = 10, 20, 30, 40, 50$). 120
5.15 Comparison of nucleolus estimation methods for a 50-prosumer coalitional energy management scheme: simple K-means clustering ($K = 8$) vs. stratified random sampling ($m = 3$). 123
5.16 Comparison of nucleolus estimation methods for a 50-prosumer coalitional energy management scheme: full estimation method with K-means clustering ($K = 5$) vs. stratified random sampling ($m = 3$). 124
5.17 Estimation performance verification of the full estimation method with K-means clustering for a 50-prosumer coalitional energy management scheme using $m = 3$ as a benchmark: varying number of clusters ($K = 3, 4, 5, 6, 7$). 125

5.18 Comparison of the nucleolus estimation time for different number of clusters selected using K-means. 126
5.19 Nucleolus estimation performance of different techniques using the stratified random sampling method ($m = 3$ as baseline). 126
5.20 Nucleolus estimation performance of different techniques using the stratified random sampling method ($m = 1$ as baseline). 127
5.21 Nucleolus estimation performance with different clustering techniques validated using the stratified random sampling based approach ($m = 1$). 129
5.22 Nucleolus estimation performance with GMM-full with marginal contribution feature using different K values, validated using the stratified random sampling based approach ($m = 1$). 130

List of Tables

1.1 Contributing Research (Published and Under Review) 14

4.1 Shapley Value Computation Time τ (s) 85

5.1 Nucleolus Computation Time (s) 120
5.2 Clustering Computation Time (s) Avg. over 10 Runs 128

List of Abbreviations

ACC . . . . . achievable cycle count

ADMM . . . . . alternating direction method of multipliers

AWC . . . . . average wear cost

BIC . . . . . Bayesian information criterion

CVI . . . . . clustering validity index

DER . . . . . distributed energy resources

DoD . . . . . depth of discharge

EM . . . . . expectation-maximization

ES . . . . . energy storage

EV . . . . . electric vehicles

FPP . . . . . federated power plant

GMM . . . . . Gaussian mixture model

ICT . . . . . information and communication technologies

LP . . . . . linear program

P2P . . . . . peer-to-peer

PCA . . . . . principal component analysis

PDF . . . . . probability distribution function

Prosumers . . . proactive-consumers with distributed energy resources that actively control their energy behaviors

SOM . . . . . self-organizing map

VPP . . . . . virtual power plant

Nomenclature

$\mathbf{0}$	Zero matrix or vector of any size
$\mathbf{1}^N$	All-ones column vector of size N: $\mathbf{1}^N = [\underbrace{1, ..., 1}_{N}]^\intercal$
$1_i^{\mathcal{S}}$	Indicator function of player i and coalition $\mathcal{S}$: $1_i^{\mathcal{S}} = \begin{cases} 1, & \text{if } i \in \mathcal{S} \\ 0, & \text{if } i \notin \mathcal{S} \end{cases}$
$2^{\mathcal{N}}$	Set of all coalitions within grand coalition $\mathcal{N}$: $\mathcal{S} \in 2^{\mathcal{N}}, \forall \mathcal{S} \subseteq \mathcal{N}$
a, a'	Indices for the nucleolus computation iterations
$\mathfrak{B} \in \mathbb{R}^{R \times N}$ . . .	ES operation variable matrix for the cooperative game $(\mathcal{N}, v)$: $\mathfrak{B} = [\mathbf{b}_1, ..., \mathbf{b}_N]$, (*kWh*)
$\overline{\mathbf{B}} \in \mathbb{R}^{R \times N}$ . . .	ES system charge limit matrix for the cooperative game $(\mathcal{N}, v)$ over time span of Δt: $\overline{\mathbf{B}}^{\mathcal{S}} = [\overline{b}_1 \mathbf{1}^R 1_1^{\mathcal{S}}, \overline{b}_2 \mathbf{1}^R 1_2^{\mathcal{S}}, ..., \overline{b}_N \mathbf{1}^R 1_N^{\mathcal{S}}]$, (*kWh*)
$\underline{\mathbf{B}} \in \mathbb{R}^{R \times N}$ . . .	ES system charge limit matrix for the cooperative game $(\mathcal{N}, v)$ over time span of Δt: $\underline{\mathbf{B}}^{\mathcal{S}} = [\underline{b}_1 \mathbf{1}^R 1_1^{\mathcal{S}}, \underline{b}_2 \mathbf{1}^R 1_2^{\mathcal{S}}, ..., \underline{b}_N \mathbf{1}^R 1_N^{\mathcal{S}}]$, (*kWh*)
$\mathbf{b}_i \in \mathbb{R}^{R \times 1}$ . . .	Prosumer i's ES energy charge (positive) and discharge (negative) variables, column vector, (*kWh*)
$\overline{b}_i$	ES system charge limit over time span of Δt for prosumer i: $\overline{b}_i \geq 0$, (*kWh*)
$\underline{b}_i$	ES system discharge limit over the time span of Δt for prosumer i: $\underline{b}_i \leq 0$, (*kWh*)
b_{it}	Prosumer i's ES system's charge (positive) and discharge (negative) energy variables at timestep t, (*kWh*)
$\mathbf{b}^{\mathcal{S}}$	ES operation variables for coalition $\mathcal{S}$: $\mathbf{b}^{\mathcal{S}} = [b_{it}], \forall i \in \mathcal{S}, \forall t \in [1, R]$, (*kWh*)
$\mathbf{b}^{\mathcal{S}^+}$	Positive part of the ES operation variables of all the prosumers in coalition $\mathcal{S}$: $\mathbf{b}^{\mathcal{S}^+} = [b_{it}^+], \forall i \in \mathcal{S}, \forall t \in [1, R]$ where $b_{it}^+ = \max\{0, b_{it}\}$, (*kWh*)
$\mathbf{b}^{\mathcal{S}^-}$	Non-positive part of the ES operation variables of all the prosumers in coalition $\mathcal{S}$: $\mathbf{b}^{\mathcal{S}^-} = [b_{it}^-], \forall i \in \mathcal{S}, \forall t \in [1, R]$ where $b_{it}^- = \min\{0, b_{it}\}$, (*kWh*)

$\mathbf{b}^{\mathcal{T}*}$ . . . . . Optimal ES operation for coalition $\mathcal{T}$: $\mathbf{b}^{\mathcal{T}*} = \arg\min_{\mathbf{b}^{\mathcal{T}}} \sum_{t=1}^{R} F_t^{\mathcal{T}}(\mathbf{b}^{\mathcal{T}})$

$\mathfrak{B}^{\langle\mathcal{Y}\rangle*}$ . . . . . Optimal decision variable matrix to $\theta[\mathcal{Y}]$: $C_{\mathcal{Y}}(\mathfrak{B}^{\langle\mathcal{Y}\rangle*}) = \theta[\mathcal{Y}]$

$\mathcal{C}$ . . . . . Core of a cooperative game $(\mathcal{N}, v)$, which is the set of all the stabilizing imputations: $\mathcal{C} = \{\mathbf{x} \in \mathcal{I} \mid v(\mathcal{S}) \leq \sum_{i\in\mathcal{S}} x_i, \forall \mathcal{S} \in 2^{\mathcal{N}}\}$

$\mathbf{c}_j \in \mathbb{R}^R$ . . . . . Cluster cl_j's centroid load profile: $\mathbf{c}_j = [c_{jt}], j \in \mathcal{K}, t \in [1, ..., R]$

cl_j . . . . . A cluster, or clustered player indexed by j: $cl_j \in cl_{\mathcal{K}}, j \in \mathcal{K}$

$cl_{\mathcal{K}}$ . . . . . Set of all clusters, or grand coalition of all clustered players, when prosumers are grouped in to K clusters: $cl_{\mathcal{K}} := \{cl_1, cl_2, ..., cl_K\}$

$cl_{\mathcal{U}}$ . . . . . A coalition of clustered players: for $\mathcal{U} \subseteq \mathcal{K}$, $cl_{\mathcal{U}} := \{cl_j | \forall j \in \mathcal{U}\}$, and $cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}$

$C_{\mathcal{S}}(\mathfrak{B})$ . . . . . Coalitional energy cost for coalition $\mathcal{S}$, expressed in the coalition matrix format as a function of the ES operation variable matrix $\mathfrak{B}$: $C_{\mathcal{S}}(\mathfrak{B}) = F(\mathfrak{B}, \mathbf{Q}^{\mathcal{S}})$

D . . . . . Function of the ES operation variable matrix $\mathfrak{B}$ as an indicator of whether $\mathfrak{B}$ meets the ES operation constraints: $D(\mathfrak{B}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}}) = \begin{cases} 1, & \text{ES operation constraints (power, energy, cycle) are all met.} \\ 0, & \text{otherwise.} \end{cases}$

$d(\mathbf{f}_i, \mathbf{c}_j)$ . . . . . Euclidean distance between player i's load profile and cluster cl_j's centroid load profile: $d(\mathbf{f}_i, \mathbf{c}_j) = \sqrt{\sum_t^R (f_{it} - c_{jt})^2}$

$\mathbf{E} \in \mathbb{R}^{R\times N}$ . . . ES system energy capacity matrix for the cooperative game $(\mathcal{N}, v)$: $\mathbf{E}^{\mathcal{S}} = [e_1 \mathbf{1}^R 1_1^{\mathcal{S}}, e_2 \mathbf{1}^R 1_2^{\mathcal{S}}, ..., e_N \mathbf{1}^R 1_N^{\mathcal{S}}]$, (*kWh*)

e_i . . . . . ES system energy capacity for prosumer i: $e_i \geq 0$, (*kWh*)

F . . . . . Coalitional energy cost used in the coalition matrix format, a function of the net energy load and ES operation variable matrices $\mathbf{Q}, \mathfrak{B} \in \mathbb{R}^{K\times N}$: $F(\mathfrak{B}, \mathbf{Q}) = \mathbf{r}^{im}[(\mathfrak{B} + \mathbf{Q})\mathbf{1}^N]^+ + \mathbf{r}^{ex}[(\mathfrak{B} + \mathbf{Q})\mathbf{1}^N]^-$, (£)

$\mathbf{f}_i \in \mathbb{R}^R$ . . . . . Player i's load profile: $\mathbf{f}_i = [f_{it}], i \in \mathcal{N}, t \in [1, ..., R]$

$F_t^{\mathcal{S}}$ . . . . . Coalitional energy cost for coalition $\mathcal{S}$ at timestep t, a function of $\mathbf{b}^{\mathcal{S}}$: $F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}}) = \sum_{i\in\mathcal{S}} \left\{ r_t^{im}[q_{it} + b_{it}]_+ + r_t^{ex}[q_{it} + b_{it}]_- \right\}$, (£)

$\boldsymbol{g} \in \mathbb{R}^N$ . . . . . Clustering assignment: $\boldsymbol{g} : g_i = j \mid i \in cl_j, \forall i \in \mathcal{N}$

$G^R(\mathcal{S})$ . . . . . Fuction to minimize the coalitional energy cost for coalition $\mathcal{S}$ by varying $\mathbf{b}^{\mathcal{S}}$: $G^R(\mathcal{S}) = \min_{\mathbf{b}^{\mathcal{S}}} \sum_{t=1}^{R} F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}})$

h . . . . . Index for the coalitions in $2^{\mathcal{N}}$: $h = 1, 2, ..., 2^N$, e.g., $\mathcal{S}_h$ is the hth coalition, λ_h is the scaling factor for $\mathcal{S}_h$

$\boldsymbol{h}$	Updated count of total number of samples taken for each coalitional stratum in the coalitional stratified sampling algorithm: $\boldsymbol{h} = \{h_{il}\}, \forall i, l \in [1, N]$
H_{il}	A set of coalitions of size $\lvert H_{il}\rvert$ randomly sampled with replacement from coalitional stratum Q_{il}
H_{il}^A, H_{il}^B	Coalition sample sets in the coalitional stratified sampling algorithm for Stage A and B respectively
$h_{il}^A, h_{il}^B, h_{il}^{tot}$	Stage A, B, and total number of samples taken for each coalitional stratum in the coalitional stratified sampling algorithm
h^*	Minimum multi-step sample size required to achieve the estimation criterion σ^*
i	Player/prosumer index: $i \in \mathcal{N}$
$\mathcal{I}$	Set of all imputations of a cooperative game $(\mathcal{N}, v)$: $\mathcal{I} = \{\mathbf{x} \mid \sum_{i\in\mathcal{N}} x_i = v(\mathcal{N}),\ x_i \geq v(\{i\}), \forall i \in \mathcal{N}\}$
j	Cluster index: $j \in \mathcal{K}$
K	Number of prosumer clusters.
$\mathcal{K}$	Set of all cluster indices: $\mathcal{K} = \{1, ..., K\}$
LP_a	The ath iteration of the series of linear problems solved for computing the nucleolus
$\mathbf{L}^+$	Positive part of the net energies of all the prosumers: $\mathbf{L}^+ = [L_{it}^+], \forall i \in \mathcal{S}, \forall t \in [1, R]$, where $L_{it}^+ = \max\{0, b_{it}^+ + b_{it}^- + q_{it}\}$
$\mathbf{L}^-$	Non-positive part of the net energies of all the prosumers: $\mathbf{L}^- = [L_{it}^-], \forall i \in \mathcal{S}, \forall t \in [1, R]$, where $L_{it}^- = \min\{0, b_{it}^+ + b_{it}^- + q_{it}\}$
m	Stratum sample size, or the number of permutation samples in P_{il}: $m = \lvert M_{il}\rvert$
M_{il}	A set of permutations of player set $\mathcal{N}$ of size $\lvert M_{il}\rvert$ randomly sampled with replacement from stratum P_{il}
$\mathfrak{M}_m$	Set of coalitions derived from the permutation sample set M_{il} with a stratum sample size m: $\mathfrak{M}_m := \{\mathcal{T} \mid \mathcal{T} = Pre^i(O) \cup i, \forall O \in M_{il}, \forall i, l \in [1, N]\}$
N	Number of players in the grand coalition.
$\mathcal{N}$	Grand coalition: $\mathcal{N} = \{1, ..., N\}$
$(\mathcal{N}, v)$	The cooperative game with $\mathcal{N}$ as the grand coalition, and v as the value function

O . . . . . A permutation of player set $\mathcal{N}$ that assigns player $O(l)$ to position l

$\mathfrak{O}_{\mathcal{Y}}$ . . . . . Set of ES operation variable matrices that satisfy all the ES constraints given the coalition combination of $\mathcal{Y}$: $\mathfrak{O}_{\mathcal{Y}} = \{\mathfrak{B} \in \mathbb{R}^{R\times N} \mid D(\mathfrak{B}, \mathbf{E}^{\mathcal{Y}}, \underline{\mathbf{B}}^{\mathcal{Y}}, \overline{\mathbf{B}}^{\mathcal{Y}}) = 1\}$

$\mathbf{p}$ . . . . . Grand coalition cooperative energy profiles: $\mathbf{p} = \{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N\}$

$\mathbf{p}_i \in \mathbb{R}^R$ . . . . . Player i's grand coalition cooperative energy profile: $\mathbf{p}_i = \mathbf{q}_i + \mathbf{b}_i^{\mathcal{N}*}, \forall i \in \mathcal{N}, t \in [1, ..., R]$

P_{il} . . . . . A stratum, or a stratified set of player permutations, containing every permutation $O \in \pi(\mathcal{N})$, in which player i is in position l: $P_{il} := \{O \in \pi(\mathcal{N}) \mid O(l) = i, \forall i, l \in [1, N]\}$. Therefore, P_{il}

$Pre^i(O)$ . . . . Set of predecessors of the player i for a given $O \in \pi(\mathcal{N})$: if $i = O(l)$, $Pre^i(O) = \{O(1), O(2), ..., O(l-1)\}$

$\mathbf{Q} \in \mathbb{R}^{R\times N}$ . . . Net energy load matrix (without ES operation) for the cooperative game $(\mathcal{N}, v)$: $\mathbf{Q}^{\mathcal{S}} = [\mathbf{q}_1 1_1^{\mathcal{S}}, \mathbf{q}_2 1_2^{\mathcal{S}}, ..., \mathbf{q}_N 1_N^{\mathcal{S}}]$, (*kWh*)

$\mathbf{q}_i \in \mathbb{R}^{R\times 1}$ . . . Prosumer i's net energy consumption (positive) and generation (negative) without ES, column vector, (*kWh*)

Q_{il} . . . . . A coalitional stratum, or a stratified set of coalitions of size $l-1$ that does not contain player i: $Q_{il} := \{\mathcal{S} \subseteq \mathcal{N} \mid i \notin \mathcal{S}, |\mathcal{S}| = l-1, \forall i, l \in [1, N]\}$

q_{it} . . . . . Prosumer i's net energy consumption (positive) and generation (negative) without ES at timestep t, (*kWh*)

$Q(cl_{\mathcal{U}})$ . . . . . A functional to convert a coalition of clustered players into a coalition of prosumers: $Q(cl_{\mathcal{U}}) = \{i \mid cl_{g_i} \in cl_{\mathcal{U}}\}$

R . . . . . Total number of timesteps considered for the cooperative energy management game

$\mathbf{r}^{ex} \in \mathbb{R}^{1\times R}$ . . . Electricity sell (export) price, row vector, (£/*kWh*)

r_t^{ex} . . . . . Electricity sell (export) price at time t

$\mathbf{r}^{im} \in \mathbb{R}^{1\times R}$ . . . Electricity buy (import) price, row vector, (£/*kWh*)

r_t^{im} . . . . . Electricity buy (import) price at time t

$\boldsymbol{s}$ . . . . . Stratum interim parameters in the coalitional stratified sampling algorithm to facilitate the computation of the sample variances $\overline{\sigma}_{il}^2$ and the mean marginal contributions $\overline{\phi}_{il}$ for each coalitional stratum: $\boldsymbol{s} = \{s_{il}\}, \forall i, l \in [1, N]$

$\mathcal{S}$ . . . . . A coalition of players

$\mathfrak{S}_a$ . . . . . Set of coalitions that have an excess value of ε_a with the nucleolus as the payoff allocation, obtained in the ath iteration of the nucleolus computation: $\mathfrak{S}_a := \{\mathcal{S} \mid v(\mathcal{S}) - \sum_{\forall i \in \mathcal{S}} x_i = \varepsilon_a, \forall \mathbf{x}_a\}$

$\mathfrak{S}_i$ . . . . . Set of prosumer coalitions, each of which is a union of a prosumer coalition from $\mathbf{U}_i$ and $\{i\}$: $\mathfrak{S}_i = \{\mathcal{T} \mid \mathcal{T} = Q(cl_{\mathcal{U}}) \cup \{i\}, \forall cl_{\mathcal{U}} \in \mathbf{U}_i\}$

$\overline{\mathbf{SoC}} \in \mathbb{R}^{N \times N}$. ES system state of charge upper limit matrix for the cooperative game $(\mathcal{N}, v)$: $\overline{\mathbf{SoC}} = \text{diag}(\overline{SoC}_1, \overline{SoC}_2, ..., \overline{SoC}_N)$

$\underline{\mathbf{SoC}} \in \mathbb{R}^{N \times N}$. ES system state of charge lower limit matrix for the cooperative game $(\mathcal{N}, v)$: $\underline{\mathbf{SoC}} = \text{diag}(\underline{SoC}_1, \underline{SoC}_2, ..., \underline{SoC}_N)$

$\overline{SoC}_i$ . . . . . ES system state of charge upper limit for prosumer i: $\overline{SoC}_i \in [0, 1]$

$\underline{SoC}_i$ . . . . . ES system state of charge lower limit for prosumer i: $\underline{SoC}_i \in [0, 1]$

$\mathbf{SoC^0} \in \mathbb{R}^{N \times N}$. ES system initial state of charge matrix for the cooperative game $(\mathcal{N}, v)$: $\mathbf{SoC^0} = \text{diag}(SoC_1^0, SoC_2^0, ..., SoC_N^0)$

SoC_i^0 . . . . . ES system initial state of charge for prosumer i: $SoC_i^0 \in [0, 1]$

t . . . . . Timestep index, $t = 1, ..., R$

$\mathcal{T}$ . . . . . A coalition of players

$\mathbf{T}_a$ . . . . . Set of coalitions whose excesses are minimized in LP_a and fixed for all subsequent iterations in Algorithm 3

$\mathfrak{T}_a$ . . . . . Set of coalitions whose excesses have been lexicographically minimized in all previous iterations up until the end of Iteration a and fixed for all subsequent iterations in Algorithm 3

$\mathfrak{T}^{cl_{\mathcal{K}}}$ . . . . . Set of prosumer coalition samples for excess minimization to estimate the nucleolus: $\mathfrak{T}^{cl_{\mathcal{K}}} = \mathfrak{S} \cup \mathfrak{U}$

$\mathfrak{T}^{\mathcal{N}}$ . . . . . Set of all the possible coalitions in an N-player cooperative energy management scheme: $\mathfrak{T}^{\mathcal{N}} = \{\mathcal{T} | \mathcal{T} \subseteq \mathcal{N}\}$

$\mathfrak{T}^*$ . . . . . Set of coalitions whose excesses are minimized lexicographically in Algorithm 3

U, u . . . . . Total number of steps in the multi-step sampling strategy, and the step counter: $u = 1, ..., U$

$\mathfrak{U}$ . . . . . Set of all prosumer coalitions that can be derived from coalitions of clustered players: $\mathfrak{U} = \{\mathcal{T} \mid \mathcal{T} = Q(cl_{\mathcal{U}}), \forall cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}\}$

$\mathbf{U}_i$ . . . . . Set of coalitions of *clustered players* that do not contain player i: $\mathbf{U}_i = \{cl_{\mathcal{U}} \mid \forall cl_{\mathcal{U}} \subseteq (cl_{\mathcal{K}} \setminus \{cl_{g_i}\})\}$

$v(\mathcal{S})$	Value or characteristic function of coalition $\mathcal{S}$
$w(\mathcal{T})$	Sum of the payoff allocations of all the players in coalition $\mathcal{T}$
$\mathbf{x} \in \mathbb{R}^N$	A payoff allocation of a cooperative game $(\mathcal{N}, v)$
$\mathbf{x}_a$	An optimal payoff allocation obtained in $LP_{a'}$, not unique until the solution converges to the nucleolus $\boldsymbol{\nu}$
x_i	Payoff allocated to player i under the payoff allocation $\mathbf{x}$
$\mathbf{x}^*$	Payoff allocation output from Algorithm 3 that lexicographically minimizes the excesses of $\mathfrak{T}^*$
$\mathcal{Y}$	Any first degree polynomial of all the coalitions with non-negative coefficients: $\mathcal{Y} \in \{\sum_{h=1}^{2^N} \lambda_h \mathcal{S}_h \mid \mathcal{S}_h \in 2^{\mathcal{N}},\ \lambda_h \geq 0\}$
$\mathbf{Z} \in \mathbb{R}^{R \times N}$ . . .	Any $R \times N$ matrix: $\mathbf{Z} = [\mathbf{z}_1, ..., \mathbf{z}_N]$
$\mathbf{z}_i \in \mathbb{R}^{R \times 1}$	Any column vector with R elements

---

γ	A balanced map that satisfies: $\sum_{\mathcal{S} \in 2^{\mathcal{N}}} \gamma(\mathcal{S}) 1_i^{\mathcal{S}} = 1, \forall i \in \mathcal{N}$
$\Delta\mathbf{F}$	Grand coalition marginal contribution profiles: $\Delta\mathbf{F} = \{\Delta\mathbf{F}_1, ..., \Delta\mathbf{F}_N\}$
$\Delta\mathbf{F}_i \in \mathbb{R}^R$ . . .	Player i's grand coalition marginal contribution profile, or cost savings profile from having i join the existing coalition $\mathcal{N} \setminus \{i\}$: $\Delta\mathbf{F}_i = \{\Delta F_{i1}, ..., \Delta F_{iR}\}$
ΔF_{it}	Player i's grand coalition marginal contribution at timestep t: $\Delta F_{it} = F_t^{\mathcal{N}\setminus\{i\}}(\mathbf{b}^{\mathcal{N}\setminus\{i\}*}) + F_t^{\{i\}}(\mathbf{b}^{\{i\}*}) - F_t^{\mathcal{N}}(\mathbf{b}^{\mathcal{N}*}), t \in [1, ..., R]$
$\delta(O)_i$	Player i's marginal contribution to a coalition made up of players that are predecessors of the player i for a given $O \in \pi(\mathcal{N})$: $\delta(O)_i = v(Pre^i(O) \cup i) - v(Pre^i(O))$
$\Delta(\mathcal{S})_i$	Player i's marginal contribution to coalition $\mathcal{S}$: $\Delta(\mathcal{S})_i = v(\mathcal{S} \cup i) - v(\mathcal{S})$
ε_a	The ath largest excess of a cooperative game with the nucleolus as the payoff allocation, obtained in the ath iteration of the nucleolus computation
$\boldsymbol{\varepsilon}_{\mathbf{x},m}$	Set of excesses of all the prosumer coalition samples in $\mathfrak{M}_m$ under the payoff allocation $\mathbf{x}$: $\boldsymbol{\varepsilon}_{\mathbf{x},m} = \{\varepsilon(\mathbf{x}, \mathcal{T}) \mid \forall \mathcal{T} \in \mathfrak{M}_m\}$
$\boldsymbol{\varepsilon}_{\mathbf{x},m}^{pos}$	Set of all positive excesses of the prosumer coalition samples in $\mathfrak{M}_m$ under the payoff allocation $\mathbf{x}$
$\varepsilon(\mathbf{x}, \mathcal{S})$	Excess of coalition $\mathcal{S}$ under payoff allocation $\mathbf{x}$

ζ . . . . . Convergence threshold for the evaluation of the nucleolus estimation using stratified random sampling. It represents the change in the percentage of the coalition samples that have positive excesses between iterations.

$\boldsymbol{\eta^I} \in \mathbb{R}^{N\times N}$ . . . ES system charge efficiency matrix for the cooperative game $(\mathcal{N}, v)$: $\boldsymbol{\eta^I} \in \mathbb{R}^{N\times N} = \text{diag}(\eta_1^{in}, \eta_2^{in}, ..., \eta_N^{in})$

η_i^{in} . . . . . ES system charge efficiency for prosumer i: $\eta_i^{in} \in (0, 1)$

$\boldsymbol{\eta^O} \in \mathbb{R}^{N\times N}$ . . ES system discharge efficiency matrix for the cooperative game $(\mathcal{N}, v)$: $\boldsymbol{\eta^O} = \text{diag}(1/\eta_1^{out}, 1/\eta_2^{out}, ..., 1/\eta_N^{out})$

η_i^{out} . . . . . ES system discharge efficiency for prosumer i: $\eta_i^{out} \in (0, 1)$

$\theta[\mathcal{Y}]$ . . . . . Function of the coalition combination $\mathcal{Y}$ indicating its equivalent minimum coalitional energy cost: $\theta[\mathcal{Y}] = \min_{\mathfrak{B}\in\mathfrak{D}_\mathcal{Y}} F(\mathfrak{B}, \mathbf{Q}^\mathcal{Y})$

λ . . . . . Scaling factor used in coalition matrix operations: $\lambda \geq 0$

$\boldsymbol{\Lambda}^R \in \mathbb{R}^{R\times R}$ . . . Scaling factor matrix: $\boldsymbol{\Lambda}^R = [\lambda_{m'n'}]:\ \lambda_{m'n'} = \begin{cases} 1, & m' \geq n' \\ 0, & m' < n' \end{cases}$

$\boldsymbol{\nu} \in \mathbb{R}^N$ . . . . . Nucleolus of a cooperative game $(\mathcal{N}, v)$

$\hat{\boldsymbol{\nu}} \in \mathbb{R}^N$ . . . . . Estimated nucleolus of a cooperative game $(\mathcal{N}, v)$

$\boldsymbol{\nu}_{cl}$ . . . . . Nucleolus of the cooperative game of clustered players: $\boldsymbol{\nu}_{cl} := \{\nu_{cl_j} | \forall cl_j \in cl_\mathcal{K}\}$

$\pi(\mathcal{N})$ . . . . . Set of all possible permutations of player set $\mathcal{N}$, thus $|\pi(\mathcal{N})| = N!$

σ_{il}^2 . . . . . True variance of player i's marginal contributions to stratum P_{il}: $\sigma_{il}^2 = \frac{\sum_{O\in P_{il}} (\delta(O)_i - \phi_{il})^2}{(N-1)!}$

$\overline{\sigma}_{il}^2$ . . . . . Estimated variance of player i's marginal contributions to stratum P_{il} based on coalition sample set H_{il}: $\overline{\sigma}_{il}^2 = \frac{\sum_{\mathcal{S}\in H_{il}} (\Delta(\mathcal{S})_i - \overline{\phi}_{il})^2}{|H|_{il}}$

σ^* . . . . . Estimation criterion for the multi-step sampling strategy

$\sigma(\hat{\mathbf{x}}, \mathbf{x})$ . . . . . standard error of the estimate, where $\hat{\mathbf{x}}$, $\mathbf{x}$ represent the estimated and the benchmark payoffs respectively: $\sigma(\hat{\mathbf{x}}, \mathbf{x}) = \sqrt{\sum_i^N (\hat{x}_i - x_i)^2/N}$

τ . . . . . Computation time of the Shapley value estimation algorithm

$\boldsymbol{\phi} \in \mathbb{R}^N$ . . . . . Shapley value of a cooperative game $(\mathcal{N}, v)$

$\boldsymbol{\Phi}$ . . . . . Stratum interim parameters in the coalitional stratified sampling algorithm to facilitate the computation of the sample variances $\overline{\sigma}_{il}^2$ and the mean marginal contributions $\overline{\phi}_{il}$ for each coalitional stratum: $\boldsymbol{\Phi} = \{\Phi_{il}\}, \forall i, l \in [1, N]$

$\hat{\boldsymbol{\phi}}$	Estimated Shapley value
ϕ_i	Payoff allocated to player i under the Shapley value $\boldsymbol{\phi}$
$\hat{\phi}_i$	Payoff allocated to player i under the estimated Shapley value $\hat{\boldsymbol{\phi}}$
ϕ_{il}	Player i's mean marginal contribution of stratum P_{il}: $\phi_{il} = \frac{1}{\lvert P_{il}\rvert} \sum_{O \in P_{il}} \delta(O)_i, \forall i, l \in [1, N]$
$\overline{\phi}_{il}$	Player i's mean marginal contribution of the sampled permutation set M_{il}: $\overline{\phi}_{il} = \frac{1}{\lvert M_{il}\rvert} \sum_{O \in M_{il}} \delta(O)_i, \forall i, l \in [1, N]$
$\hat{\phi}_i^{st}$	Player i's payoff allocation based on the estimated Shapley value using stratified random sampling: $\hat{\phi}_i^{st} = \sum_{l=1}^{N} \frac{1}{N} \overline{\phi}_{il}, \forall i \in [1, N]$
$\boldsymbol{\psi}(\mathbf{x}) \in \mathbb{R}^{2^N-2}$.	Excesses in respect to imputation $\mathbf{x}$ for all coalitions (excluding the grand coalition) arranged in a non-increasing order

It takes courage to love, but pain through love is the purifying fire which those who love generously know. We all know people who are so much afraid of pain that they shut themselves up like clams in a shell and, giving out nothing, receive nothing and therefore shrink until life is a mere living death.

— Eleanor Roosevelt

1 Introduction

Contents

1.1 Context and Motivation 2
1.1.1 Distributed Renewable Energy Generation 4
1.1.2 Energy Network Flexibility and Storage 7
1.1.3 Energy Market Transition 8
1.1.4 Game Theory 10
1.1.5 Discussion 11
1.2 Original Contributions 12
1.2.1 Contributing Publications 13
1.2.2 Thesis Scope 13
1.3 Thesis Structure 16

The main research question this thesis aims to answer is:

As the adoption of distributed energy resources continues to increase, how do we incentivize prosumers to manage these resources in an economically efficient and sustainable manner?

Inspired by emerging distributed energy markets, this thesis explores the use of cooperative game theory in managing energy resources owned by prosumers in the distribution network. In some literature, a prosumer means a producer-consumer, whereas this thesis adopts a broader definition: proactive-consumers with distributed energy resources (DER) that actively control their energy behaviors [1]. Through collaborative energy management of each prosumer's energy resources, it is shown that not only can

prosumers profit from this cooperative scheme, but the distribution network can benefit from the resulting reduced peak power flows and smoothed load profiles.

This introductory chapter focuses on the motivation and contributions of this thesis. Section 1.1 provides an overview of the recent technological development and current market environment for the energy sector, which explains the background of the research question. Section 1.2 summarizes the main contributions and how the thesis author's milestone publications during the course of the 4-year PhD program relate to this thesis. Section 1.3 details the structure of the whole thesis.

1.1 Context and Motivation

Since the 1990s, electricity markets worldwide have been going through privatization and deregulation in various forms [2]. This electricity market transformation, along with the emerging issues on climate change and air pollution, gave rise to the rapid development of renewable energy technologies [3]. With the further economic incentives such as feed-in tariffs, renewable energy credits, and tax credits, along with emission penalties such as carbon taxes, and carbon cap and trade set by the governments [4], privately owned renewable generation began to enter the electricity markets. Within two decades, the cost of renewable generation had become competitive against conventional fossil fuel generation in many parts of the world, and cheaper in some places even without taking into account additional governmental incentives [5].

However, problems emerged with the increase in wind and solar penetration. The most critical among them was the reliability issue inherent to the uncertain and intermittent nature of wind and solar resources. As a result, more back-up generation, quick in response but often less efficient, needed to be scheduled in order to secure the supply of energy [6]. Meanwhile, in countries that had very high solar penetrations, such as Germany, the power systems were suffering from over-generation, or generation higher than the total demand, during the day [7]. Given the system stability issues that were caused by uncertain and intermittent generation, energy policy makers in many parts of the world started reducing the financial incentives for solar and wind [8]. Some also placed regulations that either forced generators to curtail when the energy

supply was greater than the demand or financially penalized over-generation, introducing inefficiency into the energy system [9].

To complement the intermittent energy resources such as wind and solar, energy storage (ES) became the new focus in energy technological development in recent years [10]. ES has been seen as the solution to the intermittency issue created by solar and wind generation [11]. The primary function of ES is to absorb energy when generation is abundant and release it when generation is inadequate. The value of ES is multifold: it can not only allow flexibility in generation scheduling, reducing the required capacity of peak load generation, but also provide real-time balancing, enabling deeper renewable generation penetration [12].

To manage these energy resources on the distributed level, modern sensing and control technologies were developed to enable bidirectional metering of the energy flow and communication between system operators and end users [13]. The concept of a smart grid emerged out of this context. A smart grid is programmed based on objectives such as ensuring the system stability, matching supply and demand, and minimizing costs [14]. Smart meters, coupled with smart appliances and smart building systems, offer the possibility for individual end users to actively participate in the energy market. These active energy end users were defined as prosumers [1], a concept that is adopted throughout this thesis.

Recall that the main research question of the thesis is

As the adoption of distributed energy resources continues to increase, how do we incentivize prosumers to manage these resources in an economically efficient and sustainable manner?

The rest of this section explains in detail the motivation to raise this question. Section 1.1.1 introduces the issues imposed by distributed generation onto the modern power system. Section 1.1.2 discusses the potential of using energy storage to tackle these issues. Through reviewing the energy market transition in response to the increasing amount of DER, Section 1.1.3 discusses the importance of market mechanisms in managing prosumer energy behavior. Finally, Section 1.1.4 explains the rationale of using game theory in constructing such a prosumer-centric local energy market.

1.1.1 Distributed Renewable Energy Generation

Attributed to environmental, commercial, and regulatory causes, there has been a rapid growth of distributed generation over the past couple of decades [11]. Fossil fuel based distributed generators, such as diesel generators and micro-CHPs (combined heat and power plants) offer valuable ancillary services, but they further contribute to environmental issues. In the meantime, as a by-product of the rapid renewable technological development and energy regulations favoring renewable energy, the adoption of renewable generation resources on a distributed level gained tremendous momentum [15]. Small to mid-scale renewable energy generators (PV panels, solar thermal collector, micro wind turbines, etc.) have been gaining popularity among electricity end users ever since energy regulations made it possible for end users to use or sell the electricity generated locally [16].

Renewable generation adopted on the distributed level further reduces carbon emissions in the energy sector, increases efficiency in energy delivery by shortening the distance between generation and consumption, and enhances, if the connection to the main grid is unreliable, local flexibility and reliability of energy supply [17]. However, this new development changed the traditionally unidirectional distribution energy network to a bidirectional one, which has resulted in a variety of problems. On the system level, as the distributed energy generators are not directly controlled by the system operators, overgeneration becomes an issue because a reverse power flow can lead to a voltage increase and compromise fault detection [18]. It also adds to the complexity of load and generation prediction, making it difficult for the system operators to schedule the right amount of generation [19]. On the market level, in order to reward local generation accordingly, it requires either a separate meter or a smart meter for the measurement, which incurs additional costs. Moreover, it results in a profit reduction for the electricity retailer, who may in turn increase the electricity retail price in order to recover their revenue loss [20]. This becomes a social equity issue because people who can afford to install distributed generation tend to be wealthier than those who cannot, and it is the latter who end up with higher energy bills [21]. On the regulatory level, because energy asset investments are long-term, it is a big regulatory challenge to structure an

incentive program that lasts 15–30 years for distributed generation without knowing its full impact on the power system in the future [22].

In a conventional energy system, the energy generation follows the energy demand, which is mostly achievable given the generation capacity is enough to cover the peak, and is highly controllable. From an operational point of view, base load generation (nuclear, coal, hydro, etc.), which usually has a lower operational marginal cost but is slower in changing the power output, is dispatched before peak load generation (gas turbines, hydro, etc.), which is faster in changing the power output [23]. The reliability becomes an issue when either the generation capacity is not enough to cover the load, or the speed at which the power output can be changed is not quick enough. Power system operators schedule certain amounts of base load and peak load generation capacities ahead of time for every timestep. In order to do this, a relatively accurate prediction of both the load and the generation availability is required [24]. However, the electricity a wind turbine can generate is dependent on many factors, wind speed being the most important one, which is extremely difficult to predict; solar energy, on the other hand, follows a certain annual pattern depending on the location, but the short-term solar energy variations due to factors such as weather conditions and cloud coverage, are still difficult to foresee [25]. Another major setback of solar energy is its unavailability during the night time. The uncertainty of wind and solar generation indicates that more back-up generation needs to be scheduled in case the wind or the solar generators underperform. More scheduled capacity leads to higher redundancy and higher costs, and even so it cannot guarantee the same level of reliability as a conventional highly controllable energy system [6].

According to the International Energy Agency, globally, renewable power capacity is set to increase by 50%, or 1200 GW between 2019 and 2024, led by solar PV, accounting for almost 60% of the total expected growth[1]. Among these growths, distributed solar PV capacity is forecast to see the biggest increase of over 250%, reaching 530 GW by 2024[2]. According to Figure 1.1, The UK had a slow start in picking up solar generation, which only accounted for less than 20% of all the renewable sources by 2018[3], but 13.1

[1] https://www.iea.org/reports/renewables-2019
[2] https://www.iea.org/reports/renewables-2019/distributed-solar-pv#abstract
[3] https://www.iea.org/data-and-statistics?country=UK

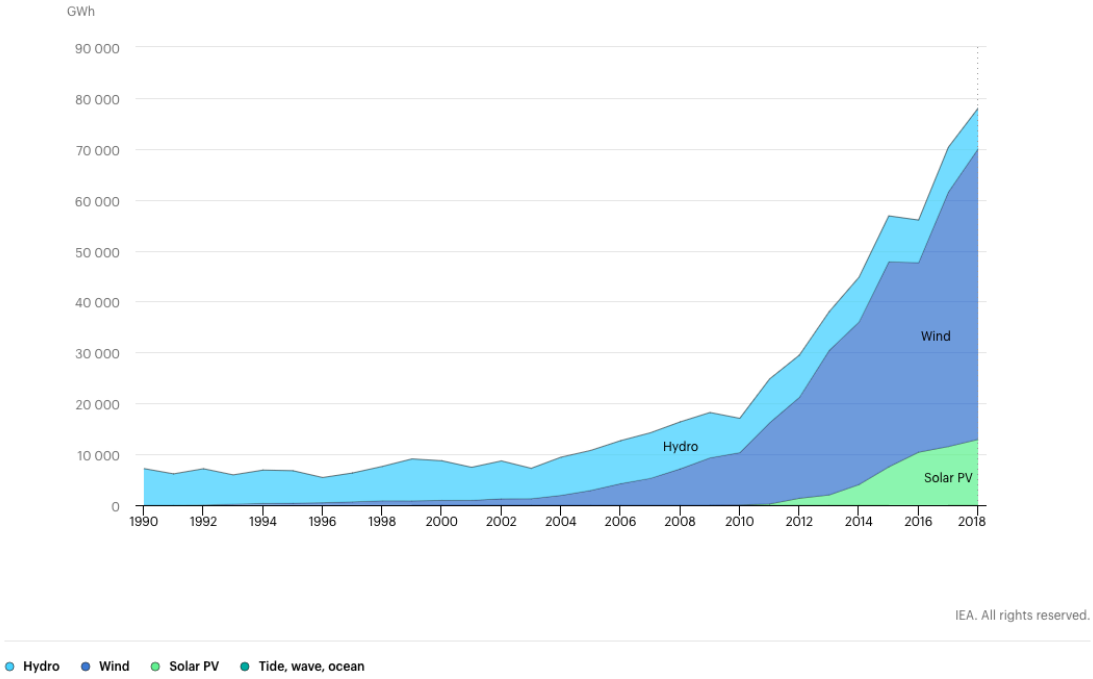


Figure 1.1: Renewable electricity generation by source (non-combustible), United Kingdom 1990-2018.

GW of solar capacity still placed the UK in the 3rd place in the EU, after Germany (45.9 GW) and Italy (20.1 GW), and over 90% of the installations are small scale PV systems[4] The residential PV capacity is expected to grow by 232-375 MW/yr until 2023[5]. Because of the major role that solar PV plays in distributed renewable generation, it is used as the sole source of distributed generation in this thesis.

Due to the negative impacts of distributed solar and wind energy generation, energy policies in recent years have a general trend of allowing retailers to curtail distributed generation under certain conditions [26], and capping the feed-in tariffs, the prices at which the distributed generation is sold back to the grid, to a level below the electricity retail price [27]. For example, the feed-in tariff for distributed rooftop PV systems in the UK has dropped from over £0.4/kWh in 2010 to less than £0.04/kWh in 2019 [28] while the retail electricity rate saw a steady increase, currently at £0.08-0.18/kWh for residential

[4]https://www.greenmatch.co.uk/blog/2019/09/uk-solar-capacity

[5]https://www.solar-trade.org.uk/wp-content/uploads/2019/11/STA-2030-Deployment-forecast-final.pdf

customers [29]. A very similar trend can be seen in Germany where the feed-in tariff was reduced from over €50/kWh to less than €10/kWh over the past 20 years [30] while the average domestic retail price climbed above €30/kWh. On a domestic electricity bill in the UK, the costs from wholesale generation only makes up 32%, while the network costs and and operational costs make up 23% and 17%, which explains why naturally distributed generation is valued at a lower price than the retail energy supplied to the distribution network. The feed-in tariff was rated significantly higher 10-20 years ago mainly due to the regulatory effort in encouraging distributed renewable development. In order to recover the capital costs, the costs of losses from the network operator's perspective, and in order for the energy suppliers to make a profit, it is only economically sustainable to set the feed-in tariff lower than the retail price.

This price difference has a very significant impact on the energy customers' energy behavior, and is a basic assumption of this thesis. Specifically, from the customers' perspective, because they are selling the local generation at a lower price than they buy electricity for, they are incentivized to increase the utilization of their generation locally. This can be done through utilizing local flexibilities and storage.

1.1.2 Energy Network Flexibility and Storage

Power networks have been evolving from a centralized model into a decentralized one with increasing amounts of distributed energy resources [31]. This has resulted in technical issues, including: 1) difficulties in generation scheduling due to the uncertainty and variability of intermittent distributed generations [6], and 2) difficulties in voltage control caused by high reverse power flows [18]. Local flexible resources and ES systems are widely seen as a solution to these problems as they can quickly respond to the generation and load imbalance in a power network and provide mitigation when needed [32].

Over the past decade, demand-side management has been widely adopted in many parts of the world, especially in industrial plants and commercial buildings [33]. It converts some of the electricity loads, traditionally viewed as inflexible, to controllable loads. Although demand-side management does not generate electricity, its ability to increase and reduce load acts in a way that is analogous to a "virtual storage" unit [34]

– it reduces the need for generation when the demand is high, and uses more electricity to absorb over-generation. Demand-side management can be implemented either by financially incentivizing end users to adjust their load themselves when needed, or by having the system operators directly control some of the electricity end uses. Bidirectional smart controls are generally needed in the second type of implementation [35].

As the popularity of electric vehicles (EV) grew among home owners, the topic of EV batteries being used as a distributed level ES resource became broadly researched in both academia and the industry. The relevant research covers not only the optimization of local energy behavior and economy, but also the potential usage of distributed ES to benefit the higher-level power system as a whole [36]. Another big driver for the increasing interest in ES is the continuous reduction in costs due to technological advances and production scaling [37].

One major difference between demand response and a physical ES unit is that demand response is unable to output energy, and is thus relatively limited in its capability to shift energy in time. EV batteries are limited by the cars availability, and are less flexible than a home ES system. The ES unit considered in this thesis does not have either of these limitations.

ES control strategies can be largely divided into centralized, decentralized, and distributed controls based on the architecture [38]. A centralized control strategy could produce the optimal results based on a given objective, but it has scalability issues [39]. Decentralized and distributed control strategies have been proposed in distribution systems [40–43], but they are inevitably limited by the assumption that the interests of all the participants are well aligned [44]. This is hardly the case in reality as each rational energy customer tends to maximize their own profits [45]. In order to create this financial incentive to efficiently manage distributed ES systems, a paradigm shift in the energy market structure becomes necessary.

1.1.3 Energy Market Transition

With the rapidly increasing deployment of DER, the energy distribution network gradually shifted from a unidirectional model to a bidirectional one [46]. Smart metering emerged as

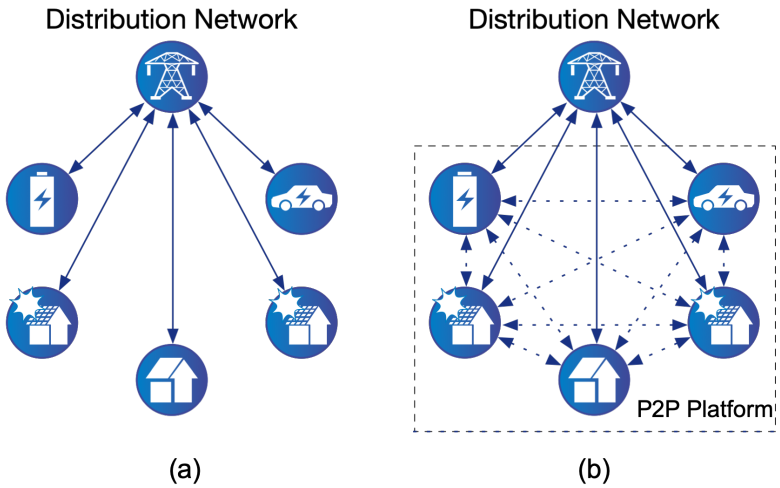


Figure 1.2: An energy network (a) without P2P trading (b) with P2P trading. The solid lines represent the transactions in the traditional centralized energy market, and the dotted lines represent the transactions among prosumers in a P2P energy network.

a technology with capabilities to measure bidirectional energy flow that could facilitate the accounting of energy payments for end users that own distributed generation assets [47].

With the development of smart appliances and energy storage, an increasing number of energy customers have become prosumers, proactive-consumers with distributed energy resources (DER) that actively control their energy behaviors [1]. A regular energy end user has little control over their local renewable generation, hence demand response has been the primary way for them to participate in the market [48]. As ES technologies matured and EV gained a larger user base, the idea of energy arbitrage provided another potential market opportunity for prosumers [39].

The increasing flexibility and controllability of energy resources and loads, coupled with the fast advancing information and communication technologies, inspired the distributed energy networking concept with mass customization and smart communication [49]. Accordingly, energy markets that used to follow a centralized clearing mechanism would need to be cleared on a distributed level. Born out of this idea was a prosumer-centric local energy market paradigm.

Under this new market paradigm, decentralized models such as microgrids (e.g., [38–40]) and virtual power plants [17] have been proposed. These market schemes tend to be optimization based with the objectives set around maximizing the overall

community benefit, which overlook the profit-drive nature of the individual prosumers [44]. However, with the fast increasing amount of user controlled flexible resources in the network, financial incentives for individual prosumers are needed to encourage coordinated operation of DER. Focusing on individual incentives, another popular concept under the prosumer-centric energy market paradigm is the peer-to-peer (P2P) energy network [50]. On a P2P platform, prosumers can actively participate in the local energy market through trading or sharing energy with other participants based on their individual objectives [51]. Figure 1.2 illustrates the comparison of a traditional centralized energy network and one with P2P energy trading.

In order to create a robust prosumer-centric local energy market, establishing a market mechanism that provides sufficient incentives for efficient DER management is key. Since prosumers are generally considered rational [52] and profit-driven, game theory offers a suitable theoretic framework for this purpose.

1.1.4 Game Theory

As discussed in Section 1.1.3, a prosumer-centric local energy market requires a robust market mechanism that offers prosumers financial incentives to efficiently manage their DER. Game theory, as a subject that studies strategic situations [53], provides a suitable methodology to structure and analyze this problem.

Game theoretic frameworks can be largely divided into two branches: non-cooperative game theory and cooperative game theory [54]. A non-cooperative game seeks to find the strategy that benefits the rational player the most, knowing that the other players are rational as well [55]. A cooperative game analyzes situations where additional values can be achieved when players operate collaboratively. It is rooted in its non-cooperative counter part [53] as forming coalitions can be seen as a strategic move. What is additionally studied in a cooperative game is the way to allocate the payoffs within each coalition.

In recent literature, non-cooperative game theory is often used in energy sharing among prosumers [56–58], who are expected to strategically adjust their energy behavior according to the dynamic dual prices. Non-cooperative frameworks enable decentralized

control of DER using prices as financial signals, but they lack the ability to guarantee that all prosumers can benefit from the market, resulting in issues such as unfairness and market instability, where prosumers may defect from the current market to seek alternative market solutions that are economically preferable.

Cooperative game theory, on the other hand, incentivizes independent decision makers to act together, investigates the fairness of profit allocation from each individual player's perspective and gives players the flexibility to decide with whom to cooperate [59]. In power system applications, cooperative game theory has been used to study ways of reducing distribution losses [60], and increasing the generation bidding price through aggregation [61]. If used to construct a prosumer-centric energy market mechanism, cooperative game theory has great potential to address the issues of unfairness and instability that are intrinsic to non-cooperative game theoretic frameworks.

1.1.5 Discussion

The advancements of distributed renewable generation, energy storage, and smart technologies have provided flexibility in the energy distribution network and enabled energy customers to become prosumers. However, they have also introduced significant variability and uncertainty that are beyond the control of the grid operators, imposing a threat to the power balance in the distribution network. In a deregulated energy market, the network operation and the market energy transactions are managed by different entities, the former by the network operator and the latter by commercial energy suppliers. Even so, the profit of the energy suppliers heavily rely on the uninterrupted operation of the network operation. Therefore, the operational challenges of the distribution network are directly reflected in the energy market in the form of differential pricing of the energy; a fundamental assumption of this thesis is that the price to import energy from the grid is higher than the price to export energy back to the grid. In turn, this price difference provides financial incentives for energy customers to increase the utilization of the local generation, potentially alleviating local uncertainty and variability. The focus of this thesis is not the quantification of the potential benefit of increased energy autonomy on

the network operation, but to explore ways to create profits for the energy customers under this new market paradigm.

Transitioning from the centralized energy market to a decentralized model, prosumer-centric markets emerged out of this context as a tool for managing the increasing amount of DER. However, there is a lack of robust market mechanisms, which require not only a technological framework to coordinate DER operation in an efficient and reliable way, but also an incentive scheme that is sustainable and fair to the participating prosumers.

Therefore, the following research question is asked, which guides the research of this thesis:

As the adoption of distributed energy resources continues to increase, how do we incentivize prosumers to manage these resources in an economically efficient and sustainable manner?

In order to construct such a market mechanism to encourage coordination of DER by using financial incentives, cooperative game theory is identified as a suitable theoretic framework. The aim of this thesis is to use cooperative game theory to develop a scalable prosumer energy management scheme, which coordinates the operation of distributed generation and energy storage with economic efficiency, and guarantees profits for each individual prosumer at a level that incentivizes participation.

1.2 Original Contributions

This thesis proposes and develops cooperative game theory as a new approach for designing prosumer-centric local energy markets. The original contributions of this thesis can be summarized as follows.

1. A coalitional energy management scheme based on cooperative game theory is proposed. In this scheme, ES systems are scheduled collaboratively to minimize the joint grand coalition energy cost. Load flattening in the distribution network is empirically demonstrated (detailed quantification of such benefits to the network operation is an interesting area for future studies). Cooperative game theory is then used to quantify the values of the coalitions, and develop a profit sharing scheme

to allocate the joint cost savings to the participating prosumers. This coalitional energy management scheme is proven to be a balanced game, meaning that there exist stabilizing payoff allocation mechanisms that would ensure that no individual prosumer or coalitions of prosumers can be financial better off defecting from the grand coalition to operate on their own.

2. A multi-step coalitional stratified sampling method to approximate the Shapley value of the proposed coalitional energy management scheme is developed. The computation of the Shapley value, a payoff allocation for cooperative games that is known for its fairness, is intractable as the number of players increases. Existing random sampling methods are modified to incorporate coalitional strata to improve its estimation performance and scalability. A novel multi-step sampling strategy is proposed and incorporated into the sampling method to further reduce the computation time.

3. A clustering based nucleolus estimation method is developed. The nucleolus is proven to be a stabilizing payoff allocation that prevents defection from the grand coalition. To address its computational complexity related to solving linear problems, the number of which is an exponential function of the number of participants, a method to estimate the nucleolus that incorporates clustering is developed. A sampling based result validation mechanism is developed to compare the estimation performances of different clustering techniques and different clustering features.

1.2.1 Contributing Publications

Material from this thesis has been submitted as papers or book chapters. Some have been published, and some under review in various journals and conference proceedings. They are listed in Table 1.1 in the order in which their relevant material appears in this thesis.

1.2.2 Thesis Scope

This thesis proposes a coalitional energy management scheme based on cooperative game theory to encourage efficient and stable coordination of prosumers' energy resources. It

Table 1.1: Contributing Research (Published and Under Review)

Contributing Research Work	**Status**	**Thesis Section**
L. Han, T. Morstyn and M. McCulloch, "P2P Platforms," in *Local Electricity Markets*, 1st ed. T. Pinto, Z. Vale, and S. Widergren, Eds. Academic Press, 2021, ch. 6.	In print	Section 2.1
L. Han, T. Morstyn, and M. McCulloch, "Incentivizing prosumer coalitions with energy management using cooperative game theory," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 303–313, Jan 2019, doi: 10.1109/TPWRS.2018.2858540	Published	Chapter 3
L. Han, T. Morstyn and M. McCulloch, "Constructing Prosumer Coalitions for Energy Cost Savings Using Cooperative Game Theory," in *2018 Power Systems Computation Conference (PSCC)*, Dublin, Jun. 2018, pp. 1-7, doi: 10.23919/PSCC.2018.8443054.	Published	Section 3.1
J J. Dixon, T. Morstyn, L. Han and M. McCulloch, "Flexible Cooperative Game Theory Tool for Peer-to-Peer Energy Trading Analysis," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1-5, Portland, OR, Aug. 2018.	Published	Section 3.2
L. Han, T. Morstyn, and M. McCulloch, "Estimation of the Shapley value of a peer-to-peer energy sharing game using multi-step coalitional stratified sampling," in *International Journal of Control, Automation and Systems*, doi: 10.1007/s12555-019-0535-1.	Accepted Aug. 2020, to be published in a future issue	Chapter 4
L. Han, T. Morstyn and M. D. McCulloch, "Scaling Up Cooperative Game Theory-Based Energy Management Using Prosumer Clustering," in *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 289-300, Jan. 2021, doi: 10.1109/TSG.2020.3015088.	Published	Chapter 5
L. Han, T. Morstyn, C. Crozier, and M. McCulloch, "Improving the scalability of a prosumer cooperative game with k-means clustering," in *2019 IEEE Milan PowerTech*, Jun. 2019, pp. 1–6,doi: 10.1109/PTC.2019.8810558.	Published	Section 5.3

is recognized that technological details and regulatory frameworks surrounding power systems operation vary significantly across the world. Although this coalitional energy management scheme is not location specific, its application may only be possible with certain infrastructure in certain regulatory environments. This section discusses the requirements and scope for the proposed scheme.

Technological and Regulatory Requirements

In order to accurately process the data of prosumers' resources, smart communication technologies are required. These technologies need to be able to control and record the energy operation of appliances (e.g. ES, flexible loads, etc.), predict generation and loads, detect the status of appliances (e.g. battery state of charge), and eventually communicate and process the information to compute the profit allocation for all the participants.

It is understood that the majority of existing energy market regulations disallow energy trading among customers. Factors such as maintaining a sustainable energy supplier business model, ensuring social equity, and preventing market manipulation all need to be considered in developing a sustainable regulatory framework to support local energy trading. There has been an increased interest in this topic internationally on a regulatory level in the forms of pilot projects [62–64] and multi-national research programs [65,66]. This thesis assumes that in the near future, such regulations will change to keep up with the prosumer-centric market development. The minimum regulatory requirement of the coalitional energy management scheme is to allow energy flow for the group of participants to be measured jointly behind-the-meter. In fact, the work is already relevant to new developments where a single meter is shared by a number of downstream consumers.

Scope of Research

Although the proposed model uses residential scenarios as case studies, it is not limited to residential applications. Rather, it can be applied to any flat or time-of-use rate structures that are known ahead of time without demand charges or real-time adjustments.

The coalitional energy management scheme focuses on coordinating resources within local distribution networks, which are not expected to have significant losses and constraints between prosumers. Therefore, the discussion of network losses and constraints is outside of the scope of this thesis.

A main assumption of this thesis is that the feed-in tariff is lower than the energy retail price. This is consistent with the majority of the current energy markets for reason explained in detail in Section 1.1.1, and it provides a financial incentive for prosumers to coordinate resources to meet energy demand locally.

Prosumers in this thesis are assumed to be individually rational [52]. In other words, they are willing to participate in a scheme that makes them economically better off than being on their own. Later chapters discuss payoff allocations that provide appropriate incentives to prevent prosumers from leaving the grand coalition, the coalition of all prosumers considered, to form smaller coalitions. The social science behind what level of financial incentives is enough to encourage participation, and to what degree a lack of financial incentives would lead to defection is outside of the scope of this thesis.

Data privacy is an important subject in implementing smart grid technologies, but is outside of the scope of this thesis.

1.3 Thesis Structure

The rest of the thesis is organized as follows:

To gather evidence for the main research question and identify research gaps, Chapter 2 provides a review of the existing literature on topics related to distributed energy markets, game theory, and load clustering. These studies provide the motivation and the fundamental data basis and methodologies for the research presented in this thesis.

Chapter 3 addresses part of the research question by proposing a coalitional energy management scheme to incentivize efficient management of DER. A cooperative game theoretic framework is adopted to construct the coalitional energy management scheme. Specifically, the basics of cooperative game theory are reviewed, and the coalitional energy management scheme is presented in a mathematical format including the coalitional energy cost minimization, the coalition value definition, and the payoff allocation.

This scheme is then proven to be balanced, meaning that there exist stabilizing payoff allocations that prevent participants from defection. The Shapley value and the nucleolus are introduced as common payoffs for cooperative games, and they are implemented in the case studies to demonstrate the effectiveness of the proposed scheme.

Computing payoffs for a cooperative game is computationally complex, which makes it difficult to scale up the coalitional energy management scheme to a size applicable for a real local energy network connected to a distribution feeder. Resolving the scalability issue is the main focus of Chapter 4 and Chapter 5.

Chapter 4 focuses on one of the payoff allocations introduced in the coalitional energy management scheme: the Shapley value. It is widely adopted in other cooperative game theory applications. Well known for its fairness, the Shapley value represents the mean of a player's marginal contributions to all the possible coalitions within the game. Although it has an analytical expression, it becomes computationally intractable as the number of players increases. A novel multi-step stratified random sampling method is proposed to approximate the Shapley value for large games, and it is shown to be able to scale up the game from less than 20 players to more than 100 players.

Despite the fairness of the Shapley value, it is not necessarily a stabilizing payoff allocation. In other words, it cannot guarantee participants higher benefits in the grand coalition than smaller ones, which may give them an incentive to defect from the grand coalition. Therefore, Chapter 5 shifts the focus to the nucleolus, another payoff allocation that is proven in Chapter 3 to be stabilizing. However, the computation of the nucleolus becomes intractable when the number of participants exceeds 15. This is mainly because the number of linear problems required to be solved is an exponential function of the number of participants. Through clustering prosumers, a novel method is proposed to estimate the nucleolus for larger games. Multiple clustering techniques coupled with different clustering features are investigated in this chapter and their performances are evaluated and compared using a novel stratified sampling based tool. The proposed method is shown to be able to effectively estimate the nucleolus for coalitional energy management schemes with 200 players in under 10^4 seconds.

Finally, Chapter 6 concludes the thesis, summarizes the main findings, and discusses promising areas for future work.

Appendix A contains additional proofs of the balancedness of the coalitional energy management scheme in conjunction with Chapter 3. Appendix B lists the main inputs for the case studies presented in this thesis.

A complete list of relevant references can be found in the Bibliography.

We are not makers of history. We are made by history.

— Martin Luther King Jr.

2 Literature Review

Contents

2.1 Prosumer-Centric Local Energy Markets 21
2.1.1 System-Focused Market Schemes 23
2.1.2 Individual-Focused Market Schemes 24
2.2 Game Theory for Prosumer-Centric Energy Market Design 28
2.2.1 Non-cooperative Game Theoretic Frameworks 28
2.2.2 Cooperative Game Theoretic Frameworks 30
2.3 Customer Load Profile Clustering 33
2.3.1 Clustering Features 34
2.3.2 Clustering Techniques 35
2.4 Gap Analysis 37

Motivated by the increasing need to coordinate DER and to keep up with the rapidly evolving energy market, this thesis aims to develop a prosumer-centric local energy market that incentivizes prosumers to efficiently manage their energy resources. This chapter reviews the state of the art literature on prosumer-centric local energy markets and identifies research gaps that motivate the thesis.

In recent years, an important research topic in the energy field is how to maintain the reliability of energy supply while encouraging distributed renewable generation, which is highly variable and intermittent [6]. The idea of setting up a prosumer-centric local energy market has gained tremendous attention both in the industry and academia in recent years, for one of its key features is using local flexibility to offset generation uncertainty.

Local flexibility often takes on the form of energy storage (ES), which can be modeled in a similar way as other types of flexible loads [67]. As the penetration of distributed generation increases, the feed-in tariff has largely fallen below the retail price of energy [8, 68]. Therefore, this thesis makes the common assumption that the price to export energy to the energy network is lower than the price to import; hence, for a single prosumer, the benefit of flexibility is directly reflected in their energy bills when they increase the local usage of their own generation by optimally scheduling their ES.

In a prosumer-centric local energy market, economic benefits can be achieved from matching local flexibility with variable generation among all the participants, so joint management of DER for maximizing community benefits has been proposed in various studies. However, these schemes tend to overlook the fact that prosumers are independent entities who are motivated by individual incentives to participate in such a market. Therefore, market mechanisms that focus on incentives for individual prosumers have emerged in recent years to encourage flexible and efficient distributed management of DER. A detailed review on the architectures, the theories, and the technological and regulatory requirements of different types of prosumer-centric local energy markets is presented in Section 2.1

Game theory studies strategic situations and is a methodology suitable for analyzing financial incentives in a prosumer-centric local energy market. Both non-cooperative game theory and cooperative game theory have been adopted in recent research to establish market frameworks to influence prosumer behavior with financial incentives. Section 2.2 describes the background of game theory, both cooperative and non-cooperative, reviews their applications in the energy market, and offers an insight into their computational challenges.

Cooperative game theory offers a framework to allocate the profits of the prosumer-centric energy market to participants in an efficient and fair manner, but its computational complexity increases exponentially with the number of participants. Section 2.3 reviews the use of clustering in categorizing prosumers based on their load profiles, which offers a potential solution to the computational issues of a cooperative game theory based prosumer-centric energy market.

2.1 Prosumer-Centric Local Energy Markets

A well-designed prosumer-centric local energy market not only satisfies prosumers' monetary or non-monetary needs, but lowers the burden for the system operators to manage and maintain the energy balance within the networks. Similar to other sharing economies, such as car sharing and home rentals, prosumer energy sharing introduces greater efficiencies into the energy market, and financial benefits to the individual prosumers [69].

Based on the exiting literature, the architecture of a prosumer-centric energy market platform can be largely divided into three main components.

The first component is the market mechanism, which sets the general objectives of the platform [51]. Among these, the main goal should be increasing the welfare, either financial or social, of the participants [70]. This component is the main focus of the thesis, and an extensive literature review on this topic is presented in this section.

The second component is the network control, which lays the technical foundation for the platform [71]. The main goal of this component is to maintain the energy balance, where robust control and real-time monitoring of the power network are necessary to ensure the energy supply security [72]. This component is responsible for power flow modeling with network constraints [73–75] and real-time power balancing [76, 77].

The third component is the information and communication system, which carries and processes data and provides a user interface for the platform [72]. The main goal of this component is to engage prosumers. Advanced information and communication technologies (ICT) enables users on the market platform to decide how they would like to participate in the market, including bidding on the platform and setting individual preferences [78]. It is also in charge of recording the cleared transactions. Blockchain with smart contracts is a popular technology for this application [79–81] and has been used in local energy market trials [62, 82].

The operation of a prosumer-centric energy market platform requires these three components to constantly interact with and support each other. The decisions made by the market component need to be monitored and adjusted by the network control component

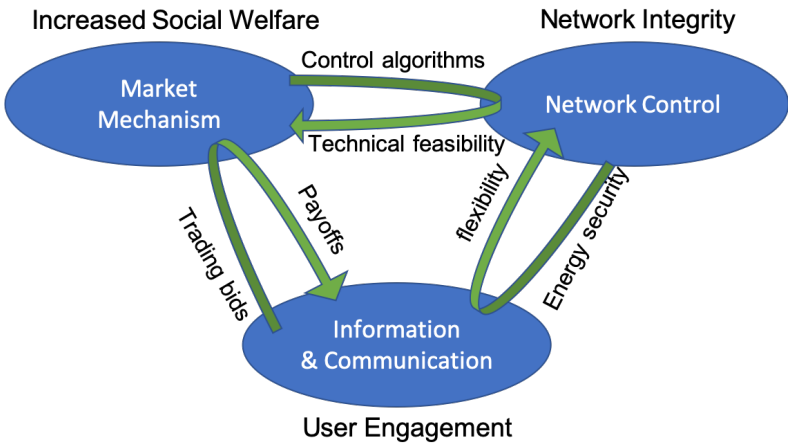


Figure 2.1: Interaction between the three main components of a prosumer-centric energy market platform. Market decisions are monitored and adjusted by the network control to ensure their technical feasibility. These decisions are realized in the form of financial transactions, implemented by the information and communication infrastructure, which also responds to network control signals by adjusting flexible resources within the network for energy balancing.

to ensure their technical feasibility [83]. These decisions also inform transactions implemented by the information and communication component. In turn, the information and communication component feeds user bids into the market component, and enables the network control component to adjust the flexible resources within the network for energy balancing [72]. Fig. 2 demonstrates the interaction among these three components.

The goal of this thesis is to establish a theoretic framework for the prosumer-centric energy market. Therefore, the market mechanism component is the main focus.

The motivations behind setting up prosumer-centric markets range from generating economic benefits, to increasing clean energy utilization, and promoting community sharing [84]. This thesis mainly focuses on prosumer-centric market schemes that aim to generate economic benefits. Such market schemes commonly assume that the value of the energy supplied to the prosumers is higher than the prosumer generation, which is often reflected in a higher energy retail price than the feed-in tariff [85]. Various factors contribute to this pricing disparity, including the need to cover the cost of the distribution infrastructure, and distributed generation's high losses and unreliability compared with energy obtained from the wholesale market [86].

Depending on the main beneficiary, prosumer-centric local energy market schemes can be divided into two categories: system-focused, and individual-focused. Figure 2.2

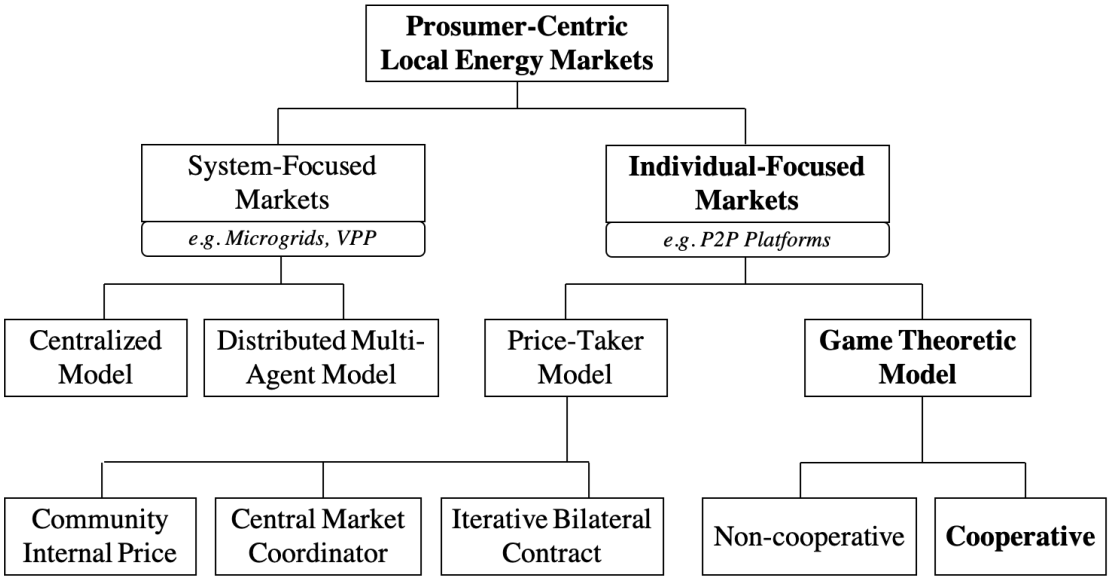


Figure 2.2: Tree diagram showing the categorization of prosumer-centric local energy markets. The bolded text shows the process of selecting the theoretic market framework for the energy management scheme proposed in this thesis.

presents a detailed categorization of different types of prosumer-centric local energy markets, and a detailed literature review of these market types are conducted in Section 2.1.1 and Section 2.1.2.

2.1.1 System-Focused Market Schemes

System-focused schemes are commonly presented in literature as microgrids [38, 40–42, 87–90], virtual power plants (VPP) [87, 91–93], or some other form of central aggregator [39, 43, 94]. Some of these schemes have multiple objectives, but one common objective of all these market schemes is to maximize the collective economic outcome of the participants through optimizing the DER operation within the network.

There are largely two types of optimization models for system-focused market schemes: the centralized model, and the distributed multi-agent model [38]. Decentralized control strategies are excluded from this list because they use local control strategies based on very limited information, which do not directly contribute to the optimal economic results for the community [40, 43].

The centralized model assumes that the central aggregator has full control over all the energy resources within the network [87]. Based on forecasts of the prosumer load

and generation, the central aggregator schedules the flexible resources within the network to minimize the cost of the whole system [92]. This is widely adopted for islanded and interconnected microgrids [88,89], and it can produce a global optimal outcome assuming no forecast error [91]. However, centralized control has very limited scalability due to the large amounts of data and the complex control sequance [39].

Instead of having the central aggregator collect information from all the controllable resources for global optimization, the distributed multi-agent model establishes communication among agents and enables control optimization based on local information and information of neighboring agents [39, 41, 42, 90]. Due to the significant reduction in the model complexity, the distributed multi-agent control significantly improves the market scheme's scalability.

However, system-focused energy markets are inevitably limited by the assumption that all of the participants align their interests with the larger community. Even though the maximum overall community welfare is achieved, a system-focused market does not address how the welfare is shared among the participants. In reality, however, rational prosumers will not change their energy behavior altruistically for the benefit of the community with no or negative financial gain for themselves [44]. The individual-focused local energy market resolves this issue by investigating the economic outcome for individual prosumers that are participants in the market.

2.1.2 Individual-Focused Market Schemes

Individual-focused local energy market schemes can be further divided into two subcategories: the price-taker model [95–100] and the game theoretic model [55–57,59,101,102]. In a price-taker model, prosumers have zero or little market power and simply accept the market prices as given. In contrast, prosumers are modeled to be strategic in game theoretic models [103], where they actively adjust their energy behavior to achieve the best personal outcome.

Within the price-taker model, there are generally three pricing strategies: community internal price, central market coordinator, and iterative bilateral contract.

Setting an internal price for a prosumer community is the simplest way of financially benefiting participants by increasing the community's energy self-sufficiency. This includes either setting a single internal price for both buying and selling energy within the prosumer-centric energy market [95, 104], or setting two different buy and sell prices [96–98]. All the participants would be exposed to the same internal buy price and the same internal sell price regardless of their energy operation. When the retail price is higher than the feed-in-tariff, as long as the community internal prices are set between the retail price and the feed-in tariff, buyers will be able to buy energy at a cheaper price within the prosumer-centric energy market, and sellers will be able to sell energy at a higher price within the market. Three basic internal pricing schemes are proposed in [95]: 1) bill sharing, in which the energy import and export prices are averaged and applied to all participants, 2) mid-market rate, in which the internal buy and sell prices are both equal to the average of the retail price and feed-in tariff, and 3) auction-based pricing strategy, in which the internal buy and sell prices are equal to the highest seller offer. Considering the network charges and the market platform's operating charges, piecewise functions of the supply demand ratio were used in [96] to model the internal sell and buy prices. Included in [97] are prosumers with concave utility functions of economic gain, a common assumption for computational properties that means diminishing marginal returns for energy. The model also includes a local trading center, which determines the internal sell price and buy price by 1) maximizing its own profit if modeled as a for-profit entity, or 2) maximizing the total utility of all the participating energy buyers and energy sellers as a non-profit entity. A more sophisticated bidding platform was developed in [98], where each flexible device type has its own bidding strategy, which is adjusted by a home energy management system depending on the user preferences, and sent to the central market platform that sets the internal price to clear the market.

One big disadvantage of the internal price strategy is that it fails to fairly reward the participating prosumers based on their diverse needs and energy profiles. Even though setting fixed prices is a common practice in the energy wholesale market, wholesale prices are usually determined through either contracts or bids based on the generator's capital and operational costs [105]. In a prosumer-centric market, however, participation

incurs almost zero costs, and it is reasonable to assume that rational prosumers are likely to choose market schemes that can offer them their fair share of the profit. To allow each prosumer to receive prices based on their own operational preferences, the concept of a central market coordinator was explored in several studies. An Energy Broker was proposed in [99] to model a for-profit coordinator who oversees the bids of the energy sellers and buyers within a local electricity market and sets the final selling and buying prices to maximize its own profit. In [100], the load aggregator recalculates the bids from energy buyers and sellers with the nodal prices from the local electricity coordinator, which aims to maximize the social surplus through pooling traditional generators, residual loads and excessive energy from prosumers.

With the central market coordinator strategy, although prosumers are assigned prices based on their own bids, these final prices, being often different from their original bids, may likely result in them deviating from their said operation, leading to inefficiency and instability in the market. To address this issue, several studies have proposed the use of iterative bilateral contracts to better reflect the prosumers' personal objectives in the trades. A multi-bilateral economic dispatch optimization problem was formulated in [78] where bilateral trades are established between all the seller-buyer pairs with trading amounts and prices as variables. The objective is to maximize the social welfare of the prosumer-centric energy community by iteratively solving for the optimal trading costs of each bilateral trade. This work provides a distributed framework to analyze the impact of product differentiation reflected on the prosumers' different utility functions. Also focusing on prosumer preferences, another distributed price-directed optimization mechanism using alternating direction method of multipliers (ADMM) was presented in [93], which enabled prosumers to trade different energy products to satisfy environmental and philanthropic preferences. The results show a clear impact of prosumers' preferences on their final trading prices.

Overall, a price-taker model provides a strategy to incentivize prosumers to participate in the local energy market. The maximum social welfare can still be achieved with the central market coordinator strategy or the iterative bilateral contracts strategy, but

they nevertheless lack the framework to verify the satisfaction of participants or market stability. Game theory, by contrast, is known to offer such a framework.

In the game theoretic model, both cooperative and non-cooperative games have been adopted to facilitate the design of the prosumer-centric energy market.

A number of non-cooperative game theory based flexible load management schemes deploy dynamic pricing of energy, and the price is set by either a for-profit retailer [56], or a non-profit aggregator such as a microgrid operator [57], or a multi-energy system [58]. The iterative communication between the electricity suppliers and the end users converges to the real-time prices. An iterative bilateral contract model was introduced in [84], where each agent (a prosumer, a supplier, or a generator) simultaneously bids and accepts contracts in the market to maximize their utility while maintaining their individual energy balance. Through an iterative distributed price-adjustment process, all the contract prices eventually reach a competitive equilibrium. A Nash equilibrium was proved in [106] to be unique among buyers. Further, the uniqueness of the Nash equilibrium among all sellers, who can be sellers or buyers, is observed empirically in various scenarios simulated in [107]. To solve for the Nash equilibrium of an energy trading game, the Nikaido-Isoda function was used in [108], and ADMM was used in [109] and [110].

Non-cooperative game theoretic schemes focus on finding the market equilibrium, such as the Nash equilibrium [55], from which no player can be financially better off deviating. This market equilibrium can often be solved in a distributed fashion through iterative communication among participants [84] or between participants and a central market operator [57, 106, 109, 110]. It is possible to achieve a unique Nash equilibrium that also yields the maximum social welfare [111].

Cooperative game theoretic schemes, on the other hand, use cooperation as a means to achieve the maximum social welfare [102]. The focus of these schemes are on the allocation of the total profit among the participants, who are assumed to be profit driven in forming coalitions [61]. Fairness and the level of incentive are two common criteria used in evaluating the profit allocations [59].

Under the assumption that prosumers are rational and strategic [52], a game theoretic model is more suitable for addressing the thesis question by creating a prosumer-centric

market that is economically efficient and sustainable. Therefore, A more detailed review of the application of game theory in prosumer-centric energy markets is conducted in Section 2.2

2.2 Game Theory for Prosumer-Centric Energy Market Design

As explained in Section 2.1.2, there are numerous ways of designing a prosumer-centric local energy market, which is individual-focused. The price-taker model may be easy to implement and simple for customers to understand, but it cannot fairly reflect the values of different types of flexibility, which may cause some participants to defect. Game theory is known for its use in dealing with strategic situations [53], and is thus a suitable methodology for constructing a local energy market that individual rational prosumers are willing to participate in. Both non-cooperative game theory and cooperative game theory can be used as frameworks for the prosumer-centric market design.

2.2.1 Non-cooperative Game Theoretic Frameworks

Non-cooperative game theory is often used to study prosumer strategic behaviors, where each prosumer aims to achieve their highest individual utility [112]. There are three main categories of non-cooperative game theory application in the electricity market: generator strategic bidding, EV charging, and flexible load management. These applications rely on finding the Nash equilibrium, which guarantees that no player has the incentive to change their strategy while the strategies of other players remain fixed [55].

The non-cooperative game theoretic way for an energy producer to determine how much capacity to bid in a competitive forward market, given a price curve as a function of demand, is by solving for the Nash equilibrium [113]. The benefit of this approach is that no player is able to find a more profitable strategy than their Nash equilibrium given all other players stick to their own equilibrium strategies. This approach, however, only works when the market is highly transparent and energy producers, especially ones with large market shares, do not attempt to manipulate the market through collusion or shill

bidding [114]. A classic example of energy market manipulation is the California energy crisis during year 1999–2001 amidst its energy market deregulation attempt [115].

As the adoption of EV continues to grow, evening charging has become a concern because it may create a new demand peak which requires additional generation capacity [116]. EV batteries can be modeled as large distributed ES systems with additional constraints related to EV battery availability due to traveling [117]. The application of non-cooperative game theory in EV charging mostly revolves around the strategy to encourage "valley filling" [24] to avoid creating a new evening peak. When the number of EV owners approaches infinite, the Nash equilibrium charging strategy of each EV owner operating on their own is shown to coincide with the social welfare optimum, under the assumption that the electricity price is a linear increasing function of the demand [118]. A distributed control strategy is proposed in [119] to achieve "valley filling" by having every EV owner operate at their own Nash equilibrium. Where bidirectional energy trading is allowed, [120] proposes an iterative control strategy via communication between the load aggregator and the EV owners to help them achieve Nash Equilibria, while encouraging energy conservation.

A number of non-cooperative game theory based flexible load management schemes deploy dynamic pricing of energy, and the price is set by either a for-profit retailer [56], or a non-profit aggregator such as a microgrid operator [57], or a multi-energy system [58]. The iterative communication between the electricity suppliers and the end users converges to the real-time prices. An iterative bilateral contract model was introduced in [84], where each agent (a prosumer, a supplier, or a generator) simultaneously bids and accepts contracts in the market to maximize their utility while maintaining their individual energy balance. Through an iterative distributed price-adjustment process, all the contract prices eventually reach a competitive equilibrium. A Nash equilibrium was proved in [106] to be unique among buyers. Further, the uniqueness of the Nash equilibrium among all sellers, who can be sellers or buyers, is observed empirically in various scenarios simulated in [107]. To solve for the Nash equilibrium of an energy trading game, the Nikaido-Isoda function was used in [108], and ADMM was used in [109] and [110]. However, the Nash

equilibrium is inefficient. In other words, the resulting strategy profiles of the prosumers are not optimal for maximizing the sum of their profits [121].

To study the independent behaviors of the sellers and buyers, it was proposed in [106] to use the framework of Stackelberg games, where players are classified as leaders and followers in a two-level game, where the leaders move first (e.g. set prices, broadcast amount of energy to sell) based on their knowledge of the followers actions (e.g. energy operation, purchasing bids) for a set objective (e.g. own profit, total social welfare). The followers are always the buyers, but the leaders could be either a single central entity (e.g. shared facility controller [112], centralized power system [122]), or all the sellers within the network [123]. A Stackelberg game approach is used in [56] to analyze the trade-off between the retail profit and the consumer surplus, which is then extended by [57] to a bi-level optimization problem for a microgrid. Natural gas and heat energy is further included in [58] in addition to electricity.

The Nash equilibrium needs to be computed in an iterative process, where sellers update the broadcasted energy prices and buyers update their projected operation profiles, and their convergence characteristics are demonstrated in [124]. It is shown in [111] that in energy trading Stackelberg games that maximize social welfare, prosumers can achieve 47% higher utility than only participating in the regular centralized energy market.

2.2.2 Cooperative Game Theoretic Frameworks

Cooperative game theory has also been used for various energy system applications. On the generation side, [61] proposed a cooperative game theoretic strategy to aggregate wind power generation in order to reduce uncertainty, thus increasing the contracted price. The nucleolus was selected in this case to prevent members from defecting from the grand coalition. In the distribution network, cooperative game theory can be used to quantify the benefit of local energy sharing in reducing transmission and distribution losses, and increasing the self-sufficiency of the local energy network [59]. For example, a cooperative game among microgrids was proposed in [60] with the objective of minimizing the costs incurred by the losses over the distribution lines.

Cooperative game theoretic approaches encourage the maximization of profits of the grand coalition, which includes all the participants. Cooperative game theory focuses on the allocation of the grand coalition profit. There are infinite ways of sharing the grand coalition profit among all the participants. Some simple solutions include equal surplus division and proportional surplus division, which can be solved analytically [125]. Taking fairness into account, the Shapley value is a widely used allocation, which represents the weighted average of a player's marginal contribution to all the possible coalitions [126]. Another important set of payoff allocations is the core [127] – the set of payoff allocations that guarantee that no subgroup of participants has an incentive to abandon the grand coalition. Using a noncooperative approach to implement a cooperative game, the Nash core is proposed in [128], which includes the Nash bargaining solution [129] that ensures all other players play the same bargaining game among themselves if any coalition deviates from the cooperation. The nucleolus [130], a payoff allocation calculated by minimizing the dissatisfaction of players [131], is guaranteed to be in the non-empty core and Nash core.

The key to constructing a stable cooperative game is by allocating the profits to each participant in a way that makes sure that no one can be better off leaving the grand coalition to operate individually or form smaller coalitions [61]. As opposed to being exposed to dynamic pricing in a non-cooperative scheme, a known centralized energy price structure is normally assumed where the energy buy price is higher than the sell price from the prosumers' perspective. Through cooperatively scheduling their energy resources, the prosumers as a coalition can achieve energy cost savings, which then are allocated to each participant following a certain algorithm. Even before the concept of microgrids was fully developed, forming multi-agent coalitions to trade energy in a "free market" had been discussed in [132] in 1999. Cooperative game theory was used to construct agent coalitions and to allocate profits. However, the network was oversimplified with the assumption that each agent could only trade with their two neighboring agents. A direct trading cooperative scheme was proposed in [102] for prosumers to sell excess PV generation to others under the same scheme. The Shapley value was used as the payoff allocation method, and was proved to be in the core. A similar scheme was

proposed in [133], where a more simplified mid-market rate was introduced to facilitate the payoff allocation. Neither of these two market schemes, however, took into account the multi-timestep operation of ES systems.

The thesis author's work published in [134] proposed a multi-timestep prosumer-centric market scheme based on cooperative game theory. It incorporated a time-of-use energy price structure and allowed the active participation of ES systems. It was proved that all participating prosumers could financially benefit from the proposed market scheme. To take it one step further, the thesis author published a subsequent work in [135], which proved that the nucleolus was guaranteed to incentivize prosumers to participate in the grand coalition and prevent them from forming smaller coalitions. Cooperative game theoretic market schemes later published by others showed similar findings and proofs [136, 137].

2.2.2.1 Computational Complexity

Under the cooperative theoretical scheme, the payoff allocation is often linked to each player's contribution to the value of coalitions, and because the number of coalitions increases exponentially with the number of participants [52], the computation of the payoffs is intractable for large games [138].

A player's Shapley value in a cooperative game is a weighted average of their marginal contribution to all the possible coalitions among all players, see [126]. Due to its definition, the Shapley value is known for its fairness. For an N-player game, there are 2^N possible coalitions, which means that the computation of Shapley value becomes intractable when increasing the number of players. The estimation of the Shapley value has been explored in some previous literature, sampling being the main methodology. An example is the cooperative scheme described in [139], but the model is constructed as simple games with a binary outcome representing whether a coalition of battery-owning households can overcome a hard network constraint. This scheme overlooks the contributions made by local generation or ES units that are not big enough to switch the binary outcome. A random sampling method was introduced in [140] to only calculate the values of randomly

selected coalition samples, based on which the Shapley value [126] could be estimated. Stratification was subsequently added in [141] to improve the estimation accuracy.

Similarly, for a cooperative energy sharing scheme such as the ones proposed in [135–137], the nucleolus becomes computationally prohibitive as the number of prosumers increases because 1) it requires solving convex optimization problems to calculate the values of all coalitions, the number of which is an exponential function of the number of prosumers [139], and 2) given the values of all coalitions, the number of iterative steps required to compute the nucleolus is also an exponential function of the number of prosumers [142].

2.3 Customer Load Profile Clustering

Clustering is an unsupervised data mining technique for pattern recognition, in which the pattern is assigned to a hitherto unknown class based on the similarity of patterns [143].

As described in Section 2.2.2.1, the scalability of a prosumer-centric energy market based on cooperative game theory is limited by the number of participants. Knowing that a prosumer's allocated payoff is directly linked to the prosumer's energy operation, it is considered that a method for grouping players based on the similarities in their energy operation could potentially reduce the number of participants needed for the computation of the payoff allocation. Therefore, clustering, as a commonly used tool for understanding the energy customer load shape [144], is identified as a promising strategy.

There are different ways of categorizing customer load patterns. Traditionally, macro-categories are used to separate residential, commercial, and industrial customers [145]. To obtain more granular information, extensive field measurements [146] together with comprehensive surveys [147] are used to categorize customers. As load measuring technologies mature, it has become possible to use smart meters to collect a large number of customer data that can be used to study customer load patterns at very low costs [148]. With the abundant data, neural networks [149] and other data mining methods [150] have been developed to model customer load profiles.

Among these, clustering, for its well-known capability to assign its objects to a set of clusters [151], has become more and more commonly used for categorizing customer

load patterns ([146, 152–159]). Some specific applications include tariff formulation ([146, 152, 154, 155]) and electric load forecasting and management [156–159].

In categorizing customer load profiles, clustering follows mainly three phases [144]:

1. *Pre-clustering phase*: clustering features selection, and size reduction.
2. *Clustering phase*: formation of clusters using the selected algorithm.
3. *Post-clustering phase*: construction of final classified load profiles.

In this thesis, the application of clustering focuses on the the first two phases.

2.3.1 Clustering Features

In a clustering application, the sample size is often limited, which means determining the type of features is crucial in order to obtain unbiased results [160]. For classifying customer load profiles, the clustering features chosen are typically the load data (e.g. 96 features with 15-min data over one day) [145]. However, it can become computationally prohibitive if the dimension of the features is too high [159].

In order to select features that are most suitable for the purpose of clustering while preventing any computational issues, techniques on feature definition, feature extraction, and feature size reduction have been widely used in the literature [152, 156, 157, 161–165].

To define features based on pre-existing knowledge of the data, some studies have come up with strategies to customize the features to fit their specific applications. For example, instead of using all the load data in a dataset, the authors in [161] define a set of shape indicators to characterize the load patterns such as the daily average to maximum load factor. In [162], each day is divided into four key time periods. The number of consumption peaks, the average consumption, and the maximum consumption are included in the features in [163].

Various feature extraction techniques have also been adopted in the application of load profile classification. Dynamic Time Warping is introduced in [164] to take into account the sequence, rather than simply the discrete magnitudes, of the profile data. Discrete Fourier transform is used in [165] to transform time-domain measurements to a reduced set of data defined in the frequency domain.

The post popular data size reduction method in energy customer clustering is principal component analysis (PCA) [166]. It is used to reduce the dimensionality of a dataset that consists of correlated variables by transforming the dataset into a new set of variables called principal components that are uncorrelated [159]. PCA is shown to be very effective in reducing the size of the features while retaining the variation in the data [152, 156, 157].

2.3.2 Clustering Techniques

Various clustering techniques have been proposed and applied to customer load pattern recognition and new ones continue to emerge. The authors in [143] discuss two major clustering techniques: iterative square-error partitional clustering and agglomerative hierarchical clustering. Fuzzy clustering is applied in [152], which also proposes the use of self-organizing maps (SOM) for electricity customer classification. Finite mixture models are adopted to cluster electricity customers [158, 162]. A new categorization proposed by [159] separates clustering techniques into three groups: distance-based techniques, SOM, and probabilistic models .

Distance-based clustering techniques usually adopt the Euclidean distance framework, which is defined as

$$d(\mathbf{a}_1, \mathbf{a}_2) = \sqrt{\frac{1}{Z}\sum_{z=1}^{Z}(a_{1z} - a_{2z})^2} \tag{2.1}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are two vectors, each with Z components.

K-means is the most widely used clustering technique [159]. With a predefined K, it sorts all vector instances $\mathbf{a}_i$ $(i = 1, ..., N)$ into K clusters, which minimize the sum of the square Euclidean distances between all instances and their cluster centroids $\mathbf{c}_j$ $(j = 1, ..., K)$:

$$\min_{\mathbf{c}_1,...,\mathbf{c}_K} \sum_{j=1}^{K} \sum_{\mathbf{a}_i \in C_j} d^2(\mathbf{a}_i, \mathbf{c}_j) \tag{2.2}$$

Fuzzy c-means clustering is a variation of K-means clustering, allowing each instance to belong to more than one cluster [167]. It introduces a fuzzy partition matrix $\mathbf{U} = \{u_{ij}\}$ into the objective function:

$$\min_{\mathbf{U},\mathbf{c}_1,...,\mathbf{c}_K} \sum_{i=1}^{N} \sum_{j=1}^{K} u_{ij}^m d^2(\mathbf{a}_i, \mathbf{c}_j) \tag{2.3}$$

where u_{ij} represents the degree of membership of ith instance in jth cluster, and $\sum_{j=1}^{K} u_{ij} = 1, \forall i$. m controls the level of fuzziness, where larger m increases the fuzziness of the algorithm [157].

Both K-means and fuzzy c-means are heuristic methods that require a predefined K to initiate the learning [168]. Hierarchical clustering, albeit distance-based, can be performed without a preset number of clusters [169]. Hierarchical clustering is done with either a divisive (top-down) approach or an agglomerative (bottom-up) approach [170]. The divisive approach starts by treating all instances as one cluster, and continues by splitting it into sub-clusters based on measures of dissimilarity. The agglomerative approach, on the other hand, treats each instance as a cluster initially and proceeds by merging clusters based on measures of similarity. For both approaches, the algorithm stops whenever the desired number of clusters is reached. Agglomerative hierarchical clustering is more computationally efficient for large datasets [169], and is the preferred approach in the power system domain [159].

Unlike the distance-based clustering techniques, SOM is an unsupervised neural network, which rearranges the original data to a lower dimensional space [171]. It is not a direct clustering technique [152], and another clustering algorithm, such as K-means or hierarchical clustering, has to be applied in conjunction with SOM to obtain the final clusters [172].

Taking a probabilistic approach, finite mixture models can also be used to cluster customer load profiles [158]. Assuming all the instances arise from a distribution that is a mixture of several probabilistic components:

$$p(\mathbf{a}_i|\boldsymbol{\theta}) = \sum_{j=1}^{K} \alpha_j p(\mathbf{a}_i|\boldsymbol{\theta}_j) \tag{2.4}$$

where $\alpha_1, ..., \alpha_K$ are the mixing probabilities: $\sum_{j=1}^{K} \alpha_j = 1$, and $\alpha_j \geq 0, \forall j$, and each $\boldsymbol{\theta}_j$ is the set of parameters defining the jth component. The finite mixture model then estimates the parameters of each component and identifies the component that predicts each instance the best [159]. All the components are usually assumed to have the same functional form, most commonly a multivariate Gaussian distribution [143]. In this case, it is called a Gaussian Mixture Model (GMM) [173]. The parameters of the

model $\{\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_K, \alpha_1, ..., \alpha_K\}$ are usually obtained using the expectation-maximization algorithm, and each instance is assigned to the component with the highest posterior probability [168]:

$$p(k|\mathbf{f}_i) = \frac{\alpha_k p(\mathbf{a}_i|\boldsymbol{\theta}_k)}{\sum_{j=1}^{K} \alpha_j p(\mathbf{a}_i|\boldsymbol{\theta}_j)}, \quad k = 1, ..., K \tag{2.5}$$

2.4 Gap Analysis

The increasing amount of DER and local flexibility is a key driver for the development of prosumer-centric energy markets. These market platforms provide tools for prosumers to become active participants in the energy market, which has traditionally been a unidirectional space where electricity users strictly purchased from electricity suppliers. Without proper regulations and robust market mechanisms, this new capability for end customers to actively control their consumption and generation could introduce significant uncertainty into the power systems. At the same time, it presents great opportunities for managing flexibility and alleviating reliability issues in the distribution network.

As explained in Section 2.1.1, a system-focused market has an objective to maximize the joint benefit of the prosumer community, which lacks the incentive for individual prosumers to participate. An individual-focused market (see Section 2.1.2), on the other hand, provides financial incentives to participating prosumers. Among these market schemes, the price-taker model does not reward participants based on the true value of their flexible resources, introducing inefficiency and unfairness into the market, while the game theoretic model can remedy this shortcoming by creating a strategic environment, where prosumers make decisions to maximize their individual profits.

ES systems, being able to quickly switch between charging and discharging, are a key component of distributed flexible resources, and a vital player in a prosumer-centric energy market.

Section 2.2.1 shows that non-cooperative game theoretic schemes are well compatible with ES systems. However, two caveats can be identified with the usage of non-cooperative game theory in prosumer energy sharing. First, optimization dual price variables can be interpreted as competitive energy prices, and it heavily relies on the

assumption that no prosumers have market power. Second, the allocation of the financial benefit to each prosumer is either unclear or suboptimal on the individual level. In other words, it might be possible for a subgroup of participants to abandon the current scheme and develop an alternative that are economically preferable.

Cooperative game theoretic schemes, as shown in Section 2.2.2, strive to allocate profit incentives to prosumer participants fairly and in a manner that prevents defection and strategic behaviors that are detrimental to others. However, few cooperative game theoretic schemes have been able to incorporate ES systems. Chapter 3 presents the thesis author's work published in [134, 135], which first proposed the coalitional energy management scheme based on cooperative game theory. This prosumer-centric market scheme, along with subsequent publications by others ([136,137]) that also proposed similar frameworks, has outstanding computational issues and is unable to scale up to over 15 players.

Two common payoff allocations are used in literature for cooperative game theoretic schemes: the Shapley value and the nucleolus (see Section 2.2.2.1). The Shapley value is known for its fairness and its analytical expression. The nucleolus minimizes the dissatisfaction of all the participants and coalitions of participants within the scheme. However, they both have significant computational issues that are primarily associated with the number of coalitions, an exponential function of the number of participants.

For the Shapley value, coalition sampling has been used in other literature as an approximation method due to the Shapley value's tie with each participant's marginal contribution to the coalitions, but the applicability of these methods has not been evaluated in an energy market context. The computational improvement these methods can offer is also unclear.

The nucleolus, however, is not directly derived from the participants' marginal contributions, but it has close linkage with the participants' consumption and generation profile patterns. Techniques discussed in existing literature can only partially reduce the computation time of the nucleolus and are unable to significantly scale up the game. On the other hand, clustering (see Section 2.3) as a widely used method in categorizing energy profiles has not been used in conjunction with computing prosumer cooperative

games. It is identified here as a promising technique to help alleviate the nucleolus's computational burden.

Chapter 3-5 aim to fill these research gaps. Specifically, Chapter 3 proposes a cooperative game theoretic energy management scheme that schedules prosumer participants' ES systems to minimize their joint energy cost. The scheme ensures financial benefits for every prosumer and is proven to have stabilizing payoff allocations that discourage defection from the grand coalition. Chapter 4 investigates the computation of the Shapley value, discusses its fairness, and develops sampling based methods to address its computational complexity. Chapter 5 focuses on the computation of the nucleolus, a proven stabilizing payoff allocation, and proposes a novel clustering based nucleolus estimation method to improve the scalability of the proposed scheme.

I have cherished the ideal of a democratic and free society in which all persons live together in harmony and with equal opportunities. It is an ideal which I hope to live for and to achieve. But if needs be, it is an ideal for which I am prepared to die.

— Nelson Mandela

3

Cooperative Game Theoretic Energy Market

Contents

3.1 Cooperative Game Framework 42
3.2 Coalitional Energy Management using Cooperative Game Theory 45
3.2.1 Coalitional Energy Cost 46
3.2.2 Optimal Energy Storage Operation 47
3.2.3 Value of Prosumer Coalitions 48
3.2.4 Payoff Allocations and Imputations 49
3.3 Balancedness of the Coalitional Energy Management Scheme . . . 50
3.3.1 Coalition Matrix Operation Definitions 50
3.3.2 Coalitional Energy Management in Matrix Format 51
3.4 Model Implementation 56
3.4.1 Energy Management Optimization 57
3.4.2 Computation of Shapley Value and Nucleolus 58
3.5 Case Studies 60
3.5.1 Load Balancing and Financial Reward 60
3.5.2 Nucleolus vs. Shapley Value 63
3.5.3 The Impact of Uncertainty 65
3.6 Concluding remarks 68

As explained in Chapter 1 and 2, the increasing amount of DER coupled with the development in communication technologies gave rise to the emergence of prosumers, proactive-consumers with distributed energy resources that actively control their energy behaviors. In order to coordinate prosumers to ensure stable and efficient operation of

the energy network, an energy market transition from the traditional centralized model to a prosumer-centric model is needed.

In a system-focused prosumer-centric energy market scheme, such as a microgrid [38, 40–42, 87–90] or a VPP [87, 91–93], achieving the best economic outcome for the community does not guarantee profit for the individual prosumers. Therefore, it raises market stability concerns. In an individual-focused prosumer-centric market, the market scheme could be either a price-taker model or a game theoretic model. The price-taker model [95–100] often fails to offer economic incentives that truly reflect the value of the prosumer's DER contributions, thus introducing unfairness and instability into the market. The game theoretic model, on the other hand, takes into account the strategic behavior of prosumers and aims to offer economic incentives at a level that stabilizes the market.

Non-cooperative game theory based prosumer-centric markets [55–58, 84, 106, 110, 112, 121] seek market equilibria at which no player can benefit more by altering their energy behavior. There are two major caveats with these schemes. First, it lacks protection against market manipulation through collusion or shill bidding [114]. Second, the market equilibrium may be compromised if prosumers are given other market options, such as rejoining the traditional centralized market, or developing a new market within a subgroup of prosumers.

Cooperative game theory based prosumer-centric markets achieves the highest profit for the grand coalition of all prosumers, and investigates the fairness and stability of the payoff allocations. It offers a robust market mechanism that incentivizes prosumer to manage their DER in an organized and economically efficient manner. However, prior to the thesis author's work in [134, 135], no ES systems were included in prosumer-centric market schemes based on cooperative game theory [102, 132, 133].

This chapter presents the cooperative game theory based prosumer energy management scheme proposed in the thesis authors publications [134, 135]. This was the first cooperative game theoretic energy market scheme that incorporated ES systems. The proposed scheme shows that cooperative game theory not only helps maximize the total welfare of the participants, but also guarantees financial benefit for each individual participant. The chapter is organized as follows:

Section 3.1 explains the basic framework of a cooperative game. Section 3.2 establishes the coalitional energy management scheme by applying the cooperative game framework to prosumer energy management. It uses the energy cost savings from cooperation to measure the value of a coalition, and allocates these savings to the prosumers in ways that economically benefit every participant. In Section 3.3, a series of detailed mathematical proofs are conducted to confirm that the coalitional energy management scheme is a balanced game, which means there exist payoff allocations that guarantee no player or coalition of players can be made better off by operating independently from the grand coalition. An explanation of the proposed model implementation is presented in Section 3.4.1. Finally, case study results are shown in Section 3.5 to demonstrate the effectiveness of the proposed scheme, compare different payoff allocations, and analyze the impact of stochastic forecast errors.

3.1 Cooperative Game Framework

The way to analyze a cooperative game is as follows:

The *grand coalition* is defined as the largest coalition in a cooperative game, which encompasses all the players within it. Assume a total of N players, the grand coalition is denoted by $\mathcal{N}$, indexed by $i \in \mathcal{N} := \{1, 2, ..., N\}$. Any subset of $\mathcal{N}$ is defined as a *coalition*. The set of all possible coalitions is defined as the power set $2^{\mathcal{N}}$ of $\mathcal{N}$. In other words, for any coalition $\mathcal{S}, \mathcal{T} \subseteq \mathcal{N}$, it satisfies $\mathcal{S} \in 2^{\mathcal{N}}$. For each coalition $\mathcal{S}$ there is a characteristic function $v(\mathcal{S})$ that represents the value of forming coalition $\mathcal{S}$. Therefore, the *characteristic* function is also called the *value* function. $v : 2^{\mathcal{N}} \to \mathbb{R}$. The pair $(\mathcal{N}, v)$ defines the cooperative game, and the value of the grand coalition is given by $v(\mathcal{N})$.

The goal of cooperating is to create more value than players can individually. Therefore, superadditivity is an essential characteristic of cooperative games.

Definition 3.1 (Superadditive Game) *A cooperative game is said to be superadditive if*

$$v(\mathcal{S}) + v(\mathcal{T}) \leq v(\mathcal{S} \cup \mathcal{T}), \forall \mathcal{S}, \mathcal{T} \in 2^{\mathcal{N}}, \mathcal{S} \cap \mathcal{T} = \emptyset \tag{3.1}$$

The characteristic function indicates the value of cooperating, which dictates the total profit that can be distributed to each player within the coalition. The subsequent step is to determine the specific amount of payoffs to be awarded to each player.

Definition 3.2 (Transferable utility) *A game that allows that allows side payments to be made among the players is called a game with transferable utility [174].*

All the games discussed in this thesis are games with transferable utility .

Definition 3.3 (Imputation) *The payoff allocation is denoted by vector* $\mathbf{x} \in \mathbb{R}^{1\times N}$ *whose entry* x_i *represents the payment to member* $i \in \mathcal{N}$. $\mathbf{x}$ *is said to be an imputation if it satisfies both the Efficiency and Individual Rationality criteria:*

1. *(Efficiency) A payoff allocation* $\mathbf{x}$ *is said to be efficient if* $\sum_{i\in\mathcal{N}} x_i = v(\mathcal{N})$
2. *(Individual Rationality) A payoff allocation* $\mathbf{x}$ *is said to be individually rational if* $x_i \geq v(\{i\}), \forall i \in \mathcal{N}$

Definition 3.4 (Core) *For a cooperative game* $(\mathcal{N}, v)$*, the core* $\mathcal{C}$ *is defined as the set of imputations such that no coalition has a value greater than the sum of the coalition's members' current allocated payoffs:*

$$\mathcal{C} := \{\mathbf{x} \in \mathcal{I} \mid v(\mathcal{S}) \leq \sum_{i\in\mathcal{S}} x_i, \forall \mathcal{S} \in 2^{\mathcal{N}}\} \tag{3.2}$$

An imputation $\mathbf{x}$ is said to be stabilizing if $\mathbf{x} \in \mathcal{C}$ because no player will have the incentives to leave the grand coalition to form smaller coalitions. In other words, no players are *dissatisfied* being in the grand coalition.

Definition 3.5 (Excess) *A coalition's dissatisfaction with respect to the imputation* $\mathbf{x}$ *is measured by its excess defined as*

$$\varepsilon(\mathbf{x}, \mathcal{S}) = v(\mathcal{S}) - \sum_{i\in\mathcal{S}} x_i \tag{3.3}$$

Comparing (3.2) and (3.3) it is noticed that if all excesses are non-positive, the given imputation is in the core. However, some games have empty cores, meaning no matter how payoffs are allocated, there will always be one or more coalitions that are dissatisfied. Two important classes of games with nonempty cores are *convex games* and *balanced games* [61].

Definition 3.6 (Convex Game) *A game is said to be convex if*

$$v(\mathcal{S}) + v(\mathcal{T}) \leq v(\mathcal{S} \cup \mathcal{T}) + v(\mathcal{S} \cap \mathcal{T}), \forall \mathcal{S}, \mathcal{T} \in 2^{\mathcal{N}} \tag{3.4}$$

Therefore, (3.4) can be rewritten as

$$v(\mathcal{S} \cup \{i\}) - v(\mathcal{S}) \leq v(\mathcal{T} \cup \{i\}) - v(\mathcal{T}), \forall i \in \mathcal{N}, \forall \mathcal{S} \subset \mathcal{T} \subset \mathcal{N} - \{i\} \tag{3.5}$$

This indicates an increasing marginal value for coalition members as the coalition grows in size. It can be shown that the core of a convex game is not empty, and that the imputation that lies in the center of a convex game's core is the Shapley Value [126]:

Definition 3.7 (Shapley Value) *The Shapley value ϕ of a cooperative game is given by*

$$\phi_i(v) = \sum_{\mathcal{S} \in 2^{\mathcal{N}},\, i \notin \mathcal{S}} \frac{(|\mathcal{S}|)!(N - |\mathcal{S}| - 1)!}{N!} [v(\mathcal{S} \cup i) - v(\mathcal{S})] \tag{3.6}$$

The Shapley value also satisfies the following axioms:

1) (*Efficiency*) $\sum_{i \in \mathcal{N}} \phi_i(v) = v(\mathcal{N})$.

2) (*Individual Rationality*) $\phi_i(v) \geq v(\{i\}), \forall i \in \mathcal{N}$.

3) (*Symmetry*) If $v(\mathcal{S} \cup \{i\}) = v(\mathcal{S} \cup \{i'\}), \forall \mathcal{S} \in 2^{\mathcal{N}}, \mathcal{S} \cap \{i, i'\} = \emptyset$, then $\phi_i(v) = \phi_{i'}(v)$.

4) (*Dummy Axiom*) If $v(\mathcal{S}) = v(\mathcal{S} \cup \{i\}), \forall \mathcal{S} \in 2^{\mathcal{N}}, \mathcal{S} \cap \{i\} = \emptyset$, then $\phi_i(v) = 0$.

5) (*Additivity*) If v and u are characteristic functions, then $\phi_i(v + u) = \phi_i(v) + \phi_i(u), \forall i \in \mathcal{N}$.

Axiom 1) and 2) qualify Shapley value as an imputation. Axiom 3) says that if the marginal value of two players are the same, then the Shapley values assigned to them should be equal. Axiom 4) says that if the marginal value of a player is zero then its Shapley value should also be zero. Axiom 5) says that the Shapley value of two games played at the same time should be the sum of the two games' Shapley values when played separately.

Definition 3.8 (Balanced Game) *A map $\gamma : 2^{\mathcal{N}} \rightarrow [0, 1]$ is called a balanced map if*

$$\sum_{\mathcal{S} \in 2^{\mathcal{N}}} \gamma(\mathcal{S}) 1_i^{\mathcal{S}} = 1, \forall i \in \mathcal{N} \tag{3.7}$$

where the *indicator function* $1_i^{\mathcal{S}} = \begin{cases} 1, & \text{if } i \in \mathcal{S} \\ 0, & \text{otherwise} \end{cases}$. A cooperative game is said to be *balanced* if for any balanced map γ,

$$\sum_{\mathcal{S} \in 2^{\mathcal{N}}} \gamma(\mathcal{S}) v(\mathcal{S}) \leq v(\mathcal{N}) \tag{3.8}$$

A balanced game always has a nonempty core [127].

Definition 3.9 (Nucleolus) *The nucleolus $\boldsymbol{\nu}$ is an imputation with lexicographically minimal excesses of all the possible coalitions, such that*

$$\psi_s(\boldsymbol{\nu}) \leq \psi_s(\mathbf{x}), \forall s \in [1, 2^N - 2], \forall \mathbf{x} \in \mathcal{I} \tag{3.9}$$

where $\boldsymbol{\psi}(\mathbf{x}) \in \mathbb{R}^{2^N - 2}$ is defined as the vector of excesses in respect to imputation $\mathbf{x}$ for all coalitions (excluding the grand coalition) arranged in a non-increasing order. Its entries are indexed by s (i.e., $\psi_s(\mathbf{x})$ is the sth entry of $\boldsymbol{\psi}(\mathbf{x})$).

It can be proven that for a game with a nonempty core (balanced), the nucleolus always belongs to the core [175].

For a three-person cooperative game, the nucleolus can be solved analytically [176]. However, when the number of players N is large, it takes a sequence of $\mathcal{O}(2^N)$ linear problems to solve for the nucleolus [142], as it has to iterate through every element in $\psi(\mathbf{x})$.

3.2 Coalitional Energy Management using Cooperative Game Theory

For a prosumer who owns a distributed PV system and an ES system, the common objective in scheduling the ES operation is to minimize the energy cost. Because the retail supplier normally offers a lower export tariff (e.g. feed-in tariff) than it charges for buying electricity, prosumers can save money by storing the excess PV generation in the ES system and using it when the generation is less than they need.

Considering a group of prosumers who have different load profiles and distributed energy resources, when they each manage their own ES system to minimize their energy bill, the joint operation of all the ES systems is unlikely to yield the minimum total energy cost for the group. However, it is possible for prosumers cooperatively operating their ES systems to minimize the total group energy cost, the additional savings from which can then be allocated among the prosumers.

This section proposes the concept of an *prosumer coalition*, a group of prosumers who collaboratively operate their ES systems to minimize the total group energy cost. A cooperative game theoretic approach is then used to analyze methods to distribute the energy cost savings among the prosumers.

3.2.1 Coalitional Energy Cost

Following the framework of cooperative game theory, let a group of N prosumers form grand coalition $\mathcal{N}$ indexed by $i \in \mathcal{N} := \{1, 2, ..., N\}$. A prosumer coalition is any subset $\mathcal{S} \subseteq \mathcal{N}$. The set of all possible prosumer coalitions is the power set $2^{\mathcal{N}}$ of $\mathcal{N}$. In other words, $\mathcal{S} \in 2^{\mathcal{N}}, \forall \mathcal{S} \subseteq \mathcal{N}$. In the rest of the thesis, all the *coalitions* refer to *prosumer coalitions*.

Each prosumer i is assumed to have no more than one ES system. Considering R timesteps ($t = 1, 2, ..., R$) with a time interval of Δt, the following terms are defined:

q_{it} : Prosumer i's net energy consumption (positive) and generation (negative) without ES at timestep t, (kWh)

b_{it} : Prosumer i's ES system operation variables (positive for charging and negative for discharging) at timestep t, (kWh)

r_t^{im} : Electricity buy price at timestep t, (£/kWh)

r_t^{ex} : Electricity sell price at timestep t, (£/kWh)

The remuneration for exporting distributed generation has been consistently declining in recent years, and here it is assumed that $r_t^{im} > r_t^{ex}, \forall t \in [1, R]$, as is already the case in countries of relatively high penetrations of distributed generation such as the UK [28, 29],

and Germany [177]. It is also assumed that the electricity price can vary by each timestep but are required to be known inputs before each modeling cycle (e.g. day-ahead).

The load consumption and generation of all prosumers are assumed to be known inputs, but they are not restricted to a specific type of load profile (e.g. residential, commercial, industrial, etc.). In reality, the model will be using predicted consumption and generation data, and the impact of uncertainties and prediction errors is analyzed in the case studies.

For coalition $\mathcal{S} \in 2^{\mathcal{N}}$, the total energy cost is the sum of energy costs of all the prosumers within $\mathcal{S}$. The controllable loads are represented by the ES operation variables $\mathbf{b}^{\mathcal{S}} = [b_{it}], \forall i \in \mathcal{S}, \forall t \in [1, R]$. Therefore, the *coalitional energy cost* for $\mathcal{S}$ at timestep t, denoted as $F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}})$, can be written as

$$F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}}) = \sum_{i \in \mathcal{S}} \left\{ r_t^{im}[q_{it} + b_{it}]_+ + r_t^{ex}[q_{it} + b_{it}]_- \right\} \tag{3.10}$$

where $[y]_{+(-)} = \max(\min)\{y, 0\}$.

3.2.2 Optimal Energy Storage Operation

Given a time span R, the operation of ES units within coalition $\mathcal{S}$ can be scheduled to minimize the coalitional energy cost, which is defined as $G^R(\mathcal{S})$:

$$G^R(\mathcal{S}) = \min_{\mathbf{b}^{\mathcal{S}}} \sum_{t=1}^{R} F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}}) \tag{3.11}$$

In order to compute the minimized coalitional energy cost for $\mathcal{S} \in 2^{\mathcal{N}}$ expressed in (3.11), the ES operation variables $\mathbf{b}^{\mathcal{S}}$ need to be constrained properly. Therefore, each prosumer's ES system is assigned an *energy capacity* of $e_i \in \mathbb{R}_{\geq 0}$, in (kWh), charging and discharging limits over the time span of Δt of $\overline{b}_i \in \mathbb{R}_{\geq 0}, \underline{b}_i \in \mathbb{R}_{\leq 0}$, in (kWh), charging and discharging efficiencies of $\eta_i^{in}, \eta_i^{out} \in (0, 1)$, *state of charge* upper and lower limits of $\overline{SoC}_i, \underline{SoC}_i \in [0, 1]$, and an *initial state of charge* of $SoC_i^0 \in [\underline{SoC}_i, \overline{SoC}_i]$. The ES operation variables $\mathbf{b}^{\mathcal{S}}$ are subject to:

1. *Power constraint:* the charge and discharge energy for each timestep is constrained by the charging and discharging limits:

$$\underline{b}_i \leq b_{it} \leq \overline{b}_i, \quad \forall i \in \mathcal{S}, \forall t \in [1, R]$$

We set $\underline{b}_i = \overline{b}_i = 0$ for prosumers that do not own an ES system.

2. *Energy constraint:* each ES system's *state of charge* is limited between its upper and lower limits for every timestep:

$$e_i\underline{SoC}_i \leq e_i SoC_i^0 + \sum_{t=1}^{r}([b_{it}]_+\eta_i^{in} + [b_{it}]_-/\eta_i^{out}) \leq e_i\overline{SoC}_i, \forall i \in \mathcal{S}, \forall r \in [1, R]$$

3. *Cycle constraint:* the state of charge at $t = R$ is required to be restored to SoC_i^0 for every ES system:

$$\sum_{t=1}^{R}([b_{it}]_+\eta_i^{in} + [b_{it}]_-/\eta_i^{out}) = 0, \;\; \forall i \in \mathcal{S}$$

The ES model used in this thesis can be used to represent a range of ES devices. It can also simulate the shift of flexible loads in time to an extent but is limited in its capability in capturing the non-economic prosumer welfare, such as comfort and satisfaction associated with consuming energy at preferred times. Modeling demand side flexible services with a linear model is a common practice in smart grid applications [33], but the ES constraints presented above need to be expanded in order to account for the specific operation cycles and discrete power output of flexible loads.

3.2.3 Value of Prosumer Coalitions

A cooperative game analyzes situations where additional values can be achieved when prosumers operate collaboratively. Here, the standard way of analyzing a cooperative game is applied on the N-prosumer coalitional energy management scheme.

For each coalition $\mathcal{S}$ there is a value function $v(\mathcal{S}) : 2^{\mathcal{N}} \rightarrow \mathbb{R}$ that represents the value of $\mathcal{S}$. This value is defined as the energy cost savings by forming $\mathcal{S}$, which is the difference between the sum of the minimum energy costs of each prosumer in $\mathcal{S}$ by optimizing their ES systems individually, and the minimum coalitional energy cost of $\mathcal{S}$ by optimizing the ES systems cooperatively:

$$v(\mathcal{S}) = \sum_{i \in \mathcal{S}} G^R(\{i\}) - G^R(\mathcal{S}) \tag{3.12}$$

The pair $(\mathcal{N}, v)$ denotes the coalitional energy management scheme, and the value of the grand coalition $v(\mathcal{N})$ denotes the total profit, or the total amount of payoffs that can be divided among the prosumers.

3.2.4 Payoff Allocations and Imputations

Having calculated the total profit, the next step is to determine how to allocate it to each prosumer.

Vector $\mathbf{x} = [x_1, ..., x_N]$ is defined as the prosumer payoff allocation whose entry x_i represents the payment to prosumer $i \in \mathcal{N}$. Based on Definition 3.3, $\mathbf{x}$ is said to be an *imputation* if it meets both the Efficiency and Individual Rationality criteria:

1. (*Efficiency*) $\sum_{i \in \mathcal{N}} x_i = v(\mathcal{N})$. It means the entirety of the grand coalition's energy cost savings is allocated to the prosumers.

2. (*Individual Rationality*) $x_i \geq v(\{i\}), \forall i \in \mathcal{N}$. Because the value function represents the energy cost savings from forming coalitions, $v(\{i\}) = 0, \forall i \in \mathcal{N}$. Therefore, this criterion guarantees that no prosumer is losing money by joining the grand coalition.

The set of all imputations for the coalitional energy management scheme is defined as

$$\mathcal{I} := \{\mathbf{x} \in \mathbb{R}^N \mid \sum_{i \in \mathcal{N}} x_i = v(\mathcal{N}),\ x_i \geq v(\{i\}), \forall i \in \mathcal{N}\} \tag{3.13}$$

However, an imputation does not guarantee every prosumer being satisfied in the grand coalition, as some prosumers may be able to achieve higher payoffs by forming smaller coalitions.

A coalition's dissatisfaction with respect to the imputation $\mathbf{x}$ is measured by its excess $\varepsilon(\mathbf{x}, \mathcal{S})$ (Definition 3.5):

$$\varepsilon(\mathbf{x}, \mathcal{S}) = v(\mathcal{S}) - \sum_{i \in \mathcal{S}} x_i \tag{3.14}$$

When $\varepsilon(\mathbf{x}, \mathcal{S}) > 0$, it indicates that $\mathcal{S}$ becoming independent of the grand coalition can offer higher payoffs to its members.

A payoff allocation $\mathbf{x}$ is said to be *stabilizing* if it is in the game's *core* (Definition 3.4):

$$\mathcal{C} := \{\mathbf{x} \in \mathcal{I} \mid \varepsilon(\mathbf{x}, \mathcal{S}) \leq 0, \forall \mathcal{S} \in 2^{\mathcal{N}}\} \tag{3.15}$$

However, some games have empty cores, meaning no matter how payoffs are allocated, there will always be one or more coalitions that are dissatisfied.

3.3 Balancedness of the Coalitional Energy Management Scheme

A balanced cooperative game has a nonempty core [127], and it has been proven that for a balanced cooperative game, the nucleolus is guaranteed to be in its nonempty core [175]. In order to prove that the coalitional energy management scheme is a balanced cooperative game, or a cooperative game with a nonempty core, the mathematical expressions used in Section 3.2 are reconstructed in matrices in this section. The matrix format is only used in this section to facilitate the operation and transformation of mathematical terms, and the nonmatrix format is used in the rest of the thesis for ease of understanding.

3.3.1 Coalition Matrix Operation Definitions

Recall that $\mathcal{N}$ is the grand coalition, and the set of all possible coalitions is defined as the power set $2^{\mathcal{N}}$ of $\mathcal{N}$. For any coalition $\mathcal{S} \subseteq \mathcal{N}$, it satisfies $\mathcal{S} \in 2^{\mathcal{N}}$.

For any matrix $\mathbf{Z} \in \mathbb{R}^{R\times N} := [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N]$, where $\mathbf{z}_i \in \mathbb{R}^{R\times 1}, \forall i \in [1, N]$, the following operations are defined using coalition superscripts: where $\mathcal{S}_h \in 2^{\mathcal{N}},$ and $\lambda_h \geq 0,$ indexed by $h = 1, 2, ..., 2^N$

1. *coalition scaling:*

$$\begin{aligned}\mathbf{Z}^{\lambda_h \mathcal{S}_h} &= [\lambda_h \mathbf{z}_1 1_1^{\mathcal{S}_h}, \lambda_h \mathbf{z}_2 1_2^{\mathcal{S}_h}, ..., \lambda_h \mathbf{z}_N 1_N^{\mathcal{S}_h}] \qquad (3.16)\\ &= \lambda_h [\mathbf{z}_1 1_1^{\mathcal{S}_h}, \mathbf{z}_2 1_2^{\mathcal{S}_h}, ..., \mathbf{z}_N 1_N^{\mathcal{S}_h}]\\ &= \lambda_h \mathbf{Z}^{\mathcal{S}_h} \qquad (3.17)\end{aligned}$$

When $\lambda_h = 1, \mathcal{S}_h = \mathcal{S}$, there is

$$\mathbf{Z}^{\mathcal{S}} = [\mathbf{z}_1 1_1^{\mathcal{S}}, \mathbf{z}_2 1_2^{\mathcal{S}}, ..., \mathbf{z}_N 1_N^{\mathcal{S}}] \qquad (3.18)$$

For a special case where $\lambda_h = 1, \mathcal{S}_h = \mathcal{N}$, because $1_i^{\mathcal{N}} = 1, \forall i$, there is

$$\mathbf{Z}^{\mathcal{N}} = [\mathbf{z}_1 1_1^{\mathcal{N}}, \mathbf{z}_2 1_2^{\mathcal{N}}, ..., \mathbf{z}_N 1_N^{\mathcal{N}}] = [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N] \qquad (3.19)$$

2. *coalition summation:*

$$\mathbf{Z}^{\sum_{h=1}^{2^N}\lambda_h\mathcal{S}_h} = [\sum_{h=1}^{2^N}\lambda_h\mathbf{z}_1 1_1^{\mathcal{S}_h}, \sum_{h=1}^{2^N}\lambda_h\mathbf{z}_2 1_2^{\mathcal{S}_h}, ..., \sum_{h=1}^{2^N}\lambda_h\mathbf{z}_N 1_N^{\mathcal{S}_h}]$$

from which the following can be derived.

$$\begin{aligned}\mathbf{Z}^{\sum_{h=1}^{2^N}\lambda_h\mathcal{S}_h} &= \sum_{h=1}^{2^N}[\lambda_h\mathbf{z}_1 1_1^{\mathcal{S}_h}, \lambda_h\mathbf{z}_2 1_2^{\mathcal{S}_h}, ..., \lambda_h\mathbf{z}_N 1_N^{\mathcal{S}_h}] \\ &\overset{(3.16)}{=} \sum_{h=1}^{2^N}\mathbf{Z}^{\lambda_h\mathcal{S}_h} \overset{(3.17)}{=} \sum_{h=1}^{2^N}\lambda_h\mathbf{Z}^{\mathcal{S}_h}\end{aligned} \tag{3.20}$$

For a special case where $\lambda_h = \begin{cases}1, & \mathcal{S}_h = \mathcal{S} \text{ or } \mathcal{T} \\ 0, & \text{otherwise}\end{cases}$, and $\mathcal{S} \cap \mathcal{T} = \emptyset$, there is $1_i^{\mathcal{S}} + 1_i^{\mathcal{T}} = 1_i^{\mathcal{S}\cup\mathcal{T}}$, so

$$\begin{aligned}\mathbf{Z}^{\mathcal{S}+\mathcal{T}} &= [\mathbf{z}_1(1_1^{\mathcal{S}} + 1_1^{\mathcal{T}}), \mathbf{z}_2(1_2^{\mathcal{S}} + 1_2^{\mathcal{T}}), ..., \mathbf{z}_N(1_N^{\mathcal{S}} + 1_N^{\mathcal{T}})] \\ &= [\mathbf{z}_1 1_1^{\mathcal{S}\cup\mathcal{T}}, \mathbf{z}_2 1_2^{\mathcal{S}\cup\mathcal{T}}, ..., \mathbf{z}_N 1_N^{\mathcal{S}\cup\mathcal{T}}] = \mathbf{Z}^{\mathcal{S}\cup\mathcal{T}}\end{aligned} \tag{3.21}$$

3.3.2 Coalitional Energy Management in Matrix Format

Given R timesteps ($t = 1, 2, ..., R$) with a time interval of Δt, the following parameters are reconstructed as vectors:

$\mathbf{q}_i \in \mathbb{R}^{R\times 1}$: Prosumer i's net energy consumption (positive) and generation (negative) without ES, (kWh)

$\mathbf{r}^{im} \in \mathbb{R}^{1\times R}$: Electricity buy price (import price), $(£/kWh)$

$\mathbf{r}^{ex} \in \mathbb{R}^{1\times R}$: Electricity sell price (export price), $(£/kWh)$

$\mathbf{b}_i \in \mathbb{R}^{R\times 1}$: Prosumer i's ES energy charge (positive) and discharge (negative) variables, (kWh)

where $r_t^{im} > r_t^{ex}, \forall t \in [1, R]$

With the grand coalition $\mathcal{N}$ indexed by prosumer i, the *net energy load* and *ES operation variable* for the grand coalition are reconstructed as matrices:

$$\mathbf{Q}^{\mathcal{N}} \in \mathbb{R}^{R\times N} := [\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_N]$$

$$\mathfrak{B} \in \mathbb{R}^{R\times N} := [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_N]$$

Transformed with (3.18), $\mathbf{Q}^{\mathcal{S}} = [\mathbf{q}_1 1_1^{\mathcal{S}}, \mathbf{q}_2 1_2^{\mathcal{S}}, ..., \mathbf{q}_N 1_N^{\mathcal{S}}]$. Notice that any column i in $\mathbf{Q}^{\mathcal{S}}$ corresponding to prosumer i that does not belong to $\mathcal{S}$ is set to zero.

The coalitional energy cost for any net energy load and ES operation variable matrices $\mathbf{Q}, \mathfrak{B} \in \mathbb{R}^{K\times N}$ are defined as:

$$F(\mathfrak{B}, \mathbf{Q}) = \mathbf{r}^{im}[(\mathfrak{B} + \mathbf{Q})\mathbf{1}^N]_+ + \mathbf{r}^{ex}[(\mathfrak{B} + \mathbf{Q})\mathbf{1}^N]_- \tag{3.22}$$

Therefore the energy cost function for any coalition $\mathcal{S}$ can then be written as

$$F(\mathfrak{B}, \mathbf{Q}^{\mathcal{S}}) = \mathbf{r}^{im}[(\mathfrak{B} + \mathbf{Q}^{\mathcal{S}})\mathbf{1}^N]_+ + \mathbf{r}^{ex}[(\mathfrak{B} + \mathbf{Q}^{\mathcal{S}})\mathbf{1}^N]_- \tag{3.23}$$

where $\mathbf{1}^N = [\underbrace{1, 1, ..., 1}_{N}]^\intercal$, and for any vector or matrix $\mathbf{Y}$, $[\mathbf{Y}]_{mn}^{+(-)} = \max(\min)\{Y_{mn}, 0\}$. It is obvious that scaling $\mathbf{Q}$ and $\mathfrak{B}$ by a factor of $\lambda \geq 0$, the energy cost is also scaled by a factor of λ:

$$F(\lambda\mathfrak{B}, \lambda\mathbf{Q}) = \lambda F(\mathfrak{B}, \mathbf{Q}) \tag{3.24}$$

Note that the net energy load matrix $\mathbf{Q}$ is superscripted and the ES operation variable matrix $\mathfrak{B}$ is not. This is because the net energy loads are known parameters, and all elements in $\mathbf{Q}^{\mathcal{S}}$ can be calculated by applying (3.18) directly to the input parameters, whereas $\mathfrak{B}$ is a matrix containing variables that can take on any values, which will be constrained later to make sure that only prosumers within coalition $\mathcal{S}$ have the capacity to operate their ES systems.

As $\mathfrak{B}$ contain all the variables in (3.23), the *coalitional energy cost function* $C_{\mathcal{S}}(\mathfrak{B})$ can be defined as

$$C_{\mathcal{S}}(\mathfrak{B}) = F(\mathfrak{B}, \mathbf{Q}^{\mathcal{S}}) \tag{3.25}$$

Therefore, the value of coalition $\mathcal{S}$ can be written as

$$v(\mathcal{S}) = \sum_{i\in\mathcal{S}} C_{\{i\}}(\mathfrak{B}^{\langle\{i\}\rangle *}) - C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle *}) \tag{3.26}$$

The ES system parameters are reconstructed as matrices as well. For the *grand coalition* $\mathcal{N}$, the energy capacity $\mathbf{E}^{\mathcal{N}}$, charge/discharge energy limits $\overline{\mathbf{B}}^{\mathcal{N}}$/$\underline{\mathbf{B}}^{\mathcal{N}}$, charging/discharging efficiencies $\boldsymbol{\eta}^{\boldsymbol{I}}$/$\boldsymbol{\eta}^{\boldsymbol{O}}$, state of charge upper/lower limits $\overline{\mathbf{SoC}}$/$\underline{\mathbf{SoC}}$ and

the initial state of charge $\mathbf{SoC^0}$ are written as

$$\begin{aligned}
\mathbf{E}^{\mathcal{N}} \in \mathbb{R}^{R\times N} &:= [e_1\mathbf{1}^R, e_2\mathbf{1}^R, ..., e_N\mathbf{1}^R] \\
\overline{\mathbf{B}}^{\mathcal{N}} \in \mathbb{R}^{R\times N} &:= [\overline{b}_1\mathbf{1}^R, \overline{b}_2\mathbf{1}^R, ..., \overline{b}_N\mathbf{1}^R] \\
\underline{\mathbf{B}}^{\mathcal{N}} \in \mathbb{R}^{R\times N} &:= [\underline{b}_1\mathbf{1}^R, \underline{b}_2\mathbf{1}^R, ..., \underline{b}_N\mathbf{1}^R] \\
\boldsymbol{\eta^I} \in \mathbb{R}^{N\times N} &:= \mathrm{diag}(\eta_1^{in}, \eta_2^{in}, ..., \eta_N^{in}) \\
\boldsymbol{\eta^O} \in \mathbb{R}^{N\times N} &:= \mathrm{diag}(1/\eta_1^{out}, 1/\eta_2^{out}, ..., 1/\eta_N^{out}) \\
\overline{\mathbf{SoC}} \in \mathbb{R}^{N\times N} &:= \mathrm{diag}(\overline{SoC}_1, \overline{SoC}_2, ..., \overline{SoC}_N) \\
\underline{\mathbf{SoC}} \in \mathbb{R}^{N\times N} &:= \mathrm{diag}(\underline{SoC}_1, \underline{SoC}_2, ..., \underline{SoC}_N) \\
\mathbf{SoC^0} \in \mathbb{R}^{N\times N} &:= \mathrm{diag}(SoC_1^0, SoC_2^0, ..., SoC_N^0)
\end{aligned}$$

Similarly, transformed using the operation defined in (3.18), $\mathbf{E}^{\mathcal{S}}$, $\overline{\mathbf{B}}^{\mathcal{S}}$ and $\underline{\mathbf{B}}^{\mathcal{S}}$ become the energy capacity, charging limit and discharging limit matrices for any coalition $\mathcal{S}$.

An ES operation variable matrix $\mathfrak{B}^{\langle\mathcal{S}\rangle}$ is defined as a *feasible ES operation matrix for* $\mathcal{S}$, thus subject to the ES operation constraints which can be rewritten in a matrix format:

1. *Power constraint:* the charge and discharge energy for each timestep is constrained by the charging and discharging limit:

$$\underline{\mathbf{B}}^{\mathcal{S}} \leq \mathfrak{B}^{\langle\mathcal{S}\rangle} \leq \overline{\mathbf{B}}^{\mathcal{S}} \tag{3.27}$$

   Matrix inequality in this section is established only between matrices of the same size such that the inequality specified is satisfied by every pair of elements of the same index on the two sides of the inequality sign. For example, (3.27) is equivalent to $\underline{B}_{ti}^{\mathcal{S}} \leq B_{ti}^{\langle\mathcal{S}\rangle} \leq \overline{B}_{ti}^{\mathcal{S}}, \forall t, \forall i$. Notice that $B_{ti}^{\langle\mathcal{S}\rangle} = 0, \forall i \notin \mathcal{S}, \forall t$, because $\underline{B}_{ti}^{\mathcal{S}} = \overline{B}_{ti}^{\mathcal{S}} = 0, \forall i \notin \mathcal{S}$ based on (3.18).

2. *Energy constraint:* each ES system's state of charge is limited between its upper and lower limits for every timestep:

$$\mathbf{E}^{\mathcal{S}}\underline{\mathbf{SoC}} \leq \mathbf{E}^{\mathcal{S}}\mathbf{SoC^0} + \boldsymbol{\Lambda}^R([\mathfrak{B}^{\langle\mathcal{S}\rangle}]_+\boldsymbol{\eta^I} + [\mathfrak{B}^{\langle\mathcal{S}\rangle}]_-\boldsymbol{\eta^O}) \leq \mathbf{E}^{\mathcal{S}}\overline{\mathbf{SoC}} \tag{3.28}$$

   where $\boldsymbol{\Lambda}^R = [\lambda_{m'n'}] \in \mathbb{R}^{R\times R} :\ \lambda_{m'n'} = \begin{cases} 1, & m' \geq n' \\ 0, & m' < n' \end{cases}$

3. *Cycle constraint:* All of the ES systems' states of charge at $t = R$ are required to be restored to $\mathbf{SoC^0}$:

$$(\mathbf{1}^R)^\intercal([\mathfrak{B}^{\langle\mathcal{S}\rangle}]_+\boldsymbol{\eta}^{\boldsymbol{I}} + [\mathfrak{B}^{\langle\mathcal{S}\rangle}]_-\boldsymbol{\eta}^{\boldsymbol{O}}) = \mathbf{0} \tag{3.29}$$

3.3.2.1 Proof of a Nonempty Core

Constraints 1–3 are denoted by $D(\mathfrak{B}^{\langle\mathcal{S}\rangle}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}}) = 1$, where $D(\mathfrak{B}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}})$ with $\mathfrak{B}$ as the variable is defined as

$$D(\mathfrak{B}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}}) = \begin{cases} 1, & \text{constraints 1–3 are all satisfied.} \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 3.1

$$D(\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}, \lambda\mathbf{E}^{\mathcal{S}}, \lambda\underline{\mathbf{B}}^{\mathcal{S}}, \lambda\overline{\mathbf{B}}^{\mathcal{S}}) = 1, \forall\lambda > 0$$

Proof: See Appendix A.1.

To compute the *optimal ES operation* a functional θ is constructed:

$$\begin{aligned} \theta[\mathcal{Y}] =& \min_{\mathfrak{B}\in\mathbb{R}^{R\times N}} F(\mathfrak{B}, \mathbf{Q}^{\mathcal{Y}}) \;\; s.t.\; D(\mathfrak{B}, \mathbf{E}^{\mathcal{Y}}, \underline{\mathbf{B}}^{\mathcal{Y}}, \overline{\mathbf{B}}^{\mathcal{Y}}) = 1 \\ =& \min_{\mathfrak{B}\in\mathfrak{O}_{\mathcal{Y}}} F(\mathfrak{B}, \mathbf{Q}^{\mathcal{Y}}) \end{aligned} \tag{3.30}$$

where $\mathcal{Y}$ denotes any first degree polynomial of all the coalitions with non-negative coefficients, and $\mathfrak{O}_{\mathcal{Y}}$ denotes the set of ES operation variable matrices that satisfy all the ES constraints given the coalition combination of $\mathcal{Y}$. They can be expressed as

$$\mathcal{Y} \in \{\sum_{h=1}^{2^N} \lambda_h\mathcal{S}_h \;\mid\; \mathcal{S}_h \in 2^{\mathcal{N}},\; \lambda_h \geq 0\} \tag{3.31}$$

$$\mathfrak{O}_{\mathcal{Y}} = \{\mathfrak{B} \in \mathbb{R}^{R\times N} \mid D(\mathfrak{B}, \mathbf{E}^{\mathcal{Y}}, \underline{\mathbf{B}}^{\mathcal{Y}}, \overline{\mathbf{B}}^{\mathcal{Y}}) = 1\} \tag{3.32}$$

Notice that for $\mathcal{S} \in 2^{\mathcal{N}}$, $\theta[\mathcal{S}]$ represents its minimum coalitional energy cost:

$$\theta[\mathcal{S}] = G^R(\mathcal{S}) \tag{3.33}$$

Lemma 3.2

$$\theta[\sum_{h=1}^{2^N} \lambda_h\mathcal{S}_h] = \theta[\sum_{i\in\mathcal{N}} \sum_{h=1}^{2^N} \lambda_h 1_i^{\mathcal{S}_h}\{i\}], \mathcal{S}_h \in 2^{\mathcal{N}}, \lambda_h \geq 0$$

Proof: See Appendix A.2.

$\mathfrak{B}^{\langle\mathcal{Y}\rangle*}$ is denoted as the *optimal decision variable matrix* to (3.30). Therefore,

$$C_{\mathcal{Y}}(\mathfrak{B}^{\langle\mathcal{Y}\rangle*}) = F(\mathfrak{B}^{\langle\mathcal{Y}\rangle*}, \mathbf{Q}^{\mathcal{Y}}) = \theta[\mathcal{Y}] \tag{3.34}$$

The minimization process provides that

$$C_{\mathcal{Y}}(\mathfrak{B}^{\langle\mathcal{Y}\rangle*}) \leq C_{\mathcal{Y}}(\mathfrak{B}^{\langle\mathcal{Y}\rangle}),\ \forall\mathfrak{B}^{\langle\mathcal{Y}\rangle} \in \mathfrak{O}_{\mathcal{Y}} \tag{3.35}$$

Lemma 3.3 *For $\mathcal{S}, \mathcal{T} \in 2^{\mathcal{N}}$ and scalars $\lambda_1, \lambda_2 \geq 0$, the functional θ has the following properties:*

1. *(Positive Homogeneity) $\theta[\lambda_1\mathcal{S}] = \lambda_1\theta[\mathcal{S}]$, based on which it can also be proven that $\theta[\lambda_1\mathcal{S}] + \theta[\lambda_2\mathcal{S}] = \theta[(\lambda_1 + \lambda_2)\mathcal{S}]$.*
2. *(Subadditivity) $\theta[\lambda_1\mathcal{S}] + \theta[\lambda_2\mathcal{T}] \geq \theta[\lambda_1\mathcal{S} + \lambda_2\mathcal{T}]$*

Proof: See Appendix A.3.

Note that the *subadditivity* of θ also guarantees the superadditivity of this prosumer cooperative game:

Proof: Rewriting (3.1) by replacing $v(\mathcal{S})$ and $v(\mathcal{T})$ with the value definition in (3.12):

$$\sum_{i\in\mathcal{S}} G^R(\{i\}) - G^R(\mathcal{S}) + \sum_{i\in\mathcal{T}} G^R(\{i\}) - G^R(\mathcal{T}) \leq \sum_{i\in(\mathcal{S}\cup\mathcal{T})} G^R(\{i\}) - G^R(\mathcal{S}\cup\mathcal{T}) \tag{3.36}$$

Since $\mathcal{S} \cap \mathcal{T} = \emptyset$, there is

$$\sum_{i\in\mathcal{S}} G^R(\{i\}) + \sum_{i\in\mathcal{T}} G^R(\{i\}) = \sum_{i\in(\mathcal{S}\cup\mathcal{T})} G^R(\{i\}) \tag{3.37}$$

Therefore, (3.36) is equivalent to

$$G^R(\mathcal{S}) + G^R(\mathcal{T}) \geq G^R(\mathcal{S}\cup\mathcal{T}) \tag{3.38}$$

Let $\lambda_1 = \lambda_2 = 1$. Based on the *subadditivity* of θ in **Lemma 3.3**:

$$\theta[\mathcal{S}] + \theta[\mathcal{T}] \geq \theta[\mathcal{S} + \mathcal{T}] \tag{3.39}$$

Given (3.33), (3.38) is proven, hence proving the superadditivity of the prosumer cooperative game.

Theorem 3.4 *The coalitional energy management scheme* $(\mathcal{N}, v)$ *is a* balanced *cooperative game.*

Proof: See Appendix A.4.

Lemma 3.5 *The nucleolus for the coalitional energy management scheme is a stabilizing payoff allocation.*

Proof: It has been proven in [175] that for a balanced cooperative game, the nucleolus is guaranteed to be in its nonempty core. According to Theorem 3.4, the coalitional energy management scheme is a balanced game. Therefore, the nucleolus of the coalitional energy management scheme is in its nonempty core. Based on the definition of a cooperative game's core in Definition 3.4, any payoff allocation that is in the game's nonempty core is said to be stabilizing. Therefore, the nucleolus is a stabilizing payoff allocation for the coalitional energy management scheme.

3.4 Model Implementation

To implement this coalitional energy management model, the following steps need to be followed.

1. Collect model inputs, including the number of prosumers, each prosumer's load and generation profiles, ES information specified in Section 3.2.2, and the electricity buy and sell prices.

2. Generate a list of all possible prosumer coalitions ($2^{\mathcal{N}}$), and optimize the coalitional ES operation to minimize each coalitional energy cost, calculate the value of each coalition using (3.12). This step is further explained in Section 3.4.1.

3. Compute the desired imputation (e.g. Shapley value, nucleolus) using the coalition values calculated in the previous step, and examine its excesses to determine whether the imputation is stabilizing. This step is further explained in Section 3.4.2.

3.4.1 Energy Management Optimization

For coalition $\mathcal{S}$, its coalitional energy cost $F^{\mathcal{S}}$, as defined in (3.10), is a function of the prosumers' ES charge/discharge energies at each timestep within $\mathcal{S}$, denoted by $\mathbf{b}^{\mathcal{S}}$. The objective of the coalitional energy management optimization is to minimize $F^{\mathcal{S}}$ using (3.11), subject to the ES systems' *power constraints*, *energy constraints*, and *cycle constraints*.

Due to the difference between the import and export energy prices, there is a piecewise term in the coalitional energy cost function in (3.10). Adapting a method from [178] to avoid this piecewise representation while maintaining the linearity of the model, the net energies of all the prosumers can be written into a positive part $\mathbf{L}^{+}$ and a non-positive part $\mathbf{L}^{-}$, constrained by other power inputs (loads, generations) and variables (battery charge/discharge).

Similarly, to obtain a linear energy management optimization problem, a piecewise linear ES model with constant charging and discharging efficiencies is commonly used [179]. Here, the discharge and charge energies are defined as $\mathbf{b}^{\mathcal{S}^-}$ and $\mathbf{b}^{\mathcal{S}^+}$ respectively, to avoid using the binary variable [180] while still maintaining the model's linearity.

The energy management optimization problem can be summarized by rewriting (3.10) and (3.11) in a linear format:

$$
\begin{aligned}
G^R(\mathcal{S}) = \min_{\substack{\mathbf{b}^{\mathcal{S}^+},\mathbf{b}^{\mathcal{S}^-},\\ \mathbf{L}^{+},\mathbf{L}^{-}}} \; & \sum_{t=1}^{R}\left\{ r_t^{im}\sum_{i\in\mathcal{S}} L_{it}^{+} + r_t^{ex}\sum_{i\in\mathcal{S}} L_{it}^{-}\right\} \\
s.t. \quad & L_{it}^{-} \le 0 \le L_{it}^{+} && (3.40)\\
& b_{it}^{+} + b_{it}^{-} + q_{it} \le L_{it}^{+} && (3.41)\\
& b_{it}^{+} + b_{it}^{-} + q_{it} = L_{it}^{+} + L_{it}^{-} && (3.42)\\
& 0 \le b_{it}^{+} \le \overline{b}_i && (3.43)\\
& \underline{b}_i \le b_{it}^{-} \le 0 && (3.44)\\
& e_i\underline{SoC}_i \le e_i SoC_i^0 + \sum_{t=1}^{r}(b_{it}^{+}\eta_i^{in} + b_{it}^{-}/\eta_i^{out}) \le e_i\overline{SoC}_i && (3.45)\\
& \sum_{t=1}^{R}(b_{it}^{+}\eta_i^{in} + b_{it}^{-}/\eta_i^{out}) = 0 && (3.46)\\
& \forall i \in \mathcal{S}, \forall t \in [1,R], \forall r \in [1,R]
\end{aligned}
$$

Because $\eta_i^{in}, \eta_i^{out} < 1, \forall i$, the cost minimization will prevent b_{it}^+ and b_{it}^- from taking on non-zero values at the same time, thus effectively setting b_{it}^+ and b_{it}^- as charge and discharge energies: $b_{it}^+ = [b_{it}]_+$ and $b_{it}^- = [b_{it}]_-$. Similarly, because $r_t^{im} > r_t^{ex}, \forall t$, the cost minimization will prevent L_{it}^+ and L_{it}^- from taking on non-zero values at the same time. Therefore, constraints (3.40), (3.41) and (3.42) set $L_{it}^+ = [b_{it}^+ + b_{it}^- + q_{it}]_+$ and $L_{it}^- = [b_{it}^+ + b_{it}^- + q_{it}]_-$. The new power constraint is denoted by (3.43) and (3.44), the new energy constraint (3.45), and the new cycle constraint (3.46).

3.4.2 Computation of Shapley Value and Nucleolus

This thesis investigates two common payoff allocations for cooperative games: the Shapley value and the nucleolus.

After a complete list of coalition values are generated using (3.12), the Shapley value can be easily calculated using Definition 3.7:

$$\phi_i = \sum_{\mathcal{S} \in 2^{\mathcal{N}}, i \in \mathcal{S}} \frac{(|\mathcal{S}| - 1)!(N - |\mathcal{S}|)!}{N!} [v(\mathcal{S}) - v(\mathcal{S} \setminus \{i\})] \tag{3.47}$$

A prosumer's Shapley value represents the mean of the prosumer's marginal contributions to all possible coalitions and is considered intuitively 'fair' [181]. However, it is not necessarily a stabilizing payoff allocation for the coalitional energy management scheme, which means that under the Shapley value payoff allocation, there could be coalitions of prosumers that would achieve a higher profit by leaving the grand coalition and to operate independently. This is shown in a counterexample in Section 3.5.

The nucleolus, on the other hand, is guaranteed to be a stabilizing imputation as proven is Lemma 3.5. To solve for the nucleolus $\boldsymbol{\nu}$, the following iterative process is implemented:

1. Solve linear problem:

$$LP_1 : \varepsilon_1 = \min_{\mathbf{x}, \varepsilon} \varepsilon \tag{3.48}$$

$$s.t. \quad \sum_{\forall i \in \mathcal{N}} x_i = v(\mathcal{N}) \tag{3.49}$$

$$v(\mathcal{S}) - \sum_{\forall i \in \mathcal{S}} x_i \leq \varepsilon, \forall \mathcal{S} \notin \{\emptyset, \mathcal{N}\} \tag{3.50}$$

where Constraint (3.49) ensures the *efficiency* criterion of the imputation, Constraint (3.50) sets ε as the largest excess for all coalitions, and the objective function (3.48) minimizes the largest excess by searching through possible imputations. Optimal solution ε_1 is obtained, along with $\mathfrak{S}_1$, a set of coalitions that are common to all optimal solutions at $\varepsilon = \varepsilon_1$, or coalitions that correspond to constraints defined by (3.50) that are binding. ε_1 is the largest excess of the nucleolus although the nucleolus has not been completely computed by this time. Therefore, if $\varepsilon_1 \leq 0$, the nucleolus is in the core.

2. Iteratively solve the following linear problem until a unique optimal solution $\mathbf{x}^*$ is obtained. $\mathbf{x}^*$ then is the nucleolus of the coalitional energy management scheme: $\boldsymbol{\nu} \leftarrow \mathbf{x}^*$.

$$
\begin{aligned}
LP_a : \varepsilon_a = \min_{\mathbf{x},\varepsilon}\ & \varepsilon \\
s.t.\quad & \sum_{\forall i \in \mathcal{N}} x_i = v(\mathcal{N}) && && (3.51)\\
& \sum_{\forall i \in \mathcal{S}} x_i = v(\mathcal{S}) - \varepsilon_{a'}, && \forall \mathcal{S} \in \mathfrak{S}_{a'}, \\
& && \forall a' \in [1, a-1] && (3.52)\\
& v(\mathcal{S}) - \sum_{\forall i \in \mathcal{S}} x_i \leq \varepsilon, && \forall \mathcal{S} \notin \{\emptyset, \mathfrak{S}_{a'}, \mathcal{N}\}, \\
& && \forall a' \in [1, a-1] && (3.53)
\end{aligned}
$$

where $a \geq 2$, numbering each linear program solved during the iteration process. Constraint (3.51) ensures the *efficiency* criterion of the imputation, Constraint (3.52) fixes the excesses of the coalitions that are binding in all previous iterations, Constraint (3.53) sets ε as the largest excess for all remaining coalitions in LP_a. $\mathfrak{S}_{a'} := \{\mathcal{S} \mid v(\mathcal{S}) - \sum_{\forall i \in \mathcal{S}} x_i = \varepsilon_{a'}, \forall \mathbf{x}_{a'}\}$, where $\mathbf{x}_{a'}$ denotes an optimal payoff allocation obtained in $LP_{a'}$, not unique until the solution converges to the nucleolus $\boldsymbol{\nu}$. $\mathfrak{S}_{a'}$ represents the set of coalitions whose excesses are binding based on (3.53) in $LP_{a'}$.

3.5 Case Studies

In this section, the coalitional energy management scheme is tested using domestic load data measured in the Customer-Led Network Revolution trials, a UK smart grid demonstration project [182]. A 24-hour time frame is selected starting from the midnight of a sunny summer day in July. The PV generation data is simulated in PVWatts [183] using a 4kW, fixed 20 degree tilt residential system, and the London Gatwick solar data. The ES model is a battery with an energy capacity of 7 kWh, a maximum charge power of 3.5 kW, a maximum discharge power of 3.2 kW, both charging and discharging efficiencies of 95%, an initial state of charge of 50%, and a state of charge range of 20-95%. The energy buy price follows a UK Economy 7 residential rate structure [29], and the energy sell price is the UK feed-in tariff [28]. Detailed model inputs can be found in Appendix B.

The cases focus on three aspects. First, it is demonstrated that the prosumer coalition does help reduce the variability of the load profile locally. Second, the Shapley value and the nucleolus are compared, showing that both imputations are able to reward all prosumers. It also demonstrates that the Shapley value is not guaranteed to be stabilizing, while the nucleolus is, providing empirical evidence for the scheme's balancedness proved in Section 3.3. Lastly, the prediction errors are taken into account to evaluate the impact of uncertainty on the value of coalitions.

3.5.1 Load Balancing and Financial Reward

For the first case study, assume there are 10 prosumers, and that there are 5 PV systems and 5 ES systems, which are randomly assigned to the prosumers. Figure 3.1 compares three different coalitional load profiles: 1) without energy management, 2) with each ES system operating to minimize its direct owner's energy cost, and 3) with all ES systems operating cooperatively to minimize the coalitional energy cost. It can be easily seen in the graph that noncooperative ES operation can reduce the reverse power flow and load variability to an extent, but cooperative ES operation is significantly more effective.

The details of the load profile results are influenced by many inputs. In this specific example using the UK load data and rate structure, the morning peak just before 7am is a result of both the high morning usage when people wake up and the last attempt

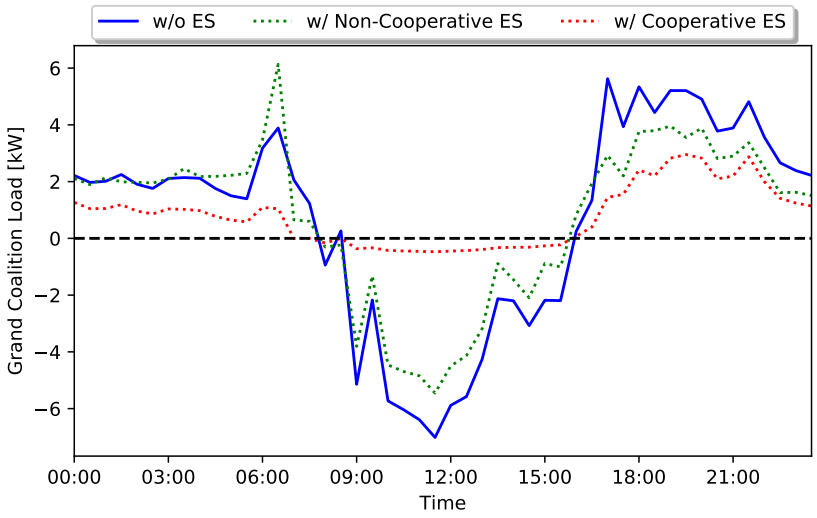


Figure 3.1: 10-prosumer grand coalition 24-hour load profile. The blue line shows the joint energy profile (consumption + PV generation) without considering the operation of ES. The dotted green line shows the joint energy profile with noncooperative ES operation where every prosumer schedules their ES systems to minimize their own energy costs. The dotted red line shows the joint energy profile with cooperative ES operation where all the prosumers collectively schedule their ES systems to minimize the energy cost of the grand coalition.

to store energy in the ES systems before the rate increases. Because the Economy 7 rate structure only has two rate tiers, and the rate remains constant within each tier, the optimal ES operation as a result of the proposed linear model is not unique. However, this non-uniqueness does not affect prosumers' payoffs as the value of coalitions is defined by the minimum energy cost achievable, not the resulting energy profile. As explained in Section 1.1.5, the focus of this thesis is the benefit of prosumer cooperation to the prosumers themselves, not to the network operation. In order to quantify the benefit of load flattening on the network operation, more detailed inputs such as network losses need to be taken into account. This is a natural extension of the current framework.

Figure 3.2 displays the monetary benefits using the nucleolus and Shapley value for each individual prosumer within the grand coalition benchmarked against their overall energy costs. It can be seen that every prosumer profits from the coalitional management scheme, but there are clear differences between the nucleolus and the Shapley value for each participating prosumer. In all the model results obtained, the nucleolus always leads to individual payoff allocations that span over a wider range of magnitude. Specifically in the example given, the smallest individual payoff (Player 10) under the nucleolus

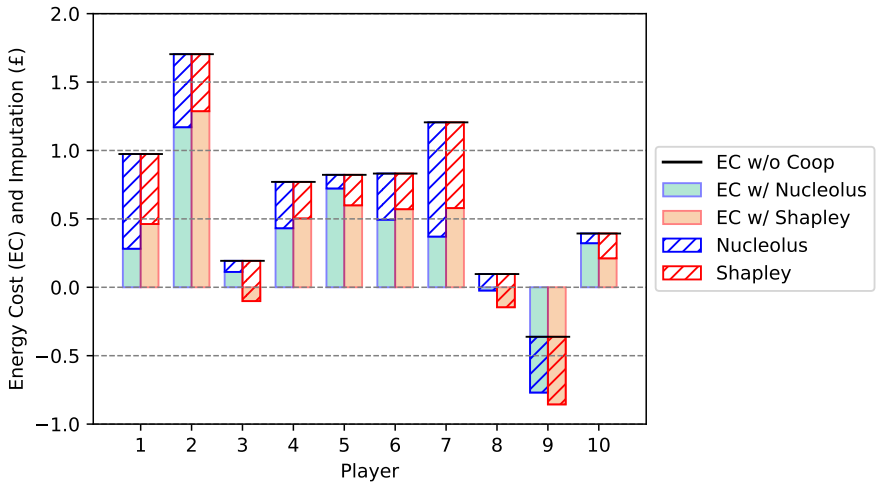


Figure 3.2: 10-prosumer coalitional energy costs and cost savings over 24 hours by end user. The black horizontal lines represent the individual energy costs if prosumers were to schedules their ES systems to minimize their own energy costs. the hashed bars represent the individual payoff allocations based on the Shapley value (red) or the nucleolus (blue) if prosumers were to participate in the coalitional energy management scheme. The solid bars represent the resulting individual energy costs under the coalitional energy management scheme.

($£0.07$) is smaller than that of the Shapley value ($£0.18$), and the largest individual payoff (Player 7) under the nucleolus ($£0.84$) is larger than that of the Shapley value ($£0.63$). Despite these differences, the distribution of their values among all the prosumers follow a similar pattern. This is because both of the imputations reward prosumers on a merit basis – whoever contributes more to reducing the coalitional energy cost tends to receive more in return.

Figure 3.3 shows the comparison of the nucleolus and Shapley value for each prosumer based on their DER ownership. It is noticed that whether a prosumer receives more under the nucleolus or the Shapley value is dependent on their resource mix. For example, prosumers without any PV or ES and prosumers with only ES tend to be rewarded more by the nucleolus, while prosumers with both PV and ES and prosumers with only PV tend to be rewarded more by the Shapley value. This is due to the high penetration of PV, where the value of cooperation relies heavily on the absorption of the PV generation during the day either through pure consumption or ES arbitrage. While the Shapley value reflects the mean of each prosumer's marginal contributions to all the coalitions, the nucleolus has to allocate more to the prosumers whose contributions are higher at

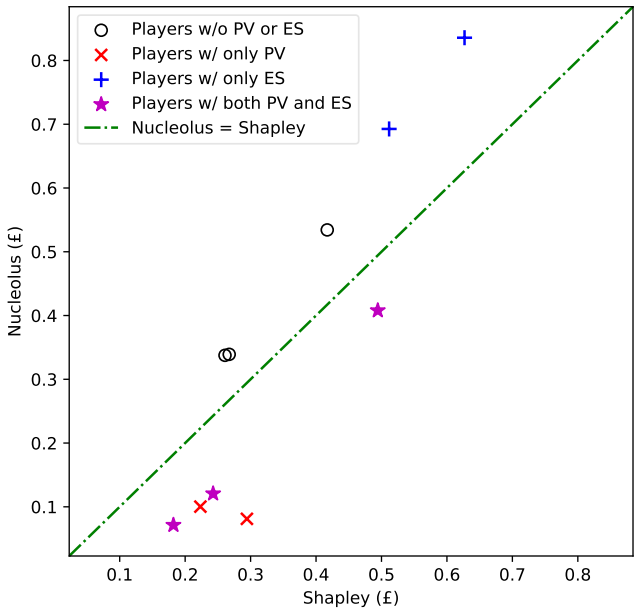


Figure 3.3: Comparison of the individual payoff allocations (nucleolus vs. Shapley value) for a 10-prosumer coalitional energy management scheme. Prosumers are differentiated by the ownership of their DER. Data points below the diagonal line indicates that the prosumer receives more payoff under the Shapley value than the nucleolus, and data points above the diagonal line indicates that the prosumer receives more payoff under the nucleolus than the Shapley value.

the expense of those whose contributions are lower. This is to ensure that out of all the coalitions that the high contributing prosumers are part of, even the one most likely to defect is satisfied with cooperating in the grand coalition.

3.5.2 Nucleolus vs. Shapley Value

For the second case study, the model is run with different numbers ($N = 4, 6, 8, 10, 12, 14$) of prosumers. To test a range of load profiles with different PV and ES assignments, the model 30 is run times for each number of prosumers, and in each run, 50% ownerships of PV and ES systems are assumed, both randomly assigned independently. The average largest excess against the number of prosumers is demonstrated in Figure 3.4, with the error bars showing the maximum and minimum values across all runs. It can been seen that regardless of the number of prosumers in the grand coalition, the largest excess with respect to the Shapley value always exceeds zero, verifying the statement in Section 3.4.2 that the Shapley value is not in the core of this coalitional energy management scheme.

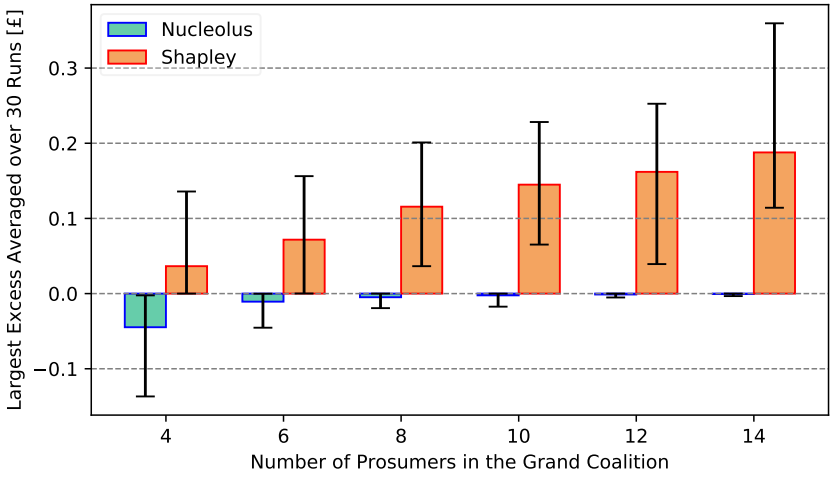


Figure 3.4: Largest excesses by number of prosumers – nucleolus vs. Shapley. The solid bars show the average value over 30 runs with different model inputs, and the error bars show the full range of values in all 30 runs.

On the other hand, the largest excess with respect to the nucleolus is always non-positive, which empirically validates that the nucleolus is a stabilizing imputation.

Taking the same 10-prosumer scheme from Section 3.5.1 as an example, Figure 3.5 maps the excesses of all the coalitions on a nucleolus vs. Shapley coordinate. The data points generally distribute along the diagonal line representing “Nucleolus Excess = Shapley Excess", but some Shapley excesses exceed zero, while no nucleolus excesses do. It is worth noting that even for the coalitions with positive Shapley excesses, their excess values split among the prosumers in those coalitions are relatively small compared to their individual Shapley values.

However, whether the Shapley value or the nucleolus is chosen as the payoff allocation, the scalability of this scheme is limited. This is mainly because the number of linear problems required to compute the values of coalitions is proportional to the total number of coalitions, which increases exponentially with the number of prosumers. For the nucleolus computation, the process of solving for the lexicographically minimal excesses of all the possible coalitions is another intractable step. For instance, it takes an average of 1, 15, and 800 minutes to compute the nucleolus for schemes of 10, 12, and 14 prosumers respectively. Developing methods to improve the scalability of the model is naturally an extension of this work (see Chapter 4, 5).

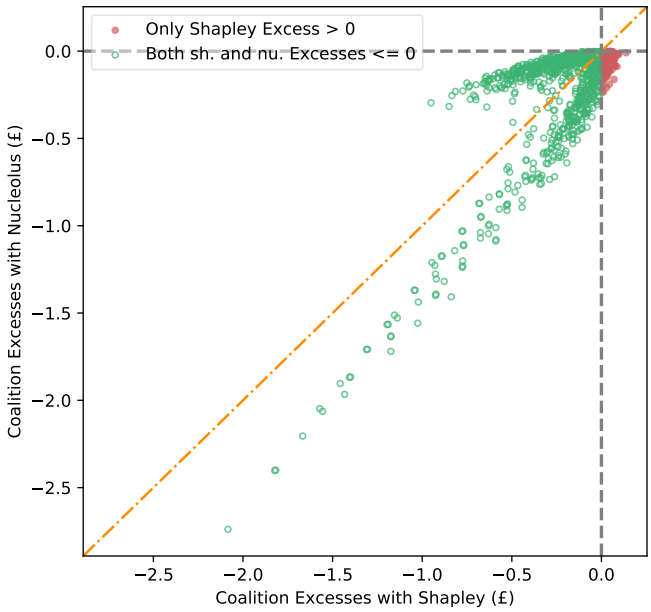


Figure 3.5: Comparison of 24-hour nucleolus excesses and Shapley excesses for all coalitions in a 10-prosumer coalitional energy management scheme. Coalitions with positive Shapley excesses are marked with solid circles, and none of the coalitions have positive nucleolus excesses.

3.5.3 The Impact of Uncertainty

In the implementation of the proposed scheme, prosumers' ES units are scheduled based on forecasts of the consumption and the generation, naturally leading to real-time discrepancies. One problem that arises from these discrepancies is the potential reduction of energy cost savings that can be awarded to the participants. Since the objective of the proposed scheme is to generate profits for the participants through cooperation, it is important to analyze the significance of the potential profit reduction due to uncertainty. The two main objectives of this analysis are: 1) to demonstrate that the costs associated with uncertainty in the traditional setup without cooperation can offset, to an extent, that of the cooperative setup; 2) to quantify the impact of uncertainty on the prosumer payoffs by calculating the percentage reduction of the total profit.

For the third case study, a Monte Carlo analysis is conducted on the impact of prediction errors on the coalitional energy cost savings. The model setup is similar to the previous case studies, except that all input load and PV profiles are treated as the prediction, and the ES operation is optimized based on predicted profiles. For each

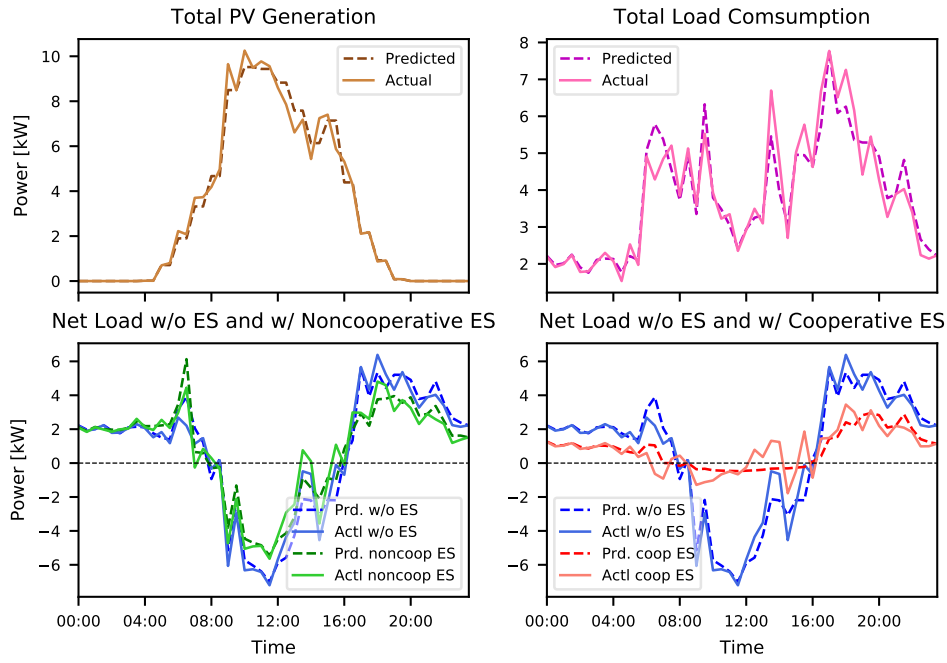


Figure 3.6: Comparison of predicted energy profiles and the actual energy profiles with ES operation ($N = 10$, $\sigma_{PV,load} = 10, 30\%$). Due to the geographic proximity assumption, a common PV profile generated with stochastic errors is applied to all prosumers. All load profiles are generated with stochastic errors independent of each other.

prosumer, the actual load and PV profiles are generated assuming the value at each step follows a normal distribution with a mean equal to the predicted value and a specified standard deviation σ. Because the prediction horizon is only one day, it is reasonable to assume low prediction errors [39], especially for the PV generation. Assuming the prediction standard deviations for PV and load are 10% and 30% for each prosumer, and using the same 10-prosumer scheme from Section 3.5.1 as an example, Figure 3.6 shows the comparison of the predicted profiles and the actual profiles.

In order to maintain the locational similarity among all the prosumers in each model run, only one actual PV profile is generated and used for all PV systems, whereas the actual load profiles are generated separately for each individual prosumer. For a range of numbers of prosumers ($N = 6, 8, 10, 12$), the model is run 200 times with different prosumer profiles for each scheme size. In each model run, the actual coalitional energy cost savings is calculated by applying the ES operations optimized using the predicted profiles to the actual profiles. Finally, the impact of uncertainty is measured by

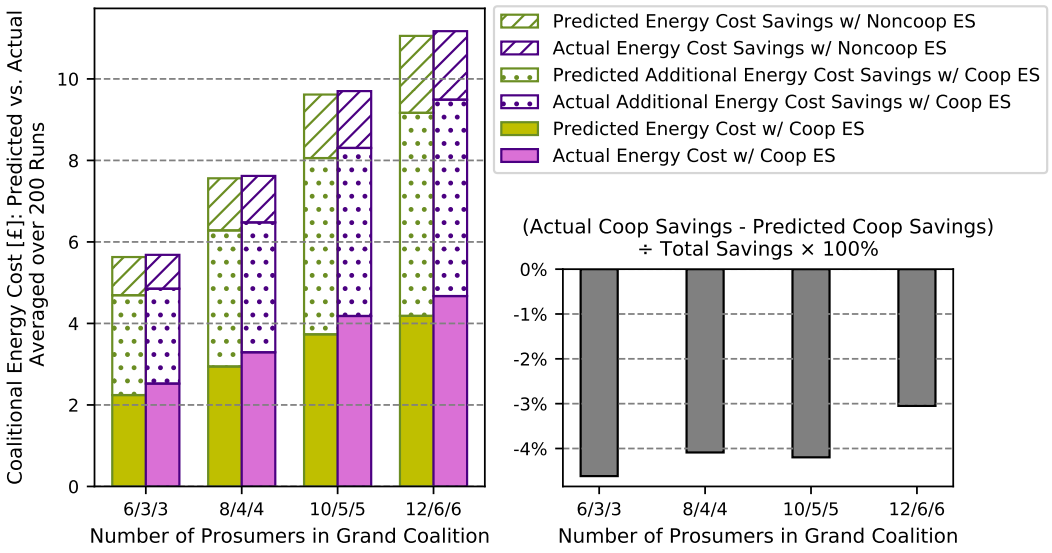


Figure 3.7: Comparison of the predicted energy cost and the actual energy cost for the grand coalition ($\sigma_{PV,load} = 10, 30\%$). The labels on the x-axes are in the format of Number of Prosumers / Number of PV / Number of ES. The dotted bars are the actual and the predicted total energy cost savings through the coalitional energy management scheme.

comparing the change in coalitional energy cost savings to the predicted savings. Figure 3.7 compares the predicted and the actual cases of coalitional energy cost for different numbers of prosumers. The savings are separated into two parts. The first part come from noncooperative ES operation used for self arbitrage, which corresponds to the dotted green line in Figure 3.1. In addition to such savings, the second part of the savings comes from the cooperative ES operation, which corresponds to the solid red line in Figure 3.1. It is shown that prediction errors can lead to reduction in coalitional energy cost savings, but the reduction is small. The change is made less significant because both the individual prosumers' energy costs and the coalitional energy cost are affected by the uncertainties in the same direction, which in turn cancel each other out to an extent when the difference between the two is taken to calculate the cost savings.

As prediction gets more accurate it can be expected that this reduction will become even smaller. To show the impact of uncertainty, two sets of standard prediction errors (σ) are tested: 1) $\sigma_{PV,load} = 5, 15\%$, and 2) $\sigma_{PV,load} = 10, 30\%$. For each set of uncertainty inputs, the average change in coalitional energy cost savings is plotted as a percentage of the predicted savings in Figure 3.8. The results show around 2% of savings reduction

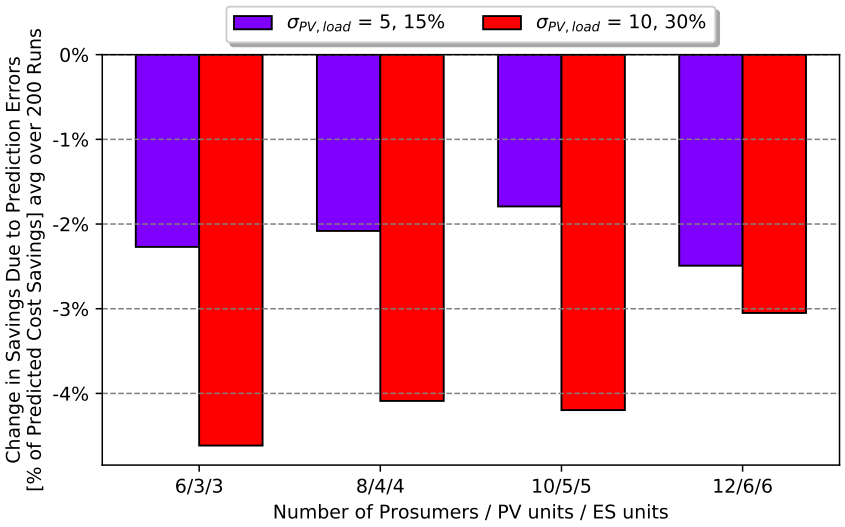


Figure 3.8: Changes in energy cost savings for the grand coalition as percentages of the predicted values due to prediction errors. Two sets of prediction uncertainty inputs are shown. For the high uncertainty case, the standard errors of prediction for the PV generation and prosumer consumption are 10% and 30% respectively, while for the low uncertainty case, the standard errors of prediction are 5% and 15% respectively. The values shown are averaged over 200 model runs.

due to uncertainty for the low uncertainty case, and around 4% for the high uncertainty case. Additionally, the reduction is shown to remain at a similar level as the number of prosumers grows. This indicates that even with uncertainty, the proposed scheme can still offer the prosumers significant amounts of financial benefits.

3.6 Concluding remarks

In this chapter, a cooperative game theoretic approach is used to construct a coalitional energy management scheme. Under this scheme, a prosumer grand coalition is formed, within which all the ES systems are scheduled collaboratively to minimize the joint energy cost. The values of all the coalitions within the grand coalition are defined by the coalitional energy cost savings, and the grand coalition energy cost savings are allocated to the prosumers to ensure that every participant receives a financial profit. Two payoff allocations within the cooperative game framework, the Shapley value and the nucleolus, are selected and discussed, and their detailed computation algorithms are presented.

Reconstructing the game in a matrix format, this coalitional energy management scheme is mathematically proved to be a balanced game, or a game with a non-empty

core. It means that there exist stabilizing payoff allocations, which ensure that all prosumers are rewarded in a way such that no coalitions of prosumers can become financially better off by defecting from the grand coalition and operating independently. It therefore proves, according to existing theorems, that the nucleolus of this balanced game is a stabilizing payoff allocation.

The case study results empirically verify that this coalitional structure helps mitigate the stress on the local power network although more in-depth analyses are required to investigate whether the formation of coalitions is always beneficial for the network operation. Both the Shapley value and the nucleolus are shown to offer economic benefits to prosumers that participate in the scheme. Through a counter example, it shows that the Shapley value is not necessarily a stabilizing payoff allocation, whereas the nucleolus is a stabilizing payoff allocation, offering empirical evidence to support the proof of the balancedness of the scheme. Additionally, the impact of stochastic errors in predicting the PV generation and load consumption is evaluated. Under the assumption that the standard prediction errors for PV generation and load consumption are 10% and 30% respectively, a reduction of about 4% of the grand coalition energy cost savings can be observed.

This chapter addresses part of the thesis research question by proposing a coalitional energy management scheme that incentivizes prosumers to manage their DER in an economically efficient manner. However, a key limitation of this scheme is the computational complexity. In fact, for both the Shapley value and the nucleolus, the computation time increases exponentially with the number of prosumers in the scheme. The scalability of the proposed scheme is the main topic discussed in Chapter 4 and Chapter 5, which develop novel sampling and clustering algorithms for addressing these computational issues.

These men ask for just the same thing, fairness, and fairness only. This, so far as in my power, they, and all others, shall have.

— Abraham Lincoln

4

Sampling for Estimating the Shapley Value - a Fair Payoff Allocation

Contents

4.1 Shapley Value of the Coalitional Energy Management Scheme . . 71
4.1.1 Shapley Value Computation and Properties 72
4.1.2 Shapley Value Computational Complexity 74
4.2 Shapley Value Estimation with Coalition Sampling 75
4.2.1 Stratified Random Sampling 76
4.2.2 Coalitional Stratified Sampling with Optimal Sample Allocation 77
4.2.3 Multi-Step Sampling 80
4.3 Case Studies 82
4.3.1 Validation of Sampling-Based Shapley Estimation 82
4.3.2 Single Prosumer Payoff Analysis 85
4.3.3 Sampling-Based Shapley Estimation for Large Schemes . . . 88
4.4 Concluding remarks 90

Chapter 3 introduced the cooperative game theory based framework for coalitional energy management. One of the main benefits of cooperative game theory for this application is its requirement of a clear strategy to allocate the profit for the prosumer participants. These strategies are called imputations, or payoff allocations that satisfy the *effiency* and the *individual rationality* criteria (see Section 3.3). In the context of a coalitional energy management scheme, an imputation guarantees that the entirety of the grand coalition cost savings is allocated to the participants, and that no prosumer is

worse off financially from participating in this scheme. Among the different types of imputations, the Shapley value [126] is known for its fairness and its analytical expression based on the coalition values. However, it is computationally intensive for large schemes. This chapter explains the Shapley value's computational complexity, and proposes a novel sampling based method to approximate it. It is organized as follows.

Section 4.1 explains the definition and exact computation of the Shapley value and discusses the cause of its computational complexity. Section 4.2 investigates sampling methods for the Shapley value estimation and proposes a multi-step coalitional stratified sampling strategy to scale up the cooperative scheme while maintaining a good estimate of the Shapley value. Section 4.3 applies the proposed estimation strategy to case study examples and demonstrates its effectiveness in reducing computation time while maintaining the estimation accuracy.

4.1 Shapley Value of the Coalitional Energy Management Scheme

The coalitional energy management scheme was introduced in Section 3.2, and it can be summarized as follows:

For an N-prosumer *grand coalition* $\mathcal{N}$, the total energy cost of a coalition $\mathcal{S} \subseteq \mathcal{N}$ at timestep t can be written as

$$F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}}) = \sum_{i \in \mathcal{S}} \left\{ r_t^{im}[q_{it} + b_{it}]_+ + r_t^{ex}[q_{it} + b_{it}]_- \right\} \tag{4.1}$$

where $[y]_{+(-)} = \max(\min)\{y, 0\}$. ***Subscripts*** i and t are indices for the prosumer and the timestep respectively: $i \in \mathcal{N} := \{1, 2, ..., N\}$, $t = 1, 2, ..., R$. The known inputs are r_t^{im}, r_t^{ex}, and q_{it}, which are electricity import price (£/kWh), electricity export price (£/kWh), and net energy consumption (positive) or generation (negative) (kWh) without ES. The variables are $\mathbf{b}^{\mathcal{S}} = [b_{it}], \forall i \in \mathcal{S}, \forall t \in [1, R]$: ES charge (positive) or discharge (negative) energy (kWh).

With the assumption $r_t^{im} > r_t^{ex}, \forall t$, all the ES units' operation within coalition $\mathcal{S}$ can be scheduled to minimize the *coalitional energy cost* $G^R(\mathcal{S})$:

$$
\begin{aligned}
G^R(\mathcal{S}) = \min_{\mathbf{b}^{\mathcal{S}}} & \sum_{t=1}^{R} F_t^{\mathcal{S}}(\mathbf{b}^{\mathcal{S}}) \\
s.t. \quad & \underline{b}_i \leq b_{it} \leq \bar{b}_i, \ \ \forall i \in \mathcal{S}, \forall t \in [1, R] && (4.2) \\
& e_i \underline{SoC}_i \leq e_i SoC_i^0 + \sum_{t=1}^{r} ([b_{it}]_+ \eta_i^{in} + [b_{it}]_- / \eta_i^{out}) \\
& \qquad \leq e_i \overline{SoC}_i, \ \ \forall i \in \mathcal{S}, \forall r \in [1, R] && (4.3) \\
& \sum_{t=1}^{R} ([b_{it}]_+ \eta_i^{in} + [b_{it}]_- / \eta_i^{out}) = 0, \ \ \forall i \in \mathcal{S} && (4.4)
\end{aligned}
$$

where (4.2), (4.3), and (4.4) represent the ES power constraint, energy constraint, and cycle constraint respectively. It is assumed that each prosumer i's ES system has an *energy capacity* (kWh) of $e_i \geq 0$, a *charging limit* (kWh) of $\bar{b}_i \geq 0$ and a *discharging limit* (kWh) of $\underline{b}_i \leq 0$ over the time span of Δt, a *charging efficiency* of $\eta_i^{in} \in (0, 1)$ and a *discharging efficiency* of $\eta_i^{out} \in (0, 1)$, an upper state of charge limit of $\overline{SoC}_i \in [0, 1]$ and a lower state of charge limit of $\underline{SoC}_i \in [0, 1]$, and an initial state of charge of $SoC_i^0 \in [\underline{SoC}_i, \overline{SoC}_i]$. For a prosumer who does not own an ES system, their energy capacity and charging/discharging limits are all set to zeros.

It is understood that the ES systems degrade with usage. In order to obtain the accurate ES inputs, it needs to undergo regular reference performance tests, which cost time and introduce additional degradation, see [184]. To reduce the cost and time, a data-driven state of health estimation method is introduced in [185], which has the potential to be implemented if the concern for privacy violation could be mitigated.

Fig. 4.1 demonstrates the effect of cooperative ES operation in a 16-prosumer scenario. As shown in (d), the cooperative ES operation tends to flatten the load as it tries to match the consumption and generation within the coalition to minimize the *coalition energy cost*.

4.1.1 Shapley Value Computation and Properties

The value of a coalition $\mathcal{S}$ is defined as the energy cost savings obtained by forming the coalition. This is given by the difference between the sum of the energy costs of all

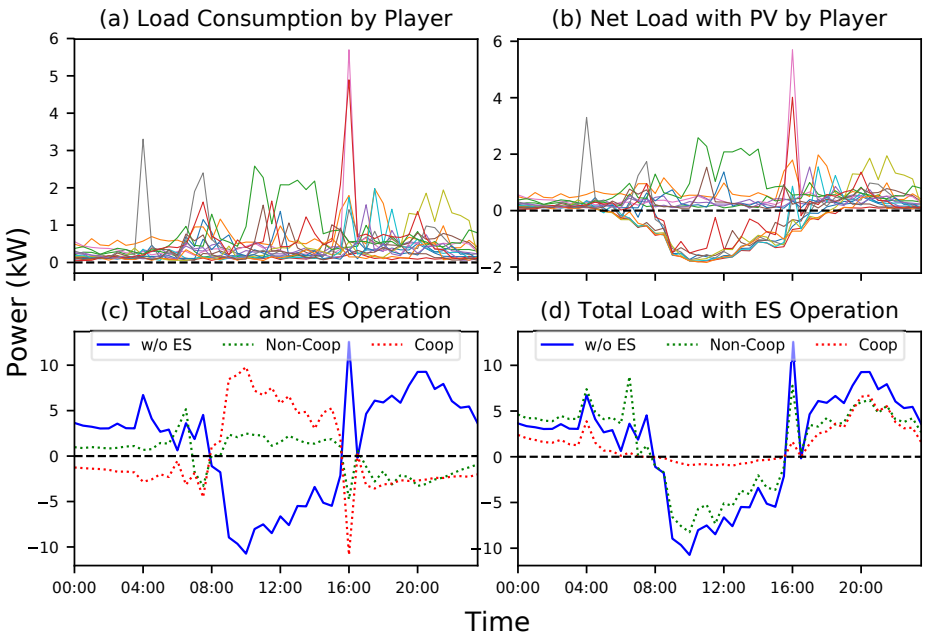


Figure 4.1: 16 Prosumer 24-hour load profiles: (a) individual load consumption only, (b) individual net load without ES (load consumption + PV generation), (c) grand coalition loads and ES operation, where the blue line shows the joint energy profile (consumption + PV generation) without considering the operation of ES, the dotted green line shows the noncooperative ES operation profile where every prosumer schedules their ES systems to minimize their own energy costs, and the dotted red line shows the cooperative ES operation profile where all the prosumers collectively schedule their ES systems to minimize the energy cost of the grand coalition, (d) grand coalition loads, where the blue line shows the joint energy profile (consumption + PV generation) without considering the operation of ES, the dotted green line shows the joint energy profile with noncooperative ES operation, and the dotted red line shows the joint energy profile with cooperative ES operation.

the prosumers in $\mathcal{S}$ when they schedule the ES systems individually, and the minimum coalitional energy cost of $\mathcal{S}$ when the prosumers schedule their ES systems collectively:

$$v(\mathcal{S}) = \sum_{i \in \mathcal{S}} G^R(\{i\}) - G^R(\mathcal{S}) \tag{4.5}$$

By this definition, the value of the *grand coalition* becomes the total energy cost savings of the coalitional energy management scheme, which denotes the total amount of payoffs that can be awarded to all the participants. Vector $\mathbf{x} \in \mathbb{R}^N$ is used as the *payoff allocation* whose entry x_i represents the payment to prosumer $i \in \mathcal{N}$. The Shapley value is denoted as $\boldsymbol{\phi} = \{\phi_i, \forall i \in \mathcal{N}\}$, which can be computed by

$$\phi_i(v) = \sum_{\mathcal{S} \subset \mathcal{N},\, i \notin \mathcal{S}} \frac{(|\mathcal{S}|)!(N - |\mathcal{S}| - 1)!}{N!} [v(\mathcal{S} \cup i) - v(\mathcal{S})] \tag{4.6}$$

The Shapley value also satisfies the following axioms:

1. (*Efficiency*) $\sum_{i\in\mathcal{N}}\phi_i = v(\mathcal{N})$. This requires the entirety of the value created by the grand coalition to be allocated to the players.

2. (*Individual Rationality*) $\phi_i \geq v(\{i\}), \forall i \in \mathcal{N}$. This ensures that no player is penalized for cooperating.

3. (*Symmetry*) If $v(\mathcal{S}\cup\{i\}) = v(\mathcal{S}\cup\{i'\}), \forall \mathcal{S}\subseteq\mathcal{N}, \mathcal{S}\cap\{i,i'\} = \emptyset$, then $\phi_i = \phi_{i'}$. This means that two players should be assigned the same Shapley value if they have the same marginal contributions to all the coalitions.

4. (*Dummy Axiom*) If $v(\mathcal{S}) = v(\mathcal{S}\cup\{i\}), \forall \mathcal{S}\subseteq\mathcal{N}, \mathcal{S}\cap\{i\} = \emptyset$, then $\phi_i = 0$. Therefore, a player's Shapley value should be zero if they add zero marginal value to any of the coalitions.

5. (*Additivity*) If v and u are characteristic functions, then $\phi_i(v+u) = \phi_i(v) + \phi_i(u), \forall i \in \mathcal{N}$. This indicates that the Shapley value of two games played at the same time should be the sum of the two games' Shapley values when played separately.

Axiom 1 guarantees that all the profits allocated to the prosumers add up to the total energy cost savings from the *grand coalition*. In this coalitional energy management scheme, $v(\{i\}) = 0, \forall i \in \mathcal{N}$, so Axiom 2 requires $\phi_i \geq 0, \forall i \in \mathcal{N}$. Axiom 4 ensures that prosumers that do not contribute to any coalition cannot gain any profit. Axiom 3 and 5 indicate the *'fairness'* of the payoff allocation. Specifically, the concept of fairness was first introduced in [186] for games in which deleting a communication relation between two players has the same effect on both their payoffs. It was then proven in [187] that symmetry and additivity guarantee fairness. Axiom 5 is not actively used in this thesis as the coalitional energy management scheme is the only game of concern.

4.1.2 Shapley Value Computational Complexity

Based on the Shapley value definition in (4.6), a player's Shapley value in a cooperative game is the weighted average of their marginal contributions to all the possible coalitions that can be constructed among all players, see [126]. In other words, for an N-player

game, the values of all 2^N possible coalitions need to be computed. This means that the computation of the Shapley value becomes intractable when increasing the number of players. In order to overcome the computational issue, [139] uses a sample-based approximation of the Shapley value of the cooperative scheme, which is constructed as simple games with binary outcomes representing whether a coalition of battery-owning households can overcome a hard network constraint. However, this scheme overlooks the potential contributions made by local generation and small ES systems.

Not limiting to the cooperative game theory application in energy management, techniques to estimate the Shapley value have been explored in some previous literature, sampling being the main methodology. A random sampling method for the Shapley value estimation was proposed in [140], and then modified in [141] by adding a stratification step to improve the accuracy of the estimation.

To improve the scalability of the proposed coalitional energy management scheme with the Shapley value as the payoff allocation, Section 4.2 develops a multi-step coalitional stratified sampling method to estimate the Shapley value.

4.2 Shapley Value Estimation with Coalition Sampling

The computation of the Shapley value is mainly limited by the cost minimization problems that are required to compute the values of the 2^N coalitions, where N is the number of participating prosumers. Since the Shapley value is the weighted average of a prosumer's marginal contributions, sampling is identified as a promising estimation technique to be applied in the coalitional energy management scheme.

Using existing stratified random sampling methods [140, 141] as a basis, this section proposes a novel multi-step sampling strategy to estimate the Shapley value of the coalitional energy management scheme. It not only further reduces the computation time but also provides a metric to measure the estimation accuracy and terminates the sampling process when the desired accuracy is achieved. The proposed method is then shown in case studies in Section 4.3 to outperform existing methods both in terms of the computation time and the estimation accuracy. With a coalitional energy management

scheme sized for a realistic distribution feeder (≥ 50 participants), it is then possible to analyze the impact of different DER adoption rates on prosumer profitability.

4.2.1 Stratified Random Sampling

The conventional definition of the Shapley value is expressed in (4.6). [188] provided an alternative definition of the Shapley value expressed in terms of all possible orders of the prosumers, which was then adopted by [140] to develop a random sampling method to estimate the Shapley value. Adapting this approach, $\pi(\mathcal{N})$ is defined as the set of all possible permutations with prosumer set $\mathcal{N}$, where each permutation is a sequence of all the prosumers in a specific order. For a given permutation $O \in \pi(\mathcal{N})$, $O(l)$ denotes the prosumer in position l, and $Pre^i(O)$ denotes the set of predecessors of Prosumer i, or the coalition of prosumers that are positioned ahead of Prosumer i in the prosumer sequence represented by O. Therefore, if $i = O(l)$, $Pre^i(O) = \{O(1), O(2), ..., O(l-1)\}$. Prosumer i's marginal contribution to the coalition $Pre^i(O)$ is thus a function of O:

$$\delta(O)_i = v(Pre^i(O) \cup i) - v(Pre^i(O)) \tag{4.7}$$

The alternative definition of the Shapley value can be written as

$$\hat{\phi}_i = \sum_{O \in \pi(\mathcal{N})} \frac{1}{N!} \delta(O)_i, i \in \mathcal{N} \tag{4.8}$$

Since $|\pi(\mathcal{N})| = N!$ and $\delta(O)_i$ are equally weighted for all $O \in \pi(\mathcal{N})$ in (4.8), the Shapley value can be estimated using the unweighted expectation of $\delta(O)_i$ given a set of randomly sampled permutations M:

$$\phi_i = \sum_{O \in M} \frac{1}{|M|} \delta(O)_i, i \in \mathcal{N} \tag{4.9}$$

Notice that when $M = \pi(\mathcal{N})$, (4.9) becomes the Shapley value's alternative definition in (4.8).

Using (4.9), the Shapley value can be estimated from randomly sampling prosumer permutations, see [140]. To improve the estimation accuracy, [141] proposed a stratified random sampling approach to divide the population of all prosumer permutations into subpopulations that have the same size of predecessors for each prosumer. This stratified random sampling method follows the following steps.

1. A stratum, or a stratified set of prosumer permutations is defined as $P_{il} := \{O \in \pi(\mathcal{N}) \mid O(l) = i, \forall i, l \in [1, N]\}$. Therefore, P_{il} contains every permutation $O \in \pi(\mathcal{N})$, in which Prosumer i is in position l. Prosumer i's mean marginal contribution of each stratum is

$$\phi_{il} = \frac{1}{|P_{il}|} \sum_{O \in P_{il}} \delta(O)_i, \forall i, l \in [1, N] \tag{4.10}$$

2. A random permutation sample M_{il} of size $|M_{il}|$ is obtained with replacement from each stratum P_{il}.

3. Adapted from (4.10), Prosumer i's mean marginal contribution of the samples from each stratum is

$$\overline{\phi}_{il} = \frac{1}{|M_{il}|} \sum_{O \in M_{il}} \delta(O)_i, \forall i, l \in [1, N] \tag{4.11}$$

The estimated Shapley value can then be calculated as

$$\hat{\phi}_i^{st} = \sum_{l=1}^{N} \frac{1}{N} \overline{\phi}_{il}, \forall i \in [1, N] \tag{4.12}$$

4.2.2 Coalitional Stratified Sampling with Optimal Sample Allocation

It is noticed that in (4.11), $\delta(O)_i$ are equally weighted for all $O \in M_{il}$. Because each prosumer set (i.e. coalition) $Pre^i(O)$ appears $(l-1)!(N-l)!$ times for a given l, they have the same probability of being sampled from P_{il} into M_{il}. The *coalitional stratum* is defined as the set of coalitions $Q_{il} := \{\mathcal{S} \subseteq \mathcal{N} \mid i \notin \mathcal{S}, |\mathcal{S}| = l-1, \forall i, l \in [1, N]\}$, and $\Delta(\mathcal{S})_i = v(\mathcal{S} \cup i) - v(\mathcal{S})$. A random sample H_{il} can then be obtained with replacement from Q_{il}. Because the order of prosumers does not matter in a coalition, H_{il} can be considered a combination sample. (4.11) can be rewritten as

$$\overline{\phi}_{il} = \frac{1}{|H_{il}|} \sum_{\mathcal{S} \in H_{il}} \Delta(\mathcal{S})_i, \forall i, l \in [1, N] \tag{4.13}$$

In order to implement the stratified random sampling method, a procedure to determine the sample size of each stratum needs to be established. [140] identified the true variance as a metric to allocate the samples among strata to minimize the estimation error, and

proposed a two-stage Shapley value estimation algorithm with optimal sample allocation. In the first stage, 50% of the samples are evenly distributed to each stratum to obtain an initial estimated Shapley value and each stratum's sample variance. In the second stage, the remaining 50% of the samples are optimally allocated to each stratum in proportion to their sample variances calculated in the first stage. The final estimated Shapley value is then calculated using the sampling results from both stages.

The true variance of Prosumer i's marginal contributions to stratum P_{il} is defined as

$$\sigma_{il}^2 = \frac{\sum_{O \in P_{il}} (\delta(O)_i - \phi_{il})^2}{(N-1)!} \tag{4.14}$$

As shown in (4.14), to obtain the true variances, each prosumer's accurate marginal contributions in all strata need to be calculated, which has an even higher computational complexity $\mathcal{O}(N!)$ than the full Shapley value calculation $\mathcal{O}(2^N)$. Therefore, in order to determine the optimal sample sizes for each stratum, the true variances also need to be estimated:

$$\overline{\sigma}_{il}^2 = \frac{\sum_{\mathcal{S} \in H_{il}} (\Delta(\mathcal{S})_i - \overline{\phi}_{il})^2}{|H|_{il}} \tag{4.15}$$

It is then recognized that $|Q_{il}| = \frac{(N-1)!}{(l-1)!(N-l)!}$, which means that evenly dividing the samples in the first stage could result in a sample size larger than the size of some *coalitional strata*: $|H_{il}| > |Q_{il}|$, especially when l is close to 1 or N. For these *coalitional strata*, calculating the accurate mean stratum marginal contribution ϕ_{il} requires less samples than its estimate $\overline{\phi}_{il}$ with random sampling:

$$\phi_{il} = \frac{1}{|Q_{il}|} \sum_{\mathcal{S} \in Q_{il}} \Delta(\mathcal{S})_i, \forall i, l \in [1, N] \tag{4.16}$$

With (4.13) and (4.16), the modified two-stage stratified random sampling method with optimal sample allocation is detailed in Algorithm 1.

For the basic coalitional stratified random sampling with optimal sample allocation, the main input is the total sample size h, and the main output is the estimated Shapley value $\hat{\boldsymbol{\phi}}$. Stratum parameters $\boldsymbol{\Phi} = \{\Phi_{il}\}, \boldsymbol{s} = \{s_{il}\}, \forall i, l \in [1, N]$ are updated throughout the process to facilitate the computation of the sample variances $\overline{\sigma}_{il}^2$ and the mean marginal contributions $\overline{\phi}_{il}$ for each coalitional stratum. $\boldsymbol{h} = \{h_{il}\}, \forall i, l \in [1, N]$ denotes

Algorithm 1 Coalitional Stratified Random Sampling with Optimal Sample Allocation (inputs: $h, \boldsymbol{\Phi}', \boldsymbol{s}', \boldsymbol{h}'$)

Stage A
$\boldsymbol{\Phi} \leftarrow \boldsymbol{\Phi}',\ \boldsymbol{s} \leftarrow \boldsymbol{s}'$
$h_{il}^A \leftarrow \frac{h}{2N^2}$
$\Omega \leftarrow \varnothing$: set of strata with sample sizes determined
for $i \in [1, N], l \in [1, N]$ **do**
  if $h_{il}^A > |Q_{il}| = \frac{(N-1)!}{(l-1)!(N-l)!}$ **then**
    $H_{il}^A \leftarrow Q_{il},\ h_{il}^{tot} \leftarrow |Q_{il}|$
    $h \leftarrow h - |Q_{il}|,\ \Omega \leftarrow \Omega \cup (i, l)$
  else
    $H_{il}^A \leftarrow h_{il}^A$ samples with replacement from Q_{il}
  for $\mathcal{S} \in H_{il}^A$ **do**
    $\Delta(\mathcal{S})_i = v(\mathcal{S} \cup i) - v(\mathcal{S})$
    $\Phi_{il} \leftarrow \Phi_{il} + \Delta(\mathcal{S})_i,\ s_{il} \leftarrow s_{il} + (\Delta(\mathcal{S})_i)^2$
  $\overline{\sigma}_{il}^2 \leftarrow \frac{1}{h_{il}^0 + \left|H_{il}^A\right| - 1}(s_{il} - \frac{(\Phi_{il})^2}{h_{il}^0 +} \left|H_{il}^A\right|)$

Stage B
$\omega \leftarrow \{(0, 0)\}$: initialize ω to start following **while** loop
while $\omega \neq \varnothing$ **do**
  for $i \in [1, N], l \in [1, N]$ **and** $(i, l) \notin \Omega$ **do**
    $h_{il}^{tot} \leftarrow h \frac{\overline{\sigma}_{il}^2}{\sum_{i=1}^N \sum_{l=1}^N \overline{\sigma}_{il}^2}$
    $h_{il}^B \leftarrow h_{il}^{tot} - h_{il}^A$
  $\omega \leftarrow \varnothing$: set of over-sampled strata in **Stage 1**
  for $i \in [1, N], l \in [1, N]$ **and** $(i, l) \notin \Omega$ **do**
    if $h_{il}^B < 0$ **then**
      $h_{il}^{tot} \leftarrow h_{il}^A,\ h \leftarrow h - h_{il}^A,\ \omega \leftarrow \omega \cup (i, l)$
  $\Omega \leftarrow \Omega \cup \omega$
for $i \in [1, N], l \in [1, N]$ **and** $(i, l) \notin \Omega$ **do**
  $H_{il}^B \leftarrow h_{il}^B$ samples with replacement from Q_{il}
  for $\mathcal{S} \in H_{il}^B$ **do**
    $\Phi_{il} \leftarrow \Phi_{il} + \Delta(\mathcal{S})_i,\ s_{il} \leftarrow s_{il} + (\Delta(\mathcal{S})_i)^2$
for $i \in [1, N]$ **do**
  for $l \in [1, N]$ **do**
    $h_{il}^{tot} \leftarrow h_{il}^{tot} + h_{il}^0$
    $\overline{\phi}_{il} \leftarrow \frac{\Phi_{il}}{h_{il}^{tot}}$
  $\hat{\phi}_i^{st,opt} \leftarrow \sum_{l=1}^N \frac{1}{N} \overline{\phi}_{il}$
$\boldsymbol{\Phi}'' \leftarrow \{\Phi_{il}\}, \boldsymbol{h}'' \leftarrow \{h_{il}^{tot}\}, \forall i, l \in [1, N],\ \boldsymbol{s}'' \leftarrow \boldsymbol{s}$
$\hat{\boldsymbol{\phi}} \leftarrow \{\hat{\phi}_i^{st,opt}\}, i \in [1, N]$
return $\hat{\boldsymbol{\phi}}, \boldsymbol{\Phi}'', \boldsymbol{s}'', \boldsymbol{h}''$

the total number of samples taken for each coalitional stratum. If the algorithm is implemented on its own without any preceding sampling steps, it is initiated by setting $\mathbf{\Phi}' = \mathbf{0}, \boldsymbol{s}' = \mathbf{0}, \boldsymbol{h}' = \mathbf{0}$. After the algorithm is implemented, the output $\mathbf{\Phi}'', \boldsymbol{s}'', \boldsymbol{h}''$ store the updated values. These three parameters are important for the multi-step sampling explained in Section 4.2.3.

In Stage A, if the evenly distributed sample size is bigger than the stratum size: $h_{il}^A > |Q_{il}|$, the prosumer's marginal contributions in all the possible samples are computed in this coalitional stratum by setting $H_{il}^A \leftarrow Q_{il},\ h_{il}^{tot} \leftarrow |Q_{il}|$. The true mean marginal contribution of Q_{il} can thus be obtained: $\phi_{il} = \overline{\phi}_{il}$. Meanwhile, the saved samples $h_{il}^A - |Q_{il}|$ can be added to Stage B. This way, The number of samples remains the same to maintain a similar computation time, while the accuracy of the estimation is improved with the precise stratum marginal contribution for some strata, and an increased number of samples for the optimal allocation.

4.2.3 Multi-Step Sampling

Regardless of the sampling method, one common input in these estimation models is the total sample size. It is intuitive that the higher the sample size is, the more accurate the estimation results tend to be, but the longer the model is going to take. Considering that the goal of the estimation is to scale up the scheme while maintaining a reasonably accurate estimation of the Shapley value, the two main metrics to evaluate the performance of an estimation method should be the computation time τ and the estimation accuracy, measured by the *standard error of the estimate* as a percentage of the average prosumer payoff:

$$\frac{\sigma(\hat{\mathbf{x}}, \mathbf{x})}{\overline{x}} = \frac{\sqrt{\sum_i^N (\hat{x}_i - x_i)^2 / N}}{v(\mathcal{N})/N} \times 100\% \tag{4.17}$$

where $\hat{\mathbf{x}}$, $\mathbf{x}$, and $\overline{x}$ represent the estimated, the benchmark, and the average payoffs respectively.

For a scheme of a small number of prosumers, the estimation accuracy can be evaluated by comparing the results to those from the full Shapley value calculation. However, a large number of tests with different sample sizes may be needed to identify

the minimum sample size h^* required to achieve a certain level of estimation accuracy $\sigma^*/\overline{x}$. Additionally, for schemes of different sizes, and even for schemes of the same size but of different prosumer compositions, h^* might be different. It thus becomes computationally cumbersome to have to identify h^* for every new scheme. For a scheme of a size too large to compute the true Shapley value, directly measuring the estimation accuracy is simply infeasible.

To address the challenge of sample size selection, a multi-step sampling strategy is proposed in this chapter. In this strategy, h is set to be the maximum sample size, and the sampling process is divided into U steps, with each step of sample size $\rho = h/U$. Functional W is used to represent Algorithm 1: $(\hat{\boldsymbol{\phi}}, \boldsymbol{\Phi}'', \boldsymbol{s}'', \boldsymbol{h}'') = W(h, \boldsymbol{\Phi}', \boldsymbol{s}', \boldsymbol{h}')$. And the implementation of the multi-step sampling strategy is presented in Algorithm 2.

Algorithm 2 Multi-Step Coalitional Stratified Random Sampling with Optimal Sample Allocation (inputs: h, W, σ^*)

$\rho \leftarrow h/U$
$\boldsymbol{\Phi}^0 \leftarrow \mathbf{0},\ \boldsymbol{s}^0 \leftarrow \mathbf{0},\ \boldsymbol{h}^0 \leftarrow \mathbf{0}, \hat{\boldsymbol{\phi}}^0 \leftarrow \mathbf{0}$
for $u \in [1, U]$ **do**
  $(\hat{\boldsymbol{\phi}}^u, \boldsymbol{\Phi}^u, \boldsymbol{s}^u, \boldsymbol{h}^u) \leftarrow W(\rho, \boldsymbol{\Phi}^{u-1}, \boldsymbol{s}^{u-1}, \boldsymbol{h}^{u-1})$
  if $\sigma(\hat{\boldsymbol{\phi}}^u, \hat{\boldsymbol{\phi}}^{u-1}) \leq \sigma^*$ **then**
    end for loop.
$\hat{\boldsymbol{\phi}} \leftarrow \hat{\boldsymbol{\phi}}^u$
$h^* \leftarrow \rho u$
return $\hat{\boldsymbol{\phi}}, h^*$

The main idea of the multi-step strategy is to verify the estimation results as the the sampling process proceeds. The sampling results are stored in memory, and as more samples are added in each sampling step, the Shapley value estimation is updated based on the results in all previous steps. After each step is completed, the updated estimation results are compared against those of the last step, and if the estimation difference between the two steps measured by the *standard error* is below the set threshold: $\sigma(\hat{\boldsymbol{\phi}}^u, \hat{\boldsymbol{\phi}}^{u-1}) \leq \sigma^*$, the algorithm stops as it is considered that the estimation $\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}}^u$ has reached the desired accuracy. At this point, if the total number of samples drawn is less than the set maximum sample size: $h^* < h$, additional computation time is saved than the basic coalitional stratified random sampling. Note that in order to obtain the

actual *standard error of the estimate* of the Shapley value $\sigma(\hat{\phi}^u, \phi)$, the true Shapley value ϕ needs to be computed. It is computationally intractable for larger schemes, but the case study in Section 4.3.1 shows a small scheme of 20 prosumers to validate the effectiveness of this proposed multi-step sampling strategy.

4.3 Case Studies

The following three case studies implement the proposed sampling method to estimate the Shapley value of the coalitional energy management scheme. The first case study compares the estimation accuracies and computation times of the Shapley value estimation methods against each other and against the full Shapley value calculation. The second case study focuses on the potential financial losses for an individual prosumer caused by stochastic error, battery degradation, and the Shapley value estimation error. In the third case study, the cooperative scheme is scaled up to evaluate the payoffs to the prosumers based on their DER types.

Some of the model inputs are as follows: the domestic load data was measured in the Customer-Led Network Revolution trial [182]. the model time frame is 24 hours starting from the midnight of a sunny summer day in July. The PV systems are 4kW with fixed 20 degree tilt, simulated in PVWatts [183] using the London Gatwick solar data. The ES model has an energy capacity of 7 kWh, a maximum charge power of 3.5 kW, a maximum discharge power of 3.2 kW, both charging and discharging efficiencies of 95%, an initial state of charge of 50%, and a state of charge range of 20-95%. The energy import price follows a UK Economy 7 residential rate structure: £0.08/kWh for midnight–7am, and £0.18/kWh for 7am–midnight [29], and the energy export price is the UK feed-in tariff [28] fixed at £0.0379/kWh. For more detailed description of the model inputs, please refer to Appendix B.

4.3.1 Validation of Sampling-Based Shapley Estimation

For a 14-prosumer scheme, the true Shapley value can be easily computed. The PV and ES adoption rates are both fixed at 50%, and both ownerships are randomly assigned independently of each other. In other words, each prosumer can have a PV system,

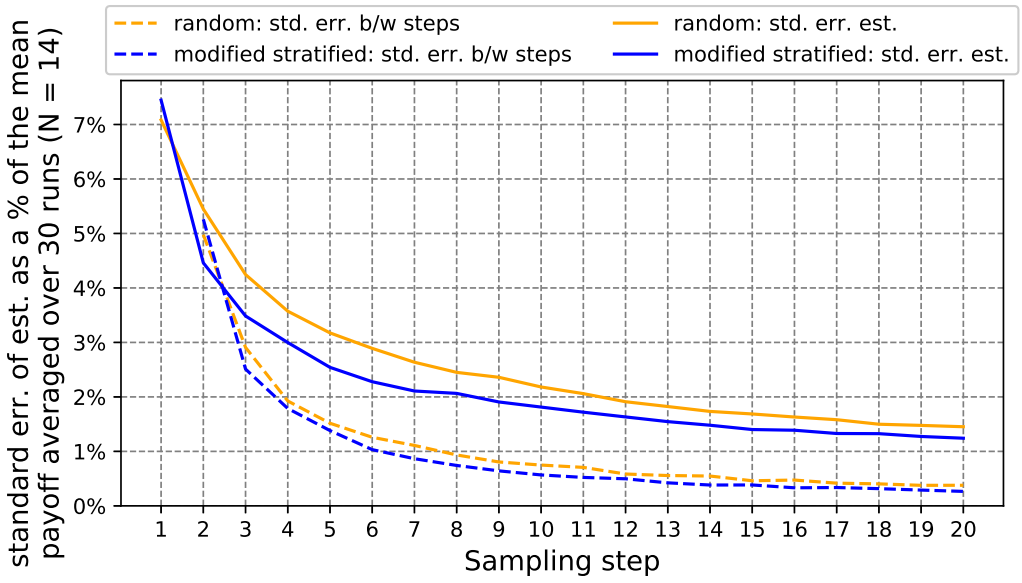


Figure 4.2: Standard error comparison using the multi-step sampling technique: random sampling vs. stratified sampling (N=14). The data for each sampling method at each sampling step are averaged over 30 runs with different inputs of prosumer load and DER ownerships. Each data point on a dotted line represents the standard error between the estimations of the current step and the previous step. Each data point on a solid line represents the standard error between the estimation of the current step and the true Shapley value. The random sampling method and the modified stratified sampling method are represented by yellow and blue lines respectively.

or an ES system, or both, or neither. The proposed multi-step sampling technique is applied to both the random sampling method and the modified stratified sampling, using a maximum sample size $h = 2000N$, and a total number of steps $U = 20$. The Shapley value estimation at the end of each step $\hat{\phi}^u$ is recorded, as well as the estimation difference between every two time steps, which is measured by $\sigma(\hat{\phi}^u, \hat{\phi}^{u-1})$. Then, $\hat{\phi}^u$ is compared with the true Shapley value ϕ to verify its accuracy based on its *standard error of the estimate* $\sigma(\hat{\phi}^u, \phi)$. The model is run 30 times, each time with randomly selected load profiles and PV, ES assignments.

The results are shown in Figure 4.2 with the standard error for each step averaged over 30 runs. The dotted lines represent the standard errors between estimation steps, while the solid lines represent the standard errors between the estimation and the true Shapley value. Even though the dotted lines decline faster than the solid lines meaning that the estimation differences between steps are smaller than the differences between the estimation and the true Shapley value, the estimation error represented by the solid lines still follow a

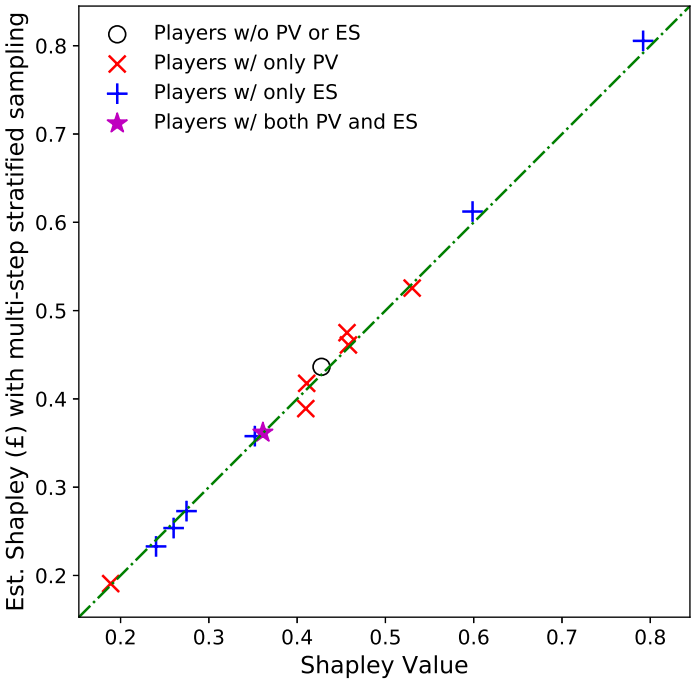


Figure 4.3: Comparison of the true Shapley value with the estimated Shapley values using multi-step stratified sampling (N=14). The closer the data points are to the diagonal line, the more accurate the estimation is.

consistent declining trajectory. Additionally, the stratified sampling method consistently produces more accurate estimates of the Shapley value as more samples are drawn. For example, it takes only 7 steps for the stratified sampling method to produce an estimate that has a 2% standard error from the true Shapley value ($\sigma(\hat{\boldsymbol{\phi}}^u, \boldsymbol{\phi}) \approx 2\%$), whereas the basic random sampling method requires 11 steps, a $\sim 60\%$ increase in computation time, to reach the same accuracy. It is also noticed that $\sigma(\hat{\boldsymbol{\phi}}^u, \boldsymbol{\phi})$ stops showing significant improvement after Step 7, which is around the time when $\sigma(\hat{\boldsymbol{\phi}}^u, \hat{\boldsymbol{\phi}}^{u-1})$ reaches 1%. Therefore, $\sigma^* = 1\% \times \overline{x}$ is chosen as the estimation accuracy criterion for the remaining of the case studies. A comparison of the true Shapley value and the multi-step sampling estimation for each individual prosumer in this 14-prosumer scheme is shown in Figure 4.3. It further demonstrates that regardless of the prosumer type, the estimated values closely resemble the true Shapley value.

Applying Algorithm 2, the cooperative scheme is able to scale up to 100 prosumers within a reasonable computation time. Table 4.1 shows the 10-run average computation

Table 4.1: Shapley Value Computation Time τ (s)

No. Prosumers	10	14	20	50	100	200
full model	2E+1	3E+2	2E+4	N/A*	N/A	N/A
full sampling	2E+1	2E+2	1E+3	4E+3	2E+4	N/A
multi-step smpl	2E+1	1E+2	3E+2	9E+2	3E+3	2E+4

*N/A means the computation takes longer than 10 hours

time[1] of the three models. Any computation time above 10 hours is considered to be impractical for this application. As predicted, the full Shapley value calculation is shown to be intractable for more than 20 prosumers. Sampling is able to scale up the scheme significantly, and the proposed multi-step sampling method can even compute the estimated Shapley value in about 1 hour for a 100-prosumer scheme.

4.3.2 Single Prosumer Payoff Analysis

This case study discusses the financial payoff for an individual prosumer using the 14-prosumer scheme shown in Figure 4.3 as an example. The main objectives of this case study is to put into context how and to what extent different factors, namely ES degradation, uncertainty, and Shapley estimation errors, affect a prosumer's payoffs. This is done by conducting an analysis on the worst case scenario for a prosumer that owns both PV and ES (marked by ★). Using the estimated Shapley value, this prosumer receives a reward of £0.36.

First, the ES degradation cost is considered. It is not directly built into the cooperative game theoretic model mainly due to the short time frame (24 hours). Instead, a state of charge range of 20-95% has been enforced to prevent operating modes that result in significant degradation, but it cannot eliminate it. The average wear cost (AWC) calculation presented in [189] is used to estimate the cost incurred by the cooperation:

$$AWC = \frac{\text{Battery Price}}{ACC(DoD) \times 2 \times DoD \times e \times \eta^{in} \times \eta^{out}} \tag{4.18}$$

where DoD is the depth of discharge, and ACC is the achievable cycle count, a function of DoD. Based on the model inputs, $e = 7kWh$, $\eta^{in} = \eta^{out} = 95\%$ and the state of charge range is 20-95%. Assuming the ES unit in the model is a battery with a

[1]Running on a computer equipped with 16 GB RAM and a 2.8 GHz Quad-Core Intel Core i5 processor.

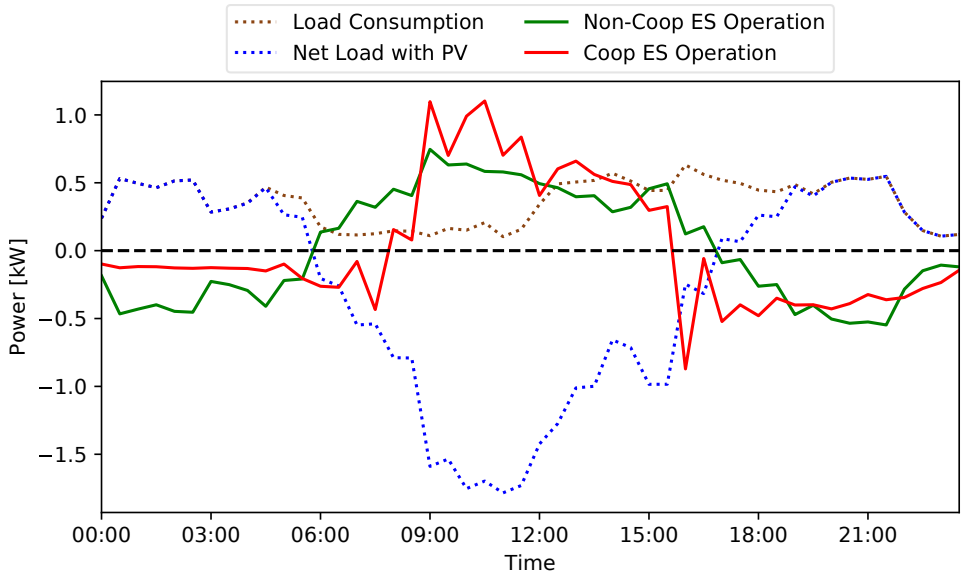


Figure 4.4: Energy profiles of a prosumer with both PV and ES in a 14-prosumer coalitional energy management scheme. The brown dotted line shows the prosumer's load consumption, the blue dotted line shows the prosumer's net energy profile (consumption + PV generation) without considering the operation of ES. The green line shows the noncooperative ES operation where the prosumer schedules their ES system to minimize their own energy cost. The red line shows the cooperative ES operation where the prosumer schedules their ES system to minimize the energy cost of the grand coalition.

unit price of \$156/kWh [190], the battery costs around £850 with a year 2019 average conversion rate of $£1 = \$1.28$[2]. Using a conservative 80% DoD, it can be estimated that $ACC(80\%) = 1000$ based on [191]. Therefore, it can be calculated for this prosumer that $AWC \approx £0.084$/kWh. The prosumer's battery operation profiles with and without cooperation are shown in Figure 4.4, which yields an additional 0.31kWh round-trip operation for the cooperative case. Therefore, the wear cost incurred by the cooperative scheme is $£0.084 \times 0.31 = £0.026$, equivalent to 7.24% of the original reward.

Second, the impact of stochastic errors on the prosumer financial payback is analyzed. A Monte Carlo analysis is conducted by treating all input load and PV profiles as predictions. The actual load and PV profiles are generated assuming the value for each time step follows a normal distribution, for which the mean value is equal to the predicted value and the standard deviation can be specified. The actual coalitional energy cost savings is calculated by applying to the actual profiles the ES operations optimized based on the predicted profiles. Assuming a standard deviation of 5% for solar, and 15%

[2]https://www.macrotrends.net/2549/pound-dollar-exchange-rate-historical-chart

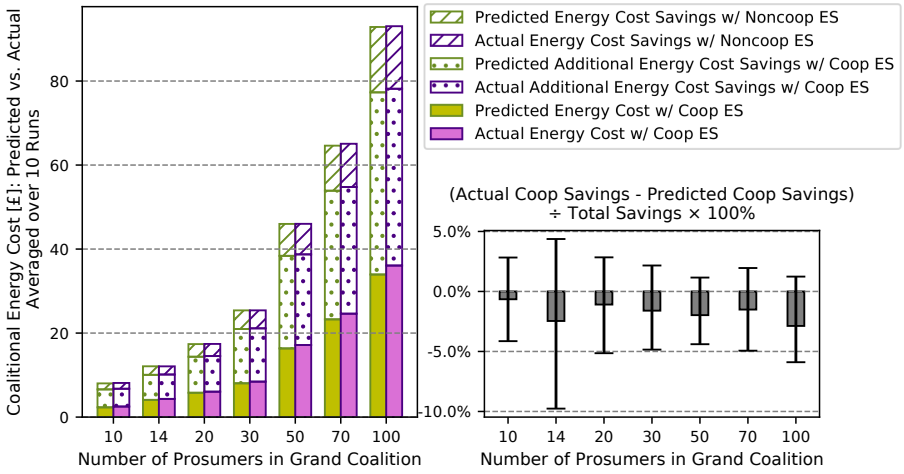


Figure 4.5: The impact of stochastic errors on the grand coalition energy cost savings. The values shown are the average of 10 independent model runs. The figure on the right shows the comparison of the energy costs and energy cost savings for different grand coalitions sizes (predicted case vs. actuality with the ES operation optimized under the prediction); the full length of all three stacked bars is the energy cost without ES operation, the top hashed bars represent the grand coalition energy cost savings if all the ES units are optimized non-cooperatively to minimize the energy costs of each prosumer, the middle dotted bars represent the additional grand coalition energy cost savings achieved through cooperatively scheduling the ES units to minimize the grand coalition energy cost, and the bottom solid bars represent the grand coalition energy costs as a result of cooperation. The middle dotted bars are the total profits of the coalitional energy management scheme, and the differences between the predicted and actual profits are a result of the prediction error, which are plotted on the figure on the right as a percentage of the total predicted savings.

for prosumer loads, the model is run 10 times for schemes of each size, which ranges from 10 to 100 prosumers. The results are shown in Figure 4.5. For the 14-prosumer case, the worst case is a 10% reduction of the total savings, which can be distributed in proportion to the prosumer payoffs. Therefore, the worst case for the prosumer in this case study is considered to be a 10% payoff reduction.

Last, based on Figure 4.2, the Shapley estimation can introduce around 2% error, which contributes to another 2% payoff reduction in the worst case. In total, the worst case can amount to a total of roughly 20% reduction of payoff, which still leaves a profit of $£0.36 \times 80\% = £0.29$ for the prosumer.

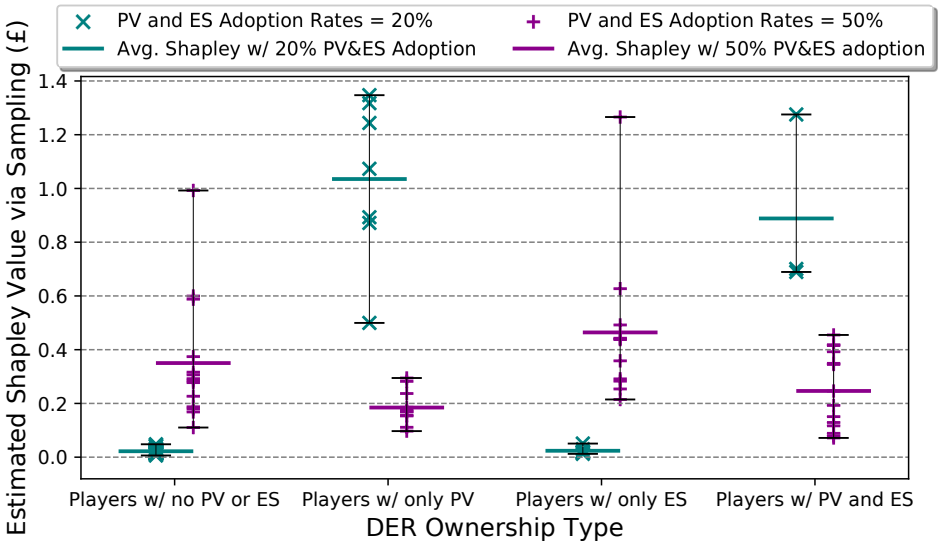


Figure 4.6: Estimated Shapley value for a 50-prosumer scheme by DER ownership type with two different DER adoption rates (20% vs. 50%). PV and ES adoption rates are set to be the same, and are varied together. The 50 prosumers are separated into four categories based on their DER ownership type: 1) no PV or ES, 2) only PV, 3) only ES, and 4) both PV and ES. Their payoff allocations are plotted on the same chart for two different DER adoption rates: 20% and 50%.

4.3.3 Sampling-Based Shapley Estimation for Large Schemes

Having demonstrated the effectiveness of the Sampling methods in estimating the Shapley value, it now becomes possible to investigate the impact of various inputs on the model results for large schemes. For a coalitional energy management scheme with 50 prosumers, first, the adoption rates for PV and ES are set to the same value, and two adoption rates are considered: 20% and 50%. Figure 4.6 compares the estimated Shapley values by prosumers' DER ownership type, where each marker represents a prosumer's estimated Shapley value. There are a few interesting observations. First, except for a few outliers, prosumers with the same DER ownership type are rewarded similar Shapley values regardless of the overall DER adoption rate. Second, as the DER adoption rates change, there is a significant shift in the Shapley values; when the DER adoption rates are low, PV owners are awarded significantly higher Shapley values likely because they provide cheaper energy to the coalitions, while when the DER adoption rates are high, pure consumers and ES owners are awarded higher Shapley values likely because they absorb more local generation. Third, as the DER adoption rates increase, the average Shapley

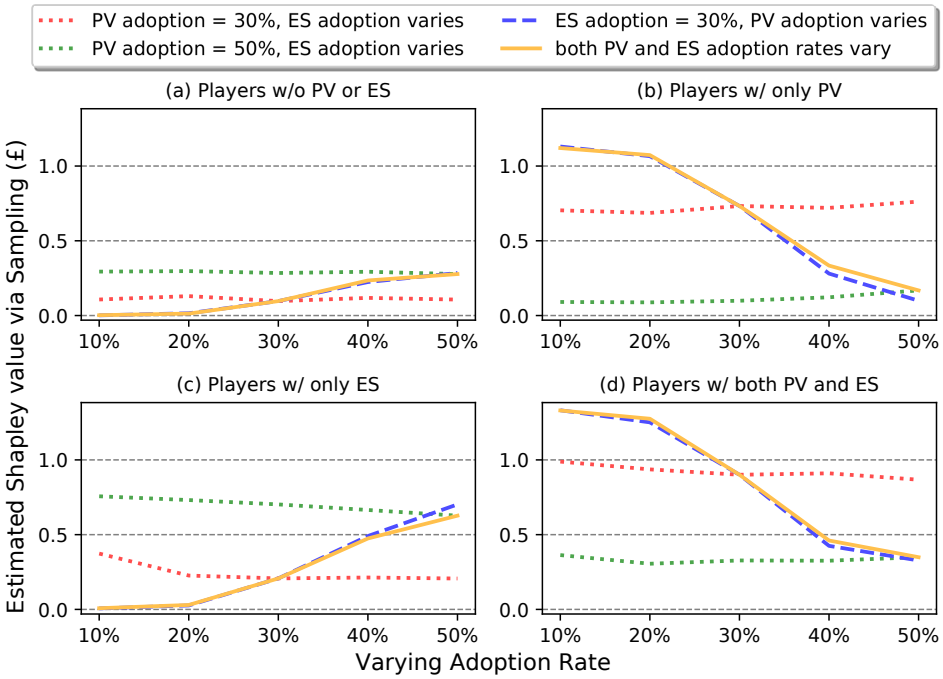


Figure 4.7: Estimated Shapley value for a 50-prosumer scheme with varying PV and ES adoption rates. PV and ES adoption rates are varied under four scenarios: 1) red dotted line shows PV adoption rate fixed at 30%, and ES adoption rate varying between 10%-50%, 2) green dotted line shows PV adoption rate fixed at 50%, and ES adoption rate varying between 10%-50%, 3) blue dashed line shows ES adoption rate fixed at 30%, and PV adoption rate varying between 10%-50%, and 4) yellow solid line shows both PV and ES adoption rates varying between 10%-50%. The average individual payoff under these four scenarios are plotted in four subplots based on the the prosumers' DER ownership type : a) no PV or ES, b) only PV, c) only ES, and d) both PV and ES.

values by DER ownership type tends to converge despite the wider spread among the pure consumers and prosumers with only ES systems.

With the same 50 prosumers, the focus is shifted to four typical prosumers with different DER ownership types, and implement the coalitional energy management scheme under four different scenarios: 1) PV adoption rate is fixed at 30%, and ES adoption rate varies from 10% to 50%, 2) PV adoption rate is fixed at 50%, and ES adoption rate varies from 10% to 50%, 3) ES adoption is rate fixed at 30%, and PV adoption rate varies from 10% to 50%, and 4) PV and ES are with the same adoption rate that varies from 10% to 50%. Figure 4.7 illustrates how the Shapley value changes with different DER adoption rates. Based on the DER ownership type, the trend at which the Shapley value changes with the varying DER adoption rates can be very different. For example, a consumer that does not own any PV or ES tends to be awarded more when

the adoption rates for the PV and ES increase together, whereas a prosumer that owns both PV and ES display the opposite trend. It is interesting to note that when the PV adoption rate is fixed, whether at 30% or 50%, varying the ES adoption rate has very little influence on the Shapley value regardless of the prosumer type. In contrast, whether the ES adoption rate is fixed at 30% or follows the PV adoption rate, varying the PV adoption rate has a significant impact on the Shapley value of all prosumer types.

It is worth noting that the main purpose of the case studies is to validate the scalability of the proposed sampling method applied in the coalitional energy management scheme. The specific results shown are dependent on the assumptions made about the PV and ES system specifications, and the energy prices. Further sensitivity analyses are required to generalize the results to other markets.

4.4 Concluding remarks

To improve the scalability of a cooperative game based energy management scheme, this chapter proposes the novel application of a coalitional stratified random sampling method to estimate the Shapley value. A multi-step sampling approach is created to divide the sampling process to smaller steps. By measuring the standard error of the estimate between steps, it provides a way to verify the estimation results and terminates the sampling process once a certain estimation accuracy threshold is achieved, further reducing the computation time. The maximum size of the scheme that can be computed in a reasonable time (< 10 hours) is thus increased from less than 20 prosumers to over 100 prosumers. For schemes that are small enough to generate the true Shapley value, the effectiveness of the proposed method is demonstrated by showing a standard error of the estimate smaller than 2% of a prosumer's average payoff. The proposed model is then run on a coalitional energy management scheme of 50 prosumers to demonstrate the patterns of the Shapley value for different prosumer DER ownership types and with varying DER adoption rates. It shows that compared to the ES adoption rate, the PV adoption rate has a much larger impact, both negative and positive depending on the DER ownership type, on the prosumer payoffs.

So far, this thesis has explained how to use cooperative game theory to construct a coalitional energy management scheme, and how to allocate the energy cost savings of the scheme to the prosumer participants through the estimated Shapley value. Through the proposed multi-step coalitional stratified sampling, the scheme is able to scale up to over 100 prosumers, a realistic size to apply to a normal distribution network. It has been shown in Chapter 3 that the Shapley value is not necessarily stabilizing, whereas the nucleolus is. That means that the nucleolus can guarantee that no prosumers or coalitions of prosumers can leave the grand coalition to create higher profits for themselves. However, the nucleolus is even more computationally complex than the Shapley value. How to address this computational issue is the main focus of Chapter 5.

If we cannot now end our differences, at least we can help make the world safe for diversity.

— John F. Kennedy

5

Clustering for Estimating the Nucleolus - a Stabilizing Payoff Allocation

Contents

5.1 Nucleolus of the Coalitional Energy Management Scheme 94
5.1.1 Nucleolus: A Stabilizing Imputation of the Coalitional Energy Management Scheme 94
5.1.2 Nucleolus Computational Complexity 98
5.2 Prosumer Clustering Techniques 99
5.2.1 K-Means Clustering 99
5.2.2 Hierarchical Clustering 102
5.2.3 Gaussian Mixture Model 103
5.2.4 Prosumer Clustering Features and Principal Component Analysis 105
5.3 Nucleolus Estimation with Prosumer Clustering 111
5.3.1 Simple K-Means for Computing the Nucleolus for Clustered Players 111
5.3.2 Full Nucleolus Estimation with Clustering 115
5.3.3 Evaluating Nucleolus Estimation with Stratified Random Sampling 117
5.4 Case Studies 122
5.4.1 Validation of Nucleolus Estimation 123
5.4.2 Trade-Off between Estimation Accuracy and Computation Time 125
5.4.3 Comparison of Clustering Techniques 128
5.4.4 Scaling up with Relaxed Computation Time 130
5.5 Concluding remarks 131

The Shapley value, albeit a fair payoff allocation that can be computed analytically,

has been shown in Section 3.5 that it is not guaranteed to be stabilizing. In other words, some coalitions within the grand coalition may find it economically beneficial to defect from the grand coalition and operate on their own. The nucleolus, on the other hand, is a stabilizing imputation for the coalitional energy management scheme as it has been proven to be a balanced game in Section 3.3.

However, as shown in this chapter, the nucleolus requires even more computational power to calculate than the Shapley value. The sampling algorithms developed in Chapter 4 suits the Shapley value because the definition of the Shapley value is the average marginal contribution of a player, which can be well estimated by unbiased random sampling. The nucleolus computation includes a ranking of the coalitions' excesses, thus requiring a method that can capture the dissimilarities among the prosumers and prosumer coalitions.

Clustering, known as an unsupervised learning technique, has been widely used in the energy sector for classifying customer loads. It is identified as a promising method for capturing the similarities and differences among the prosumers, which may facilitate the estimation of the nucleolus.

The main contributions of this chapter include

- a method to estimate the nucleolus that incorporates clustering, which significantly reduces the computation time of coalitional energy management schemes.
- the formulation of a marginal contribution profile instead of the load profile as the clustering features, which significantly improves the nucleolus estimation accuracy.
- a performance evaluation approach inspired by a stratified random sampling technique to compare the estimation performances of different clustering techniques.

This chapter is structured as follows: Section 5.1 summarizes the coalitional energy management scheme and the nucleolus computation algorithm, and explains in detail the computational complexity of the nucleolus. Section 5.2 provides a description of various clustering techniques and proposes suitable clustering features for the coalitional energy management scheme. Section 5.3 proposes a nucleolus estimation method that

incorporates the clustering techniques described in the previous section, and formulates sampling based strategies to evaluate the performance of the proposed method. Finally, Section 5.4 uses case studies to demonstrate the effectiveness of the proposed nucleolus estimation method, and compares different clustering techniques both in terms of the computation time and the estimation accuracy.

5.1 Nucleolus of the Coalitional Energy Management Scheme

This section summarizes the coalitional energy management scheme in a succinct format, and rewrites the computation of the nucleolus in a clear algorithm. Based on the computation algorithm for the nucleolus, the section then analyzes the reasons for its computational complexity of the nucleolus and discusses potential solutions.

5.1.1 Nucleolus: A Stabilizing Imputation of the Coalitional Energy Management Scheme

In an N-prosumer cooperative game, the *grand coalition* $\mathcal{N}$ is defined as the set of all N prosumers. Any subset of the grand coalition $\mathcal{T} : \mathcal{T} \subseteq \mathcal{N}$ is called a *coalition*[1]. Each prosumer is indexed by i and $i \in \mathcal{N} := \{1, 2, ..., N\}$. Given R timesteps ($t = 1, 2, ..., R$) each with duration Δt, the *coalitional energy cost* for $\mathcal{T}$ at timestep t is written as

$$F_t^{\mathcal{T}}(\mathbf{b}^{\mathcal{T}}) = \sum_{i \in \mathcal{T}} \left\{ r_t^{im}[q_{it} + b_{it}]_+ + r_t^{ex}[q_{it} + b_{it}]_- \right\} \tag{5.1}$$

where $[y]_+ = \max(z, 0),\ [y]_- = \min(z, 0)$, and *subscripts* i and t are indices for the prosumers and the timesteps. The electricity import price r_t^{im} (£/kWh) and export price r_t^{ex} (£/kWh) are known parameters, as well as q_{it} (kWh), which denotes the net energy consumption (positive) or generation (negative) over Δt without ES. The variables are ES charge (positive) or discharge (negative) energy amounts (kWh) over Δt, denoted by $\mathbf{b}^{\mathcal{T}} = [b_{it}], \forall i \in \mathcal{T}, \forall t \in [1, R]$.

[1] As a distinction from the previous chapters, $\mathcal{T}$ is used in this chapter as a default coalition symbol instead of $\mathcal{S}$.

With the assumption $r_t^{im} > r_t^{ex}, \forall t$, the operation of ES units within coalition $\mathcal{T}$ can be scheduled to minimize the *coalitional energy cost*:

$$G^R(\mathcal{T}) = \min_{\mathbf{b}^{\mathcal{T}}} \sum_{t=1}^{R} F_t^{\mathcal{T}}(\mathbf{b}^{\mathcal{T}}) \tag{5.2}$$

$$s.t. \quad \underline{b}_i \leq b_{it} \leq \bar{b}_i, \ \ \forall i \in \mathcal{T}, \forall t \in [1, R] \tag{5.3}$$

$$\begin{aligned} e_i \underline{SoC}_i &\leq e_i SoC_i^0 + \sum_{t=1}^{r} ([b_{it}]_+ \eta_i^{in} + [b_{it}]_- / \eta_i^{out}) \\ &\leq e_i \overline{SoC}_i, \ \ \forall i \in \mathcal{T}, \forall r \in [1, R] \end{aligned} \tag{5.4}$$

$$\sum_{t=1}^{R} ([b_{it}]_+ \eta_i^{in} + [b_{it}]_- / \eta_i^{out}) = 0, \ \ \forall i \in \mathcal{T} \tag{5.5}$$

where (5.3), (5.4), and (5.5) represent the ES power constraint, energy constraint, and cycle constraint respectively. It is assumed that each prosumer i's ES system has an *energy capacity* (kWh) of $e_i \geq 0$, a *charging limit* (kWh) of $\bar{b}_i \geq 0$ and a *discharging limit* (kWh) of $\underline{b}_i \leq 0$ over the time span of Δt, a *charging efficiency* of $\eta_i^{in} \in (0, 1)$ and a *discharging efficiency* of $\eta_i^{out} \in (0, 1)$, an upper state of charge limit of $\overline{SoC}_i \in [0, 1]$ and a lower state of charge limit of $\underline{SoC}_i \in [0, 1]$, and an initial state of charge of $SoC_i^0 \in [\underline{SoC}_i, \overline{SoC}_i]$. For a prosumer who does not own an ES system, their energy capacity and charging/discharging limits are all set to zeros.

An example of a 20-prosumer coalitional energy management scheme is demonstrated in Figure 5.1. It shows that in order to minimize the *coalition energy cost*, the scheme tries to match consumption and generation within the cooperative scheme, which results in a flatter cooperative load profile.

In a coalitional energy management scheme, the value function of a coalition is defined as the energy cost savings by forming the coalition:

$$v(\mathcal{T}) = \sum_{i \in \mathcal{T}} G^R(\{i\}) - G^R(\mathcal{T}) \tag{5.6}$$

Based on this definition, the value of the *grand coalition* becomes the total energy cost savings of a coalitional energy management scheme, which determines the total amount of payoffs that can be allocated to the participants.

Vector $\mathbf{x}$ is denoted as the *payoff allocation*, and its entry x_i represents the payment to prosumer $i \in \mathcal{N}$. Based on Definition 3.3, $\mathbf{x}$ is an *imputation* if it is

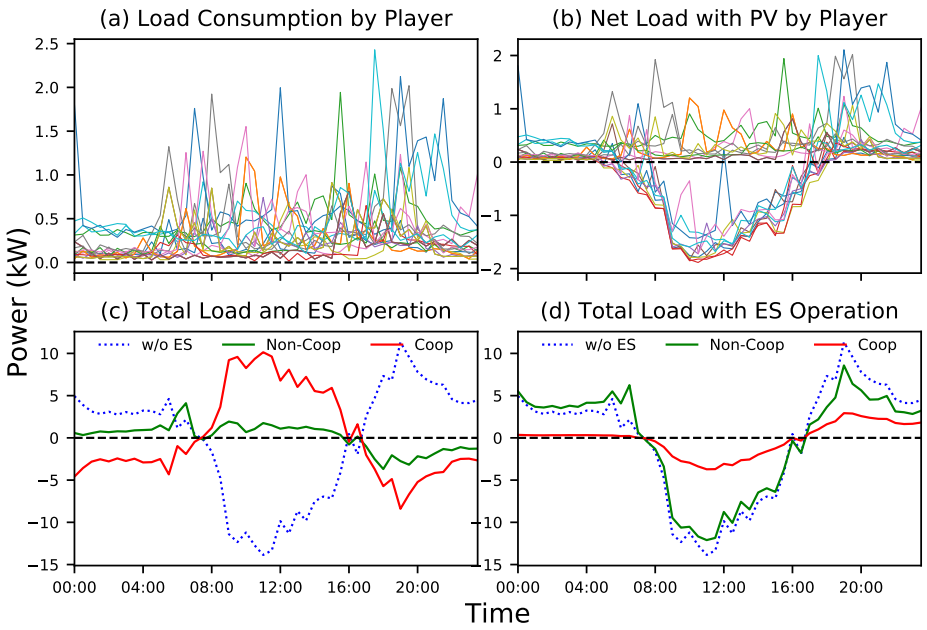


Figure 5.1: 20 Prosumer 24-hour load profiles: (a) individual load consumption, (b) individual consumption + PV generation, (c) joint consumption + PV generation (blue dotted line), joint ES operation if all ES systems are scheduled non-cooperatively to minimize individual energy costs (green solid line), and joint ES operation if all ES systems are scheduled cooperatively to minimize the joint energy cost (red solid line), (d) joint consumption + PV generation (blue dotted line), joint consumption + PV + non-cooperative ES operation (green solid line), and joint consumption + PV + cooperative ES operation (red solid line).

1. *Efficient:* $\sum_{i\in\mathcal{N}} x_i = v(\mathcal{N})$. This requires the entirety of the value created by the grand coalition to be allocated to the prosumers.

2. *Individually Rational:* $x_i \geq v(\{i\}), \forall i \in \mathcal{N}$. This ensures that no prosumer is penalized for cooperating.

However, an imputation does not necessarily guarantee everyone being *satisfied* in the grand coalition, as some prosumers may be able to achieve higher payoffs by forming smaller energy coalitions. A coalition's *dissatisfaction* with respect to the imputation $\mathbf{x}$ is measured by its excess (Definition 3.5):

$$\varepsilon(\mathbf{x}, \mathcal{T}) = v(\mathcal{T}) - \sum_{i\in\mathcal{T}} x_i \tag{5.7}$$

If $\varepsilon(\mathbf{x}, \mathcal{T}) > 0$, then $\mathcal{T}$ is better off on its own and can offer higher payoffs to its members than the grand coalition.

An imputation that satisfies all coalitions is called a stabilizing imputation: $\varepsilon(\mathbf{x}, \mathcal{T}) \leq 0, \forall \mathcal{T} \subseteq \mathcal{N}$.Since the coalitional energy management scheme has been proven to be a balanced game in Section 3.3, the nucleolus is guaranteed to be a stabilizing imputation [175].

The nucleolus $\boldsymbol{\nu}$ is a payoff allocation with lexicographically minimal excesses of all the possible coalitions [61]. It is proven in Lemma 3.5 to be *stabilizing* for the coalitional energy management scheme. Algorithm 3 explains the iterative steps of lexicographically minimizing the excesses of a given set of coalitions. In order to compute $\boldsymbol{\nu}$, the set of all the possible coalitions in an N-prosumer coalitional energy management scheme is denoted as $\mathfrak{T}^{\mathcal{N}} := \{\mathcal{T} | \mathcal{T} \subseteq \mathcal{N}\}$. By setting $\mathfrak{T}^* = \mathfrak{T}^{\mathcal{N}}$, Algorithm 3 outputs $\mathbf{x}^*$, and $\boldsymbol{\nu} = \mathbf{x}^*$.

Algorithm 3 Lexicographical Minimization of Coalition Excesses

$\mathfrak{T}_0 \leftarrow \{\mathcal{N}\}, \quad w(\mathcal{N}) \leftarrow v(\mathcal{N}), \quad a \leftarrow 1$
while $a \leq 2^N$ **do**

$$LP_a : \varepsilon_a = \min_{\mathbf{x}_a, \varepsilon} \varepsilon$$
$$s.t. \sum_{\forall i \in \mathcal{T}} x_{a_i} = w(\mathcal{T}), \forall \mathcal{T} \in \mathfrak{T}_{a-1} \tag{5.8}$$
$$v(\mathcal{T}) - \sum_{\forall i \in \mathcal{T}} x_{a_i} \leq \varepsilon, \forall \mathcal{T} \in \mathfrak{T}^* \setminus \mathfrak{T}_{a-1} \tag{5.9}$$

$\mathbf{T}_a \leftarrow \{\mathcal{T} \mid v(\mathcal{T}) - \sum_{\forall i \in \mathcal{T}} x_{a_i} = \varepsilon_a, \forall \mathbf{x}_a\}$
$w(\mathcal{T}) \leftarrow v(\mathcal{T}) - \varepsilon_a, \forall \mathcal{T} \in \mathbf{T}_a$
$\mathfrak{T}_a \leftarrow \mathfrak{T}_{a-1} \cup \mathbf{T}_a$
if $\mathfrak{T}_a = \mathfrak{T}^*$ **then**
  end while
else
  $a \leftarrow a + 1$
  continue while
$\mathbf{x}^* \leftarrow \mathbf{x}_a$
return $\mathbf{x}^*$

Here, subscript a numbers each linear program (LP) solved during the iteration process. Constraints (5.8) ensure the *efficiency* criterion of the payoff allocation, and fix the excesses of the coalitions that have been minimized in all previous iterations. Constraints (3.53) cap the excesses for all remaining coalitions at ε that is then minimized in LP_a. At the end of each iteration, all the binding constraints among (5.9) are identified,

with their corresponding coalitions defined as $\mathbf{T}_a$. $\mathbf{T}_a$'s excesses are then fixed for all subsequent iterations through (5.8).

Using the nucleolus as the payoff allocation for this cooperative game theoretic scheme offers a way to encourage DER collaboration while ensuring that all prosumers are financially incentivized to remain in the *grand coalition*.

5.1.2 Nucleolus Computational Complexity

Despite the nucleolus being a stabilizing payoff allocation, it faces significant computational issues. As the number of prosumers in this coalitional energy management scheme increases, the nucleolus becomes computationally intractable. The reason for this computational complexity is two-folded. First, the values need to be calculated for all possible coalitions, the number of which is an exponential function of the number of prosumers $|\mathfrak{T}^{\mathcal{N}}| = 2^N - 1$. Second, in solving for the lexicographically minimal excesses of all coalitions, both the number of linear problems and the number of constraints in each linear problem are $\sim \mathcal{O}(2^N)$. Both are closely linked to the number of coalitions, so the key to reducing the computation time is to limit the number of coalitions considered in the computation.

A stabilizing payoff allocation in an analytical expression was proposed in [136], and a worst-case excess method was used in [137] to compute a partial nucleolus: a payoff allocation that is obtained by only implementing the first iteration of the nucleolus computation. However, both of these payoff allocations only address the second computational problem, leaving the coalition value calculations still intractable. Among all the case studies for the coalitional energy management scheme in [136, 137], the highest number of prosumers shown is 15 [137].

The nucleolus payoff each prosumer receives is determined by their contribution to the coalitions, which means prosumers of similar energy behaviors tend to receive similar payoffs. Meanwhile, the value of cooperation relies on the matching of generation and consumption within the coalition, which means it is the heterogeneity within a coalition that creates value for its members. By capturing the similarities and differences among prosumers, a reduced number of coalitions that are most likely to be dissatisfied can

be identified. It then becomes possible to estimate the nucleolus by only minimizing the excesses of these coalitions. For that purpose, Section 5.2 investigates the use of different clustering techniques to identify similarities and difference, and Section 5.3 adopts these clustering techniques for estimating the nucleolus.

5.2 Prosumer Clustering Techniques

In the energy sector, clustering is often used to identify typical load patterns for energy pricing structure design [154] and electric load forecasting and management [156]. Considering the payoffs derived from the nucleolus of the coalitional energy management scheme heavily depend on the participants' consumption and generation profile patterns, clustering is therefore a promising technique to help estimate the nucleolus and overcome its computational hurdle.

K-means clustering is widely used as a reliable baseline clustering technique [145, 158], but hierarchical clustering and Gaussian mixture model (GMM) are shown to have superior performances measured by the Clustering Validity Indexes (CVIs) according to [159], which provides a comprehensive review of various load profile clustering techniques. This section provides a detailed description of how to implement these three clustering methods in the context of the coalitional energy management scheme.

For a set of prosumers grouped into K clusters, let these clusters be denoted by $cl_{\mathcal{K}} := \{cl_1, cl_2, ..., cl_K\}$, where j is the cluster index and $j \in \mathcal{K} := [1, 2, ..., K]$. The clustering assignment is denoted by $\boldsymbol{g} : g_i = j \mid i \in cl_j, \forall i \in \mathcal{N}$. Here, K is applicable for not only K-means, but all clustering techniques.

5.2.1 K-Means Clustering

K-means clustering is a widely used data clustering technique for unsupervised learning tasks. It requires a predetermined number of clusters K, and an initial set of K centroids usually chosen at random among the given instances to be clustered. Clustering is then done through an iterative process of minimizing the sum of all the squared errors between all instances and their corresponding centroids. To cluster prosumer feature profiles,

the error is measured by the Euclidean distance. The Euclidean distance between any prosumer i and any cluster j's centroid is defined as

$$d(\mathbf{f}_i, \mathbf{c}_j) = \sqrt{\sum_{t}^{R} (f_{it} - c_{jt})^2} \tag{5.10}$$

where $\mathbf{f}_i$ denotes prosumer i's feature profile, and $\mathbf{c}_j$ denotes cluster cl_j's centroid feature profile.

Each K-means iteration completes two tasks:

1. Assign each instance to the cluster with the closest centroid profile:

$$g_i \leftarrow \arg\min_{j} d(\mathbf{f}_i, \mathbf{c}_j) \tag{5.11}$$

2. Recompute the centroid profiles for each cluster based on the instances assigned to it using:

$$\begin{aligned} \mathbf{c}_j &= \arg\min_{\mathbf{c}} \sum_{i \in cl_j} d^2(\mathbf{f}_i, \mathbf{c}) \\ &= \frac{\sum_{i \in cl_j} \mathbf{f}_i}{|cl_j|} \end{aligned} \tag{5.12}$$

where $|cl_j|$ is the size of cl_j, or the number of prosumers in cluster cl_j.

These two tasks are repeated until the convergence criteria are met, which either means no changes in cluster assignment are observed or the number of iterations reaches the prespecified limit. Assuming the maximum number of iterations is set to $iter_{max}$, the iterative steps of K-means clustering for classifying feature profiles are detailed in Algorithm 4.

The random initial centroids selection coupled with this iteration process often traps the solution in the local minima, so it is advised to perform K-means repeatedly with different randomized initializations and pick the run with the lowest K-means objective function value denoted by $\min \sum_{i \in \mathcal{N}} d^2(\mathbf{f}_i, \mathbf{c}_{g_i})$. As an example, K-means clustering is run 1000 times with different initial centroids, and 100 prosumers are grouped into 6 clusters using their *grand coalition* cooperative load profiles as the feature vectors. Results are shown in Figure 5.2.

Algorithm 4 K-means Clustering of Feature Profiles

Initialize $\mathbf{c}_j, \forall j \in [1, K]$
$g_i^0 \leftarrow 0, \forall i \in \mathcal{N}$
for $iter \in [1, iter_{max}]$ **do**
  for $i \in \mathcal{N}$ **do**
    assign clusters: $g_i^{iter} \leftarrow \arg\min_j d(\mathbf{f}_i, \mathbf{c}_j)$
  for $j \in [1, K]$ **do**
    compute centroids: $\boldsymbol{c}_j \leftarrow \sum_{i \in cl_j} \mathbf{f}_i / |cl_j|$
  if $\boldsymbol{g}^{iter} = \boldsymbol{g}^{iter-1}$ **then**
    end iteration.
return $\boldsymbol{g}, \mathbf{c}$

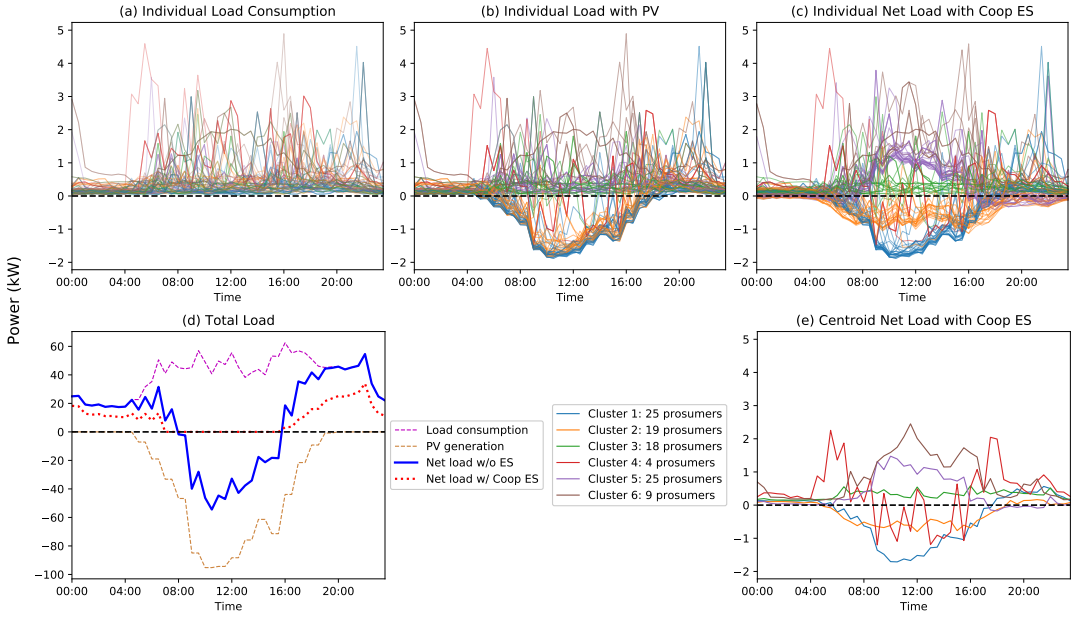


Figure 5.2: Clustering of 100 prosumers' load profiles: (a) individual load consumption only, (b) individual load without ES (load consumption + PV generation), (c) individual net load with cooperative ES operation (load consumption + PV generation + ES optimized to minimize the energy cost of the grand coalition), (d) grand coalition loads with and without cooperative ES operation, (e) 6 centroid net load profiles as a result of K-means clustering of individual net load with cooperative ES operation.

5.2.2 Hierarchical Clustering

Unlike K-means clustering, hierarchical clustering can be performed deterministically without a preset total number of clusters [159]. It is done by either an agglomerative (bottom-up) or a divisive (top-down) approach. In this thesis, only the agglomerative approach is investigated as it is preferred in power system applications for its higher computational efficiency [169]. It begins by treating each instance as an individual cluster, and proceeds by iteratively merging clusters that are most similar to each other. The Euclidean distance, defined in (5.10), is again used to measure the dissimilarity between clusters [152].

A linkage criterion needs to be selected in order to perform hierarchical clustering. There are four types of linkage criteria, namely *single*, *complete*, *average* and *ward*. Single linkage and complete linkage define the similarity between two clusters as the distance between the closest and farthest members respectively from the two clusters. Average linkage considers all instances in the two clusters by setting the similarity of two clusters as the average distance between all pairs of instances between the two. Using the ward criterion, clusters are merged with the objective to minimize the increase in the sum of squared distances between all members of the clusters and their corresponding cluster centroids. this increase in the sum of squared distances from merging clusters cl_j and cl_k is $\sqrt{\frac{|cl_j||cl_k|}{|cl_j|+|cl_k|}}d(\mathbf{c}_j, \mathbf{c}_k)$. Regardless of the linkage criterion, the merging of clusters stops when the desired number of clusters is reached.

Using the same inputs as Figure 5.2, the results of hierarchical clustering of 100 prosumers into 6 groups are shown in Figure 5.3. Because the process of hierarchical clustering does not involve the computation of centroids, all the centroid profiles are computed post clustering using (5.12). The goal of clustering in this chapter is to identify both similarities and dissimilarities among the prosumers, so clustering that yields very large or very small groups is not desirable. It can be seen that the ward linkage is able to produce the most even spread of prosumers among clusters. In this specific example it is the only linkage criterion that produces four clusters of more than 10 prosumers, and also the only one that does not single out one prosumer as its own cluster. This observation is also consistent with the results in [159], where ward linkage demonstrates

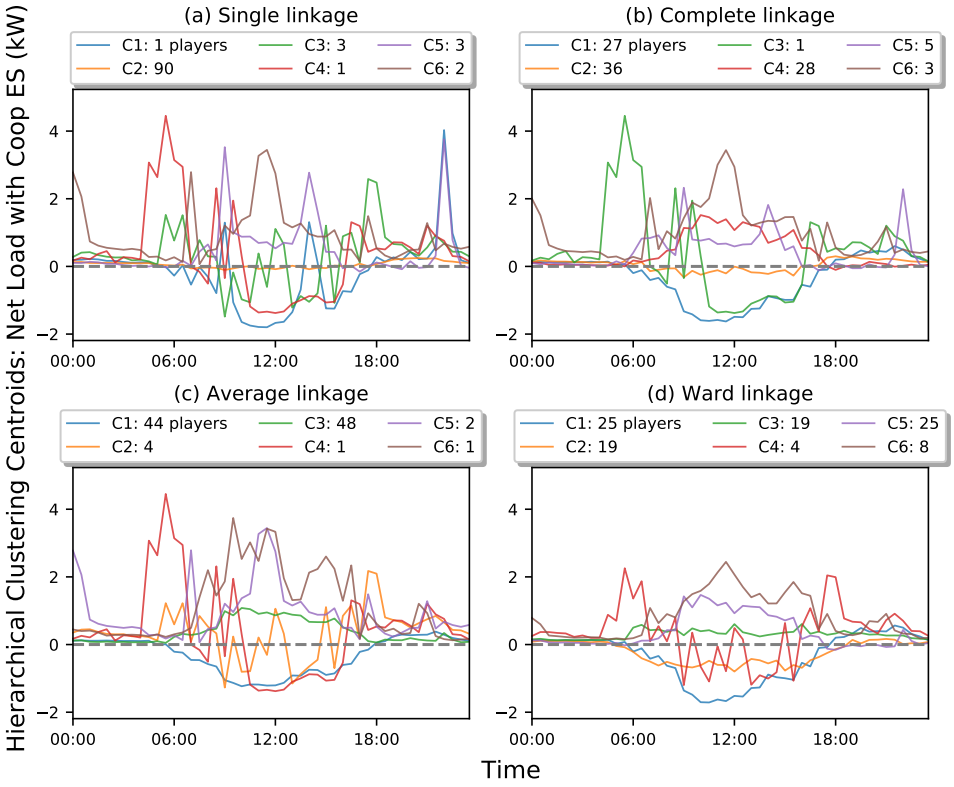


Figure 5.3: Centroid cooperative net load profiles of 6 clusters through hierarchical clustering of 100 prosumers with different linkage criteria: (a) single linkage, (b) complete linkage, (c) average linkage, and (d) ward linkage.

the best performance. Therefore, for estimating the nucleolus only the ward linkage criterion is considered for hierarchical clustering.

5.2.3 Gaussian Mixture Model

Both K-means clustering and hierarchical clustering are distance-based clustering techniques. GMM, on the other hand, is a probabilistic technique. It is part of the finite mixture models (FMM) family, which develops a probability distribution function (PDF) that is a mixture of K components: $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_K\}$. This PDF has the *maximum likelihood* to fit the given instances, given by

$$p(\mathbf{f}_i|\boldsymbol{\theta}) = \sum_{j=1}^{K} \alpha_j p(\mathbf{f}_i|\boldsymbol{\theta}_j) \tag{5.13}$$

where $\mathbf{f}_i$ denotes prosumer i's load profile, $\boldsymbol{\theta}_j$ is the jth component defined by $\boldsymbol{\theta}_j = \{\theta_{jt} \mid t \in [1, R]\}$, and α_j is the mixing probability or weight of component $\boldsymbol{\theta}_j$. It requires $\alpha_j \geq 0$, $j = 1, 2, ..., K$, and $\sum_{j=1}^{K} \alpha_j = 1$.

In the case of GMM, $\boldsymbol{\theta}_j = (\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, where $\boldsymbol{\mu}_j$ is the R-dimensional mean vector, and $\boldsymbol{\Sigma}_j$ is the covariance matrix. In other words, θ_{jt} follows a Gaussian distribution with μ_{jt} as the mean, and $\{\theta_{j1}, ..., \theta_{jR}\}$ are correlated defined by the covariance matrix $\boldsymbol{\Sigma}_j$. (5.13) can be rewritten as

$$p(\mathbf{f}_i|\boldsymbol{\theta}) = \sum_{j=1}^{K} \alpha_j \frac{exp\Big\{ -\frac{1}{2}(\mathbf{f}_i - \boldsymbol{\mu}_j)^\intercal \boldsymbol{\Sigma}_j^{-1}(\mathbf{f}_i - \boldsymbol{\mu}_j)\Big\}}{(2\pi)^{\frac{R}{2}}|\boldsymbol{\Sigma}_j|^{\frac{1}{2}}}$$

The maximum likelihood estimate of $\{\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}\}$ is often obtained through an iterative *expectation-maximization* (EM) algorithm [158, 168]. To prevent convergence to local optima, GMM can be run repeatedly with different initial conditions [162]. Finally, clustering is completed by assigning each instance to the component with the highest *posterior probability*:

$$p(k|\mathbf{f}_i) = \frac{\alpha_k p(\mathbf{f}_i|\boldsymbol{\theta}_k)}{\sum_{j=1}^{K} \alpha_j p(\mathbf{f}_i|\boldsymbol{\theta}_j)}, \quad k = 1, ..., K$$

To implement a Gaussian Mixture Model (GMM), the structure of the covariance matrix $\boldsymbol{\Sigma}$ needs to be specified. For each component, a *full* covariance matrix has uncorrelated dimensions, while a *diagonal* covariance matrix has correlated dimensions. A *spherical* covariance matrix is a special type of diagonal covariance matrix that has the same covariance across all dimensions. Lastly, if all the components share the same covariance matrix, the covariance matrix type becomes *tied*. In this study, full, diagonal, spherical covariance matrices indicate they are unshared among components, and a tied covariance matrix has uncorrelated dimensions.

GMM is used to cluster 100 prosumers into 6 groups using the same inputs as Figure 5.2 and postprocessed centroid profiles. The widely used Bayesian information criterion (BIC) is applied to different covariances with lower values indicating better performances of GMM [162]. BIC scores for spherical, diagonal, full, and tied variance matrices are [3051, 1430, -4988, 1200], suggesting that the full covariance outperforms the others. This is consistent with the results shown in Figure 5.4, where the most even spread of prosumers among clusters can be achieved with a full covariance matrix. For the same reason explained in Section 5.2.2, only GMM clustering with full covariance matrices is considered for the rest of the study.

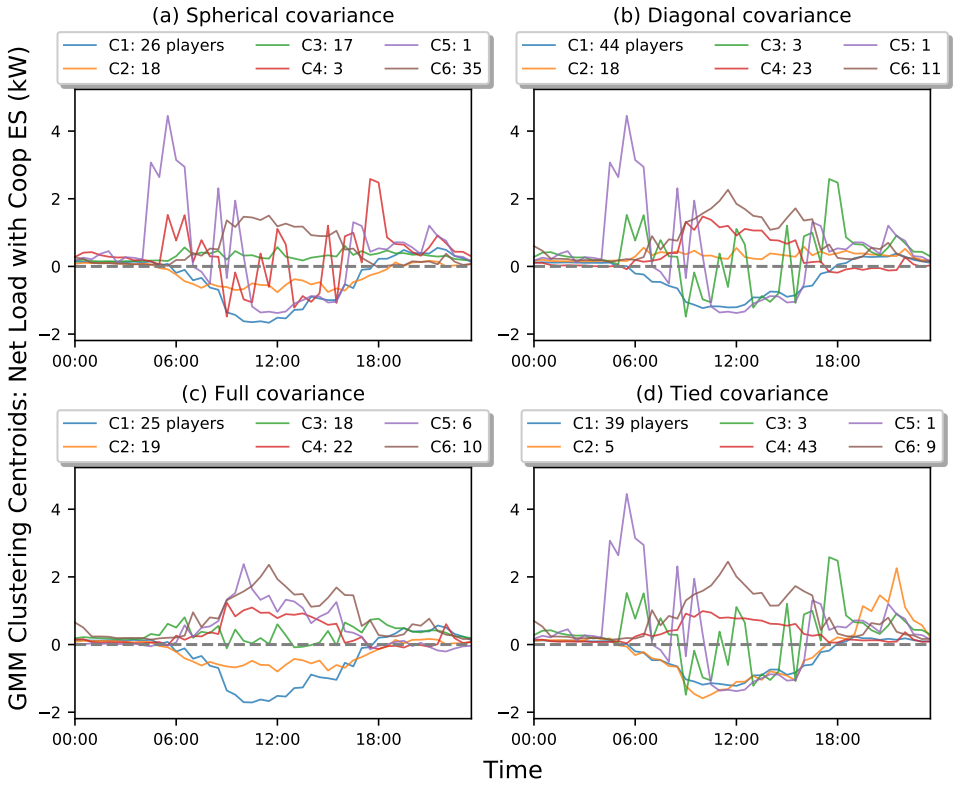


Figure 5.4: Centroid cooperative net load profiles of 6 clusters through GMM clustering of 100 prosumers with different covariance matrices: (a) spherical covariance matrix, (b) diagonal covariance matrix, (c) full covariance matrix, and (d) tied covariance matrix.

5.2.4 Prosumer Clustering Features and Principal Component Analysis

When applying clustering to classify electricity customers, the choice of the clustering features has a direct impact on the clustering results [144]. In identifying typical load patterns, customer load profiles are naturally used as the clustering features. However, these might not be the best choice for categorizing prosumers in coalitional energy management schemes as some prosumers energy profiles are not fixed, but dependent on the coalition they are in. Therefore, the following two clustering features suitable for identifying similarities among prosumers are proposed and investigated for a coalitional energy management scheme.

1. *Grand coalition cooperative energy profile* $\mathbf{p} := \{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N\}$. Although prosumers' DER operation needs to be simulated under different coalition configurations to calculate these coalitions' values, the grand coalition yields the actual total

profit. The cooperative energy profile of Prosumer i may differ based on the coalition they are participating in, but their grand coalition cooperative energy profile $\mathbf{p}_i$ is the most representative of their actual contribution to the scheme. $\mathbf{p}_i$ captures information about both the prosumer's fixed energy consumption/generation and their energy storage operation. Based on (3.10),

$$\mathbf{b}^{\mathcal{N}*} = \arg\min_{\mathbf{b}^{\mathcal{N}}} \sum_{t=1}^{R} F_t^{\mathcal{N}}(\mathbf{b}^{\mathcal{N}}) \tag{5.14}$$

$$\mathbf{p}_i = \mathbf{q}_i + \mathbf{b}_i^{\mathcal{N}*}, \forall i \in \mathcal{N} \tag{5.15}$$

Figure 5.4 shows an example of using GMM clustering to group prosumers based on their *grand coalition cooperative energy profile*.

2. *Grand coalition marginal contribution profile* $\Delta\mathbf{F} := \{\Delta\mathbf{F}_1, \Delta\mathbf{F}_2, ..., \Delta\mathbf{F}_N\}$. Specific to the coalitional energy management scheme, the value of a coalition corresponds to its energy cost savings, and the payoff is closely linked to a prosumer's marginal contributions to coalitions. Therefore, $\Delta\mathbf{F}$ is investigated as a clustering feature. Expanding the definition of the optimal ES operation in (5.14) to any coalition $\mathcal{T}$:

$$\mathbf{b}^{\mathcal{T}*} = \arg\min_{\mathbf{b}^{\mathcal{T}}} \sum_{t=1}^{R} F_t^{\mathcal{T}}(\mathbf{b}^{\mathcal{T}})$$

For any prosumer i, $\Delta\mathbf{F}_i := \{\Delta F_{i1}, \Delta F_{i2}, ..., \Delta F_{iR}\}$ is defined as the cost savings profile from having i join the existing coalition of $\mathcal{N} \setminus \{i\}$. For each timestep t, the cost savings are calculated as:

$$\Delta F_{it} = F_t^{\mathcal{N}\setminus\{i\}}(\mathbf{b}^{\mathcal{N}\setminus\{i\}*}) + F_t^{\{i\}}(\mathbf{b}^{\{i\}*}) - F_t^{\mathcal{N}}(\mathbf{b}^{\mathcal{N}*})$$

Figure 5.5 shows an example of using GMM clustering to group prosumers based on their *grand coalition marginal contribution profile*.

In clustering applications, the dimensionality of the clustering feature can have a large impact on the clustering computation time. In order to reduce the dimensionality of the dataset, principal component analysis (PCA) is a popular method [166]. It computes the eigenvectors of the covariance matrix of the input data and ranks them based on their corresponding eigenvalues in a descending order. Based on the desired level of covariance

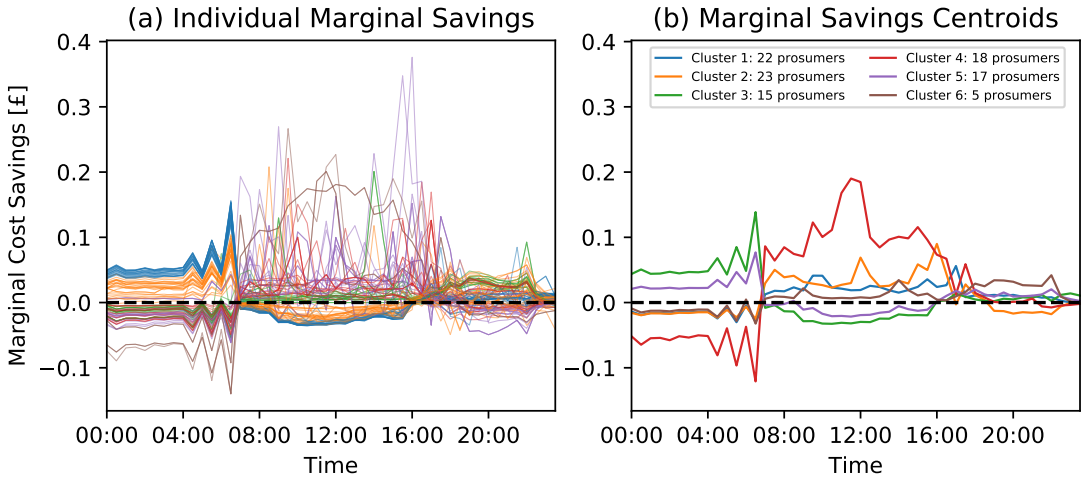


Figure 5.5: GMM clustering of 100 prosumers' marginal contribution profiles: (a) prosumers' individual marginal contribution profiles color-coded based on the clusters they belong to, (b) centroids of 6 clusters' marginal contribution profiles.

captured, the first R_{pc} eigenvectors are then selected as the principal components, and the data is projected from the original R-dimension space to this new R_{pc} subspace ($R_{pc} < R$) through linear transformation. Note that the number of dimensions of both the cooperative energy profile and the marginal contribution profile are equal to the number of timesteps considered in the model. For a 24-hour timespan with half-hour time intervals, this dimensionality is only $R = 48$. Due to this low dimensionality, the computation time of clustering is negligible compared to the nucleolus computation. As a result, the benefit of PCA becomes insignificant in this application. Therefore, besides using PCA to demonstrate the effectiveness of clustering in categorizing prosumers, it is not directly implemented in the nucleolus estimation models. A detailed discussion of this point can be found in Section 5.4.2.

Nevertheless, PCA can be used to visualize the clustering results by plotting the main components of every prosumer's input data on a 2-D map. Using the cooperative energy profile as an example, Figure 5.6 illustrates the distribution of K-means clusters based on their principal components, where each circle represents a prosumer coded with their corresponding cluster color. It can be seen that K-means clustering effectively separates prosumers based on their first two principal components. Therefore, to demonstrate the principal components of the results from hierarchical clustering with ward linkage and GMM clustering with full covariance matrix, only the first two components are shown,

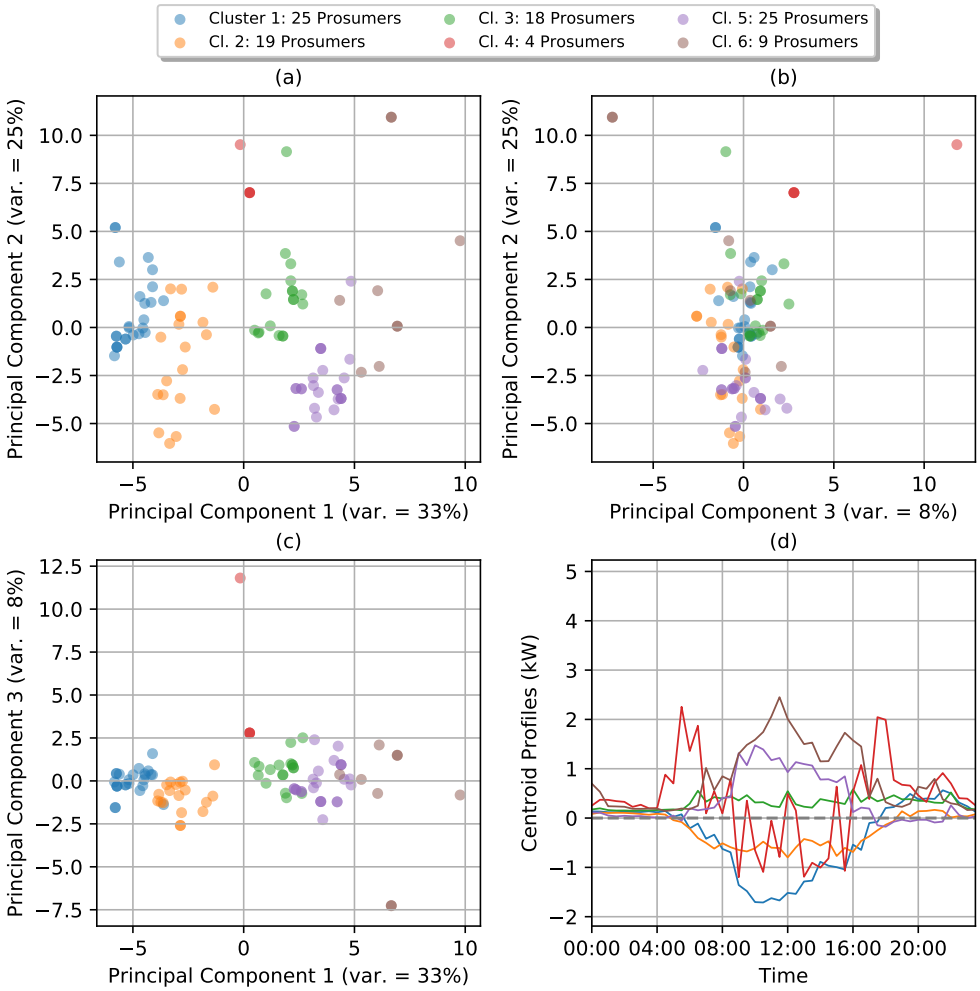


Figure 5.6: Visualization of 3 principal components of the K-means clustering results of 100 prosumers' cooperative energy profiles ($K = 6$). Each data point in (a)-(c) represents a prosumer color-coded by its cluster in a 2-D space with two principal components as the coordinates, (d) centroid cooperative energy profiles of the 6 clusters obtained through K-means clustering.

see Figure 5.7 and 5.8. It is interesting to see that hierarchical clustering produces very similar partitions as K-means, and GMM differs quite significantly on Clusters 1, 3 and 4. Roughly compared to the clustering results of K-means and hierarchical, GMM groups Clusters 3 and 4 into one cluster, and divides Cluster 1 into two clusters.

As a comparison, PCA is also implemented on the marginal contribution profiles, and the GMM clustering results are shown in Figure 5.9. There are a few interesting observations. First, the first principal component using the marginal contribution profiles accounts for over 50% of the variance in the data, significantly higher than that of the cooperative energy profiles (33%). This indicates that the formulation of the marginal

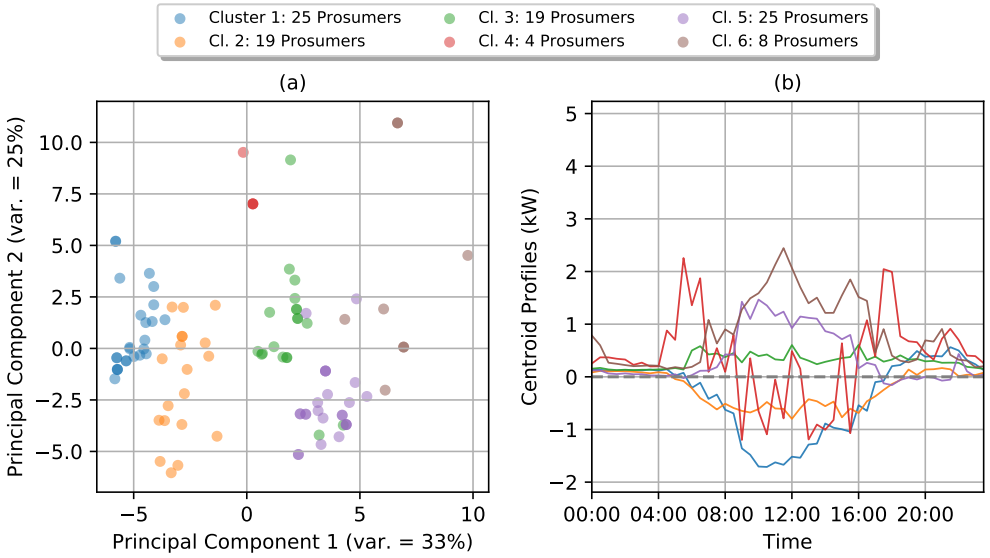


Figure 5.7: Visualization of 2 principal components of the hierarchical clustering results of 100 prosumers' cooperative energy profiles ($K = 6$). Each data point in (a) represents a prosumer color-coded by its cluster in a 2-D space with the two main principal components as the coordinates, (d) centroid cooperative energy profiles of the 6 clusters obtained through hierarchical clustering.

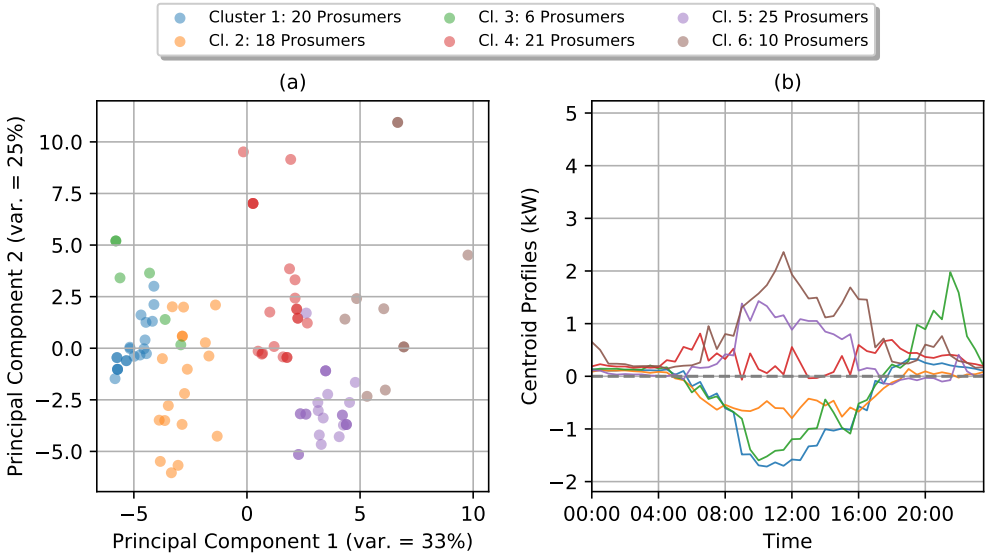


Figure 5.8: Visualization of 2 principal components of the GMM clustering results of 100 prosumers' cooperative energy profiles ($K = 6$). Each data point in (a) represents a prosumer color-coded by its cluster in a 2-D space with the two main principal components as the coordinates, (d) centroid cooperative energy profiles of the 6 clusters obtained through GMM clustering.

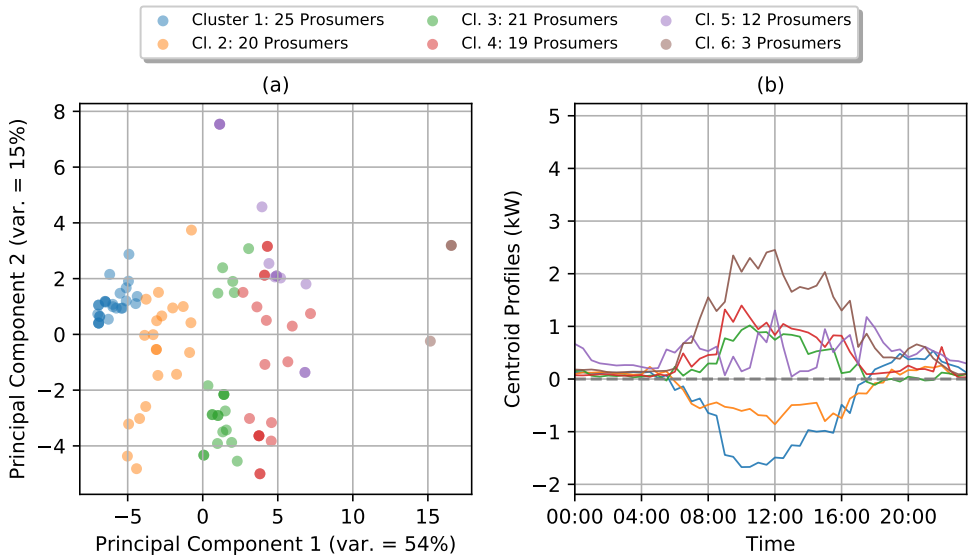


Figure 5.9: Visualization of 2 principal components of the GMM clustering results of 100 prosumers' marginal contribution profiles ($K = 6$). Each data point in (a) represents a prosumer color-coded by its cluster in a 2-D space with the two main principal components as the coordinates, (d) centroid cooperative energy profiles of the 6 clusters obtained through GMM clustering using marginal contribution profiles as the clustering feature.

contribution profiles re-maps the prosumer data in a way that magnifies the first principal component, which facilitates the clustering process in identifying the major differences within the prosumer data. Second, the clustering results of the marginal contribution profiles have one less cluster with a centroid where energy is exported during the day. On the other hand, it introduces a finer categorization of prosumers that predominantly import energy. Third, using the marginal contribution profiles as clustering features produces smoother centroid energy profiles, which means clustering based on the marginal contribution profiles is less likely to consider prosumers with large fluctuations and spikes as outliers and isolate them into small clusters. These observations suggest that applying marginal contribution profiles as clustering features is likely to produce preferable clustering results for nucleolus estimation which is confirmed in a case study in Section 5.4.3.

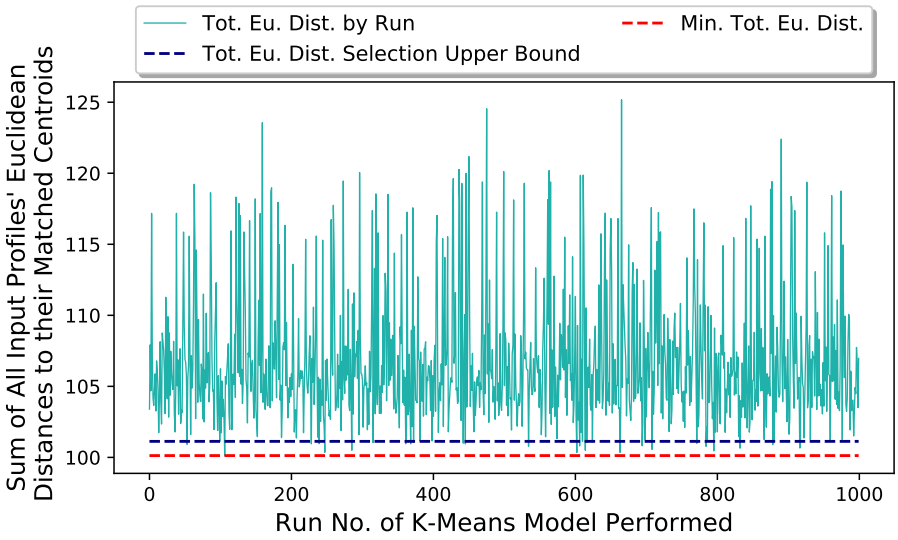


Figure 5.10: Total Euclidean distances for 1000 K-means (K=8) clustering runs on 50 prosumers' cooperative energy profiles. The maximum total Euclidean distance considered is set at 1% above the lowest total Euclidean distance found in the 1000 runs.

5.3 Nucleolus Estimation with Prosumer Clustering

This section focuses on the application of various clustering techniques in estimating the nucleolus, and the validation of the estimated results both in terms of the reduced computation time and the estimation accuracy. It is organized as follows.

Section 5.3.1 develops the concept of *clustered players* and formulates a simple K-means clustering based method to compute their nucleolus. Section 5.3.2 proposes a novel method to estimate the full nucleolus of the coalitional energy management scheme using different clustering techniques. Section 5.3.3 uses stratified random sampling to develop a verification tool to evaluate the performance of the proposed nucleolus estimation method.

5.3.1 Simple K-Means for Computing the Nucleolus for Clustered Players

Various clustering methods and clustering features were introduced in Section 5.2. As a benchmark, this subsection incorporates a simple K-means clustering algorithm based on the grand coalition cooperative energy profile as the clustering feature in order to reduce the full scheme's computational complexity.

As mentioned in Section 5.2.1, while implementing K-means clustering, some initial seed selections can lead to local optimal solutions. Due to the low number of features (i.e. 48 features for one day) and the small sample size (i.e. ≤ 400 prosumers), it would not significantly impact the computational time if the K-means model was run 1000 times with different initial centroids (see case study in Section 5.4.2). An important benefit of this approach is that all the K-means results can be compared and the final clusters can be selected based on additional criteria. As the selection of K is limited by the computational complexity, it is preferable that the prosumers are as evenly represented in the clusters as possible to avoid having small clusters of outliers. To this end, using the total Euclidean distance of each K-means run as a metric, an upper bound of 1% above the lowest total Euclidean distance is set. All the clustering results that have a total Euclidean distance below this upper bound can then be compared against each other, and the one with the most even distribution of prosumers is selected. Figure 5.10 shows an example of the Euclidean distance variation as a result of clustering 50 prosumer load profiles. It can be seen that at least 20 out of 1000 clustering runs produce a total Euclidean distance within the allowed range.

The purpose of clustering in this chapter is to reduce the number of modeled players in the cooperative game model. Therefore, each cluster of prosumers is defined as a *clustered player* cl_j, and the model becomes a K-player cooperative game, where K is the input number of clusters for K-means. The grand coalition of all clustered players is $cl_{\mathcal{K}} := \{cl_j | j \in \mathcal{K}\}$ where $\mathcal{K} = \{1, 2, ..., K\}$. For $\mathcal{U} \subseteq \mathcal{K}$, the corresponding coalition of clustered players is defined as $cl_{\mathcal{U}} := \{cl_j | \forall j \in \mathcal{U}\}$. Therefore, $cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}$. While running the cooperative game model for the clustered players, each clustered player is assigned the sum of all its member prosumers' original consumption and generation profiles as its own consumption and generation, and all of its member prosumers' ES systems as its own ES capacity.

The following are the steps to implement this cooperative game model with clustered players.

1. *Load processing*: all the model inputs are gathered, including the number of prosumers N, their net load profiles $\mathbf{q}$, and their ES system constraints. Following

(5.14), the coalitional energy cost is minimized for the grand coalition and the optimized ES operation as $\mathbf{b}^{\mathcal{N}*}$ is recorded.

2. *Prosumer clustering*: each prosumer's grand coalition cooperative energy profile is constructed: $\mathbf{p}_i = \mathbf{q}_i + \mathbf{b}_i^{\mathcal{N}*}, \forall i \in \mathcal{N}$. The K-means clustering model detailed in Algorithm 4 is then run with $\mathbf{p}$ as inputs, along with a chosen K. It yields the clustering assignment $\boldsymbol{g} : g_i = j \mid i \in cl_j, \forall i \in \mathcal{N}$. Therefore, the cluster that prosumer i belongs to can be denoted as cl_{g_i}.

3. *Clustered player cooperative game formulation*: in order to preserve as much of the original inputs as possible, all the prosumers' load and ES inputs are directly applied to the clustered player coalitions. A functional Q is defined to convert $cl_{\mathcal{U}}$ to a coalition of prosumers, where $Q(cl_{\mathcal{U}}) = \{i \mid cl_{g_i} \in cl_{\mathcal{U}}\}$. Therefore, for any coalition of clustered players $cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}$, the total energy cost at any timestep t is defined as

$$F_t^{Q(cl_{\mathcal{U}})}(\mathbf{b}^{Q(cl_{\mathcal{U}})}) = \sum_{cl_{a_i} \in cl_{\mathcal{U}}} \left\{ r_t^{im}[q_{it} + b_{it}]_+ + r_t^{ex}[q_{it} + b_{it}]_- \right\} \tag{5.16}$$

Since all the clustered player inputs have been converted into prosumer inputs, it becomes possible to use (5.2) to compute the coalitional energy costs for all the clustered player coalitions $G^R(Q(cl_{\mathcal{U}}))$. Using a similar conversion, the value of clustered player coalitions can be calculated as

$$v(Q(cl_{\mathcal{U}})) = \sum_{cl_{a_i} \in cl_{\mathcal{U}}} G^R(\{i\}) - G^R(Q(cl_{\mathcal{U}}))$$

The nucleolus of this clustered player cooperative game $\boldsymbol{\nu}_{cl} := \{\nu_{cl_j} | \forall cl_j \in cl_{\mathcal{K}}\}$ can then be computed by iteratively minimizing the excesses of the clustered player coalitions using Algorithm 3.

An example of 14 prosumers grouped into five clustered players is shown in Figure 5.11. It displays the energy costs, cost savings from both cooperating within the clusters and the clustered player nucleolus. It can be seen that all the clustered players are guaranteed savings, and the savings from cooperating within the clusters are significantly lower than the nucleolus, which further confirms that the clustering process tends to group prosumers that exhibit similar behaviors together.

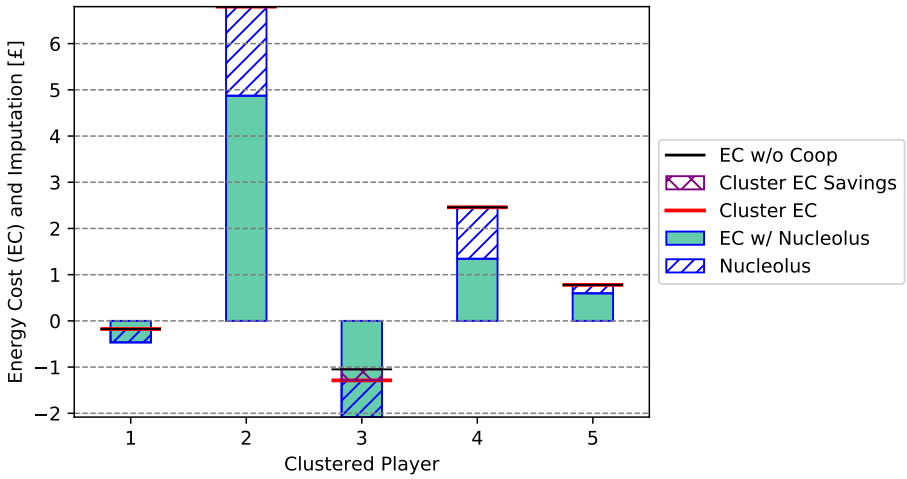


Figure 5.11: Energy costs (EC) and imputation for clustered players. For each clustered player, the black line represents the energy cost without cooperation, the mesh bar represents the energy cost savings from in-cluster cooperation, the red line represents the energy cost with only in-cluster cooperation, the hashed bar represents the additional nucleolus payoff allocation from participating in the grand coalition, and the solid bar represents the final energy cost for participating in the grand coalition.

4. *De-clustering*: the nucleolus computed for the clustered players needs to be distributed to their member prosumers. Although the value for each single prosumer coalition is always zero: $v(\{i\}) = 0, \forall i \in \mathcal{N}$, the in-cluster cooperation sometimes results in the value of a clustered player being positive: $v(Q(\{cl_j\})) > 0$. This can be considered additional savings for each clustered player that also needs to be distributed to its member prosumers. There are numerous ways to allocate these in-cluster savings. Considering a prosumer's contributions are loosely linked to their energy load magnitude, which is reflected in their individual energy cost, the savings are distributed, as a simple example, within clusters in proportion to the absolute value of each individual's energy cost without cooperation:

$$\hat{\nu}_i = [\nu_{cl_{g_i}} + v(Q(\{cl_{a_i}\}))](\left|G^R(\{i\})\right| / \sum_{i \in cl_j} \left|G^R(\{i\})\right|)$$

Using the same inputs as Figure 5.11, Figure 5.12 compares the estimated nucleolus with the true nucleolus. Each marker represents a prosumer, and the closer they are to the diagonal line the better the estimation is. It can be seen that large discrepancies do occur for some of the prosumers.

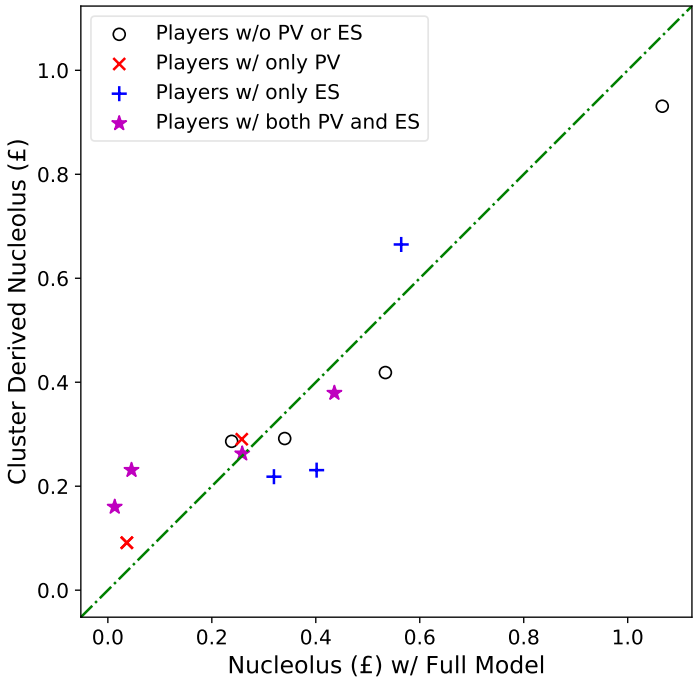


Figure 5.12: Comparison of the estimated nucleolus with simple K-mean clustering and the true nucleolus. The closer the data points are to the diagonal line, the more accurate the estimation is.

Clustering reduces the number of coalitions of the cooperative game from 2^N to 2^K, where $K \leq N$. Intuitively, the lower K is, the shorter the model takes to run, but the more information the model loses, and the less accurate the model becomes. Because there is no means to distinguish the contributions of individual prosumers within each clustered player, the payoff division within clustered players is somewhat arbitrary and the resulting individual payoffs could deviate significantly from the actual nucleolus of the original scheme. With the aim to scale up the cooperative game theory based energy management scheme with stabilizing payoffs, the following subsection proposes a new method for estimating the full nucleolus with high accuracy.

5.3.2 Full Nucleolus Estimation with Clustering

Coalitions with the most diverse mix of prosumers tend to create the highest values for their members when independent from the grand coalition, and therefore they are most likely to defect. Reversely, a payoff allocation that prevents these coalitions from defecting is likely to keep other coalitions in the scheme as well, just like the nucleolus

does. Therefore, the nucleolus can be estimated by minimizing the excesses of these coalitions. Clustering, as described in the previous section, offers a foundational step to diversify coalition samples. In this subsection a nucleolus estimation algorithm is proposed through clustering prosumers using different feature vectors. All the clustering techniques described in Section 5.2 are investigated to find the most suitable clustering technique for estimating the nucleolus.

Clustering separates dissimilar prosumers, but it does not provide any information about the differences between prosumers within the same cluster. The nucleolus estimation method presented in Section 5.3.1 can be highly inaccurate because the contributions of individual prosumers cannot be distinguished within each cluster, leading to arbitrary in-cluster division of payoffs. To solve this, instead of treating each cluster as a distinctive prosumer, every prosumer is paired with clusters they do not belong to to form coalition samples for excess minimization. This way, not only is the diversity within the coalition samples ensured, but each prosumer's individuality is preserved as well. The proposed method for estimating the nucleolus is as follows:

1. Based on a certain prosumer *feature* (see Section 5.2.4), prosumers are separated into K clustered players, the complete set of which is denoted as $cl_{\mathcal{K}} := \{cl_j | j \in \mathcal{K}\}$ where $\mathcal{K} = \{1, 2, ..., K\}$. The clustering assignment is $\boldsymbol{g} : g_i = j \mid i \in cl_j, \forall i \in \mathcal{N}$.

2. A coalition of clustered players is defined as $cl_{\mathcal{U}} := \{cl_j | \forall j \in \mathcal{U}\}, cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}$. Any $cl_{\mathcal{U}}$ can be converted to a coalition of prosumers by using $Q(cl_{\mathcal{U}}) = \{i \mid cl_{g_i} \in cl_{\mathcal{U}}\}$. Therefore, all the prosumer coalitions derived from coalitions of clustered players can be written as $\mathfrak{U} = \{\mathcal{T} \mid \mathcal{T} = Q(cl_{\mathcal{U}}), \forall cl_{\mathcal{U}} \subseteq cl_{\mathcal{K}}\}$.

3. For each prosumer $i \in \mathcal{N}$, the set of coalitions of clustered players that do not contain i is $\mathbf{U}_i = \{cl_{\mathcal{U}} \mid \forall cl_{\mathcal{U}} \subseteq (cl_{\mathcal{K}} \setminus \{cl_{g_i}\})\}$. If prosumer i is then added to each of the coalitions in $\mathbf{U}_i$, this new set of prosumer coalitions becomes $\mathfrak{S}_i = \{\mathcal{T} \mid \mathcal{T} = Q(cl_{\mathcal{U}}) \cup \{i\}, \forall cl_{\mathcal{U}} \in \mathbf{U}_i\}$. $\mathfrak{S}_i$ is then constructed for every prosumer, all of which are combined into a single set $\mathfrak{S} = \{\mathcal{T} \mid \mathcal{T} \in \mathfrak{S}_i, \forall i \in \mathcal{N}\}$.

4. Finally, the set of prosumer coalition samples for excess minimization is constructed, denoted by $\mathfrak{T}^{cl_\mathcal{K}} = \mathfrak{S} \cup \mathfrak{U}$.

In addition to the coalitions of whole clusters denoted by $\mathfrak{U}$ (used in Section 5.3.1), the coalition sample set $\mathfrak{T}^{cl_\mathcal{K}}$ also includes $\mathfrak{S}$ that groups every prosumer i with all possible coalitions of clustered players that do not contain i. This diversity based paring means $\mathfrak{S}_i$ very likely contains the coalition samples that can provide the highest payoffs to prosumer i. Therefore, by minimizing the excesses of these coalitions, prosumer i's tendency to defect from the grand coalition can be significantly decreased. The nucleolus estimation $\hat{\boldsymbol{\nu}}$ can thus be obtained by implementing Algorithm 3 with $\mathfrak{T}^* = \mathfrak{T}^{cl_\mathcal{K}}$. The size of $\mathfrak{T}^{cl_\mathcal{K}}$ is bounded by

$$|\mathfrak{T}^{cl_\mathcal{K}}| \leq |\mathfrak{U}| + |\mathfrak{S}| \tag{5.17}$$

$$\leq (2^K - 1) + (2^{K-1}N) \tag{5.18}$$

$$= 2^{K-1}(N+2) - 1 \tag{5.19}$$

Where the inequalities in (5.17) and (5.18) both occur when clustering yields one or more single-prosumer clustered players. Therefore, given a selected number of clusters K, both the number of coalitions that require a value calculation and the number of iterations in the nucleolus calculation become $\sim \mathcal{O}(N)$, a significant reduction compared to the full nucleolus calculation $\sim \mathcal{O}(2^N)$ as explained in Section 5.2.

5.3.3 Evaluating Nucleolus Estimation with Stratified Random Sampling

The proposed nucleolus estimation method involves three different clustering techniques, namely K-means clustering, Hierarchical clustering with ward linkage, GMM with full covariance matrices. For each of these clustering techniques, two types of clustering features are investigated, namely the grand coalition cooperative energy profile, and the grand coalition marginal contribution profile. In order to compare these methods, there needs to be a tool that is independent of all the clustering techniques to evaluate the estimated nucleolus $\hat{\boldsymbol{\nu}}$.

Coalitional stratified sampling has been previously used to reduce the computation time of the Shapley value in Chapter 4. In this approach, $\pi(\mathcal{N})$ is defined as the set of all possible permutations with prosumer set $\mathcal{N}$, where each permutation is a sequence of all the prosumers in a specific order. For a given permutation $O \in \pi(\mathcal{N})$, $O(l)$ denotes the prosumer in position l, and $Pre^i(O)$ denotes the set of predecessors of prosumer i, or the coalition of prosumers that are positioned ahead of prosumer i in the prosumer sequence represented by O. Therefore, if $i = O(l)$, $Pre^i(O) = \{O(1), O(2), ..., O(l-1)\}$. $\pi(\mathcal{N})$ is then divided into subpopulations, each of which is called a stratum that has the same number of predecessors for each prosumer. A stratum, or a stratified set of prosumer permutations, is denoted as $P_{il} := \{O \in \pi(\mathcal{N}) \mid O(l) = i, \forall i, l \in [1, N]\}$. Therefore, P_{il} contains every permutation $O \in \pi(\mathcal{N})$, in which prosumer i is in position l.

Considering the computational complexity of nucleolus is also directly linked to the number of coalitions, this sampling method has the potential to be used for direct nucleolus estimation as well.

5.3.3.1 Stratified Random Sampling for Direct Nucleolus Estimation

To adapt this method to the nucleolus estimation, the following steps are implemented:

1. a fixed stratum sample size m is set. For each stratum P_{il}, a random permutation sample M_{il} of this size m.

2. A set of coalitions $\mathfrak{M}_m := \{\mathcal{T} \mid \mathcal{T} = Pre^i(O) \cup i, \forall O \in M_{il}, \forall i, l \in [1, N]\}$ is defined. Note that in Strata P_{iN}, regardless of the chosen prosumer i, $Pre^i(O) \cup i = \mathcal{N}$. And in strata P_{i1}, $Pre^i(O) \cup i = \{i\}$. This means it is guaranteed that $\mathcal{N} \in \mathfrak{M}_m$, and $\{i\} \in \mathfrak{M}_m, \forall i \in \mathcal{N}$. Therefore, the total number of coalitions in the sample set $\mathfrak{M}_m$ is limited to:

$$|\mathfrak{M}_m| \leq mN^2 - (3m-2)N + 1 \tag{5.20}$$

3. use (5.6) to calculate $v(\mathcal{T}), \forall \mathcal{T} \in \mathfrak{M}_m$.

4. The nucleolus estimation $\hat{\boldsymbol{\nu}}$ can thus be obtained by implementing Algorithm 3 with $\mathfrak{T}^* = \mathfrak{M}_m$.

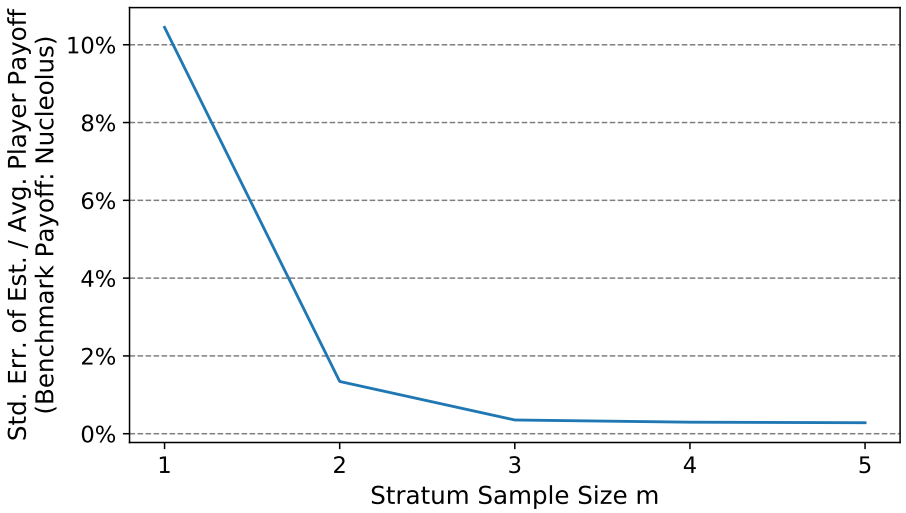


Figure 5.13: Estimation performance verification of the stratified random sampling method for a 14-prosumer coalitional energy management scheme using the full nucleolus as a benchmark: varying stratum sample sizes ($m = 1, 2, 3, 4, 5$). For an N-prosumer scheme, the total number of samples is limited by: $mN^2 - (3m - 2)N + 1$

.

To test the sampling method in estimating the nucleolus of the coalitional energy management scheme, a range of prosumer numbers is selected ($N = 10, 14, 20, 30, 40, 50$). For the lower numbers of prosumers ($N \leq 14$), it is feasible to compute the actual nucleolus despite the exponential complexity, and evaluate both the computation time reduction and the estimation accuracy of the proposed sampling method. Then the numbers of prosumers is increased to test the scalability of the proposed sampling method.

First, the number of prosumers is set to 14, which is the maximum number of prosumers that the model could compute the actual nucleolus $\boldsymbol{\nu}$ for within 24 hours. Then with a varying stratum sample size m the estimation errors under each sample size are compared. The metric chosen here is the *standard error of the estimate* as a percentage of the average prosumer payoff:

$$\frac{\sigma(\hat{\mathbf{x}}, \mathbf{x})}{\overline{x}} = \frac{\sqrt{\sum_i^N (\hat{x}_i - x_i)^2 / N}}{v(\mathcal{N})/N} \times 100\% \tag{5.21}$$

where $\hat{\mathbf{x}}$ and $\mathbf{x}$ represent the estimated and benchmark imputations respectively. Let $\hat{\mathbf{x}} = \hat{\boldsymbol{\nu}}$ and $\mathbf{x} = \boldsymbol{\nu}$, $\left(\frac{\sigma(\hat{\nu}, \nu)}{v(\mathcal{N})/N}\right) \times 100\%$ is plotted for a varying stratum sample size ($m \in [1, 5]$) on Figure 5.13. A knee point at $m = 3$ can be clearly identified with a

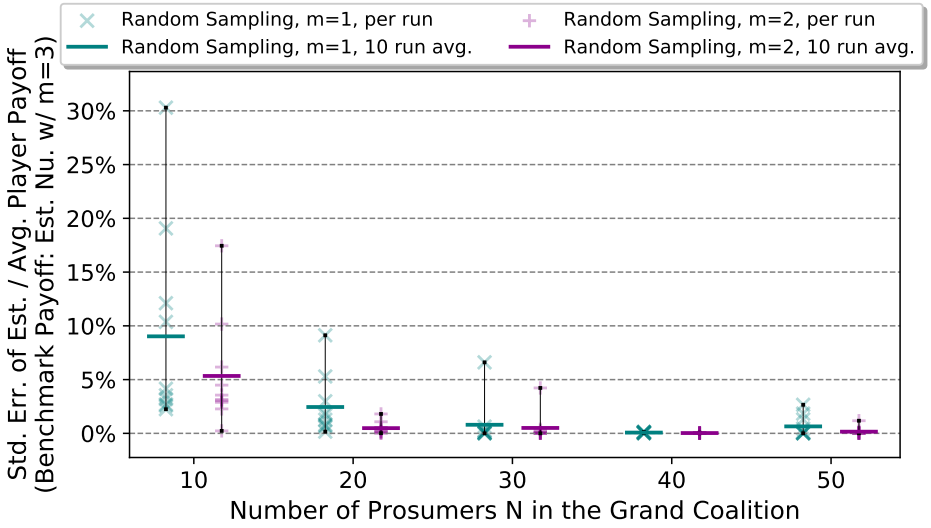


Figure 5.14: Estimation performance verification of the stratified random sampling method with smaller sample sizes ($m = 1, 2$) using $m = 3$ as a benchmark: varying the scheme size ($N = 10, 20, 30, 40, 50$). Each '$\times$' or '$+$' represents the estimation error of a model run, and the solid horizontal line represents the 10-run average.

Table 5.1: Nucleolus Computation Time (s)

No. prosumers	**14**	**20**	**30**	**50**	**70**	**100**
Full nucleolus	5E+4	N/A	N/A	N/A	N/A	N/A
Est. nu. ($m = 3$)	18	91	8E+2	2E+4	N/A	N/A
Est. nu. ($m = 2$)	10	43	3E+2	5E+3	4E+4	N/A
Est. nu. ($m = 1$)	5	15	81	1E+3	6E+3	5E+4

standard error of the estimate equal to only 0.35% of the average payoff allocation to the prosumers. Note that the number of prosumers in this example is limited to 14, and it is possible that the estimation performances of different m values may change as the size of the scheme increases.

In order to test the scalability of this stratified sampling method, first the knee point $m = 3$ is used to verify the estimation performances of lower m values for bigger schemes. Figure 5.14 plots the estimation performance of the stratified random sampling method with $m = 1, 2$ using $m = 3$ as the benchmark. It shows that despite the higher estimation error for $m = 1, 2$ in smaller schemes, their estimation results become closer to the $m = 3$ case as the size of the scheme increases.

Then the computation times of these methods are compared with each other. Because

the time scale for the coalitional energy management scheme is 24 hours, the model computation time needs to be within this time window so that the prosumers can be remunerated on time. Any computation time above 10 hours is considered to be impractical for this application. As can be seen in Table 5.2, stratified random sampling is able to significantly reduce the computation time compared to the full nucleolus calculation. It still has trouble scaling up the coalitional energy management scheme to more than 100 prosumers even with $m = 1$, but combining the findings of Figure 5.14 and Table 5.2, it is reasonable to use $m = 3$ for $N < 50$ and $m = 1$ for $50 \leq N \leq 100$ as benchmarks for other nucleolus estimation methods that may further improve the scalability of the scheme.

5.3.3.2 Stratified Random Sampling for Evaluating Nucleolus Estimation

Although it remains computationally challenging to use sampling to directly estimate the nucleolus, it can still be used as an unbiased way of evaluating the performance of the clustering based nucleolus $\hat{\boldsymbol{\nu}}$ estimation methods. Because the main desired feature of the nucleolus is that it prevents coalitions from defecting from the grand coalition, the coalitions' excesses can be used to measure the accuracy of the estimation.

To adapt this approach to evaluate the nucleolus estimation results,

1. A stratum sample size $m = 1$ is set.

2. For each stratum P_{il}, a random permutation sample M_{il} of size m is obtained.

3. A set of coalitions $\mathfrak{M}_m := \{\mathcal{T} \mid \mathcal{T} = Pre^i(O) \cup i, \forall O \in M_{il}, \forall i, l \in [1, N]\}$ is defined. The total number of coalitions in the sample set $\mathfrak{M}_m$ is limited to $|\mathfrak{M}_m| < mN^2$.

4. $v(\mathcal{T})$ is calculated using (5.6), $\forall \mathcal{T} \in \mathfrak{M}_m$. Under the estimated nucleolus $\hat{\boldsymbol{\nu}}$, the coalition samples' excesses $\boldsymbol{\varepsilon}_{\hat{\boldsymbol{\nu}},m} = \{\varepsilon(\hat{\boldsymbol{\nu}}, \mathcal{T}) \mid \forall \mathcal{T} \in \mathfrak{M}_m\}$ are calculated using (5.7). The number of positive excesses $|\boldsymbol{\varepsilon}^{pos}_{\hat{\boldsymbol{\nu}},m}|$ is obtained.

5. $m = m + 1$, and then Step 2 - 4 is repeated until:

$$\left| \frac{|\boldsymbol{\epsilon}^{pos}_{\hat{\boldsymbol{\nu}},m}|}{|\mathfrak{M}_m|} - \frac{|\boldsymbol{\epsilon}^{pos}_{\hat{\boldsymbol{\nu}},m-1}|}{|\mathfrak{M}_{m-1}|} \right| \times 100\% \leq \zeta \tag{5.22}$$

where ζ is the convergence threshold, representing the change in the percentage of the coalition samples that have positive excesses. In this thesis it is set to be 0.5%.

5.4 Case Studies

In this section, the proposed nucleolus estimation algorithms are tested with different clustering techniques. In the first case study, in order to evaluate the the estimation performance of the simple clustered player method (Section 5.3.1) and the proposed full estimation model (Section 5.3.2), their estimation results are compared against the benchmark stratified sampling method (Section 5.3.3.1. In the second case study, by varying the number of clusters, the trade-off between the computation time and the estimation accuracy of the proposed full estimation model is analyzed. The third case study compares the estimation performances of different clustering techniques, first on the computation time, and then on the estimation accuracy with different clustering features. Finally, the fourth case study scales up the scheme to higher numbers of prosumers and demonstrates the estimation performance of the chosen clustering technique.

Some of the model inputs are as follows: the domestic load data was measured in the Customer-Led Network Revolution trials [182]. the model time frame is 24 hours starting from the midnight of a sunny summer day in July. The PV systems are 4kW with fixed 20 degree tilt, simulated in PVWatts [183] using the London Gatwick solar data. The ES model has an energy capacity of 7 kWh, a maximum charge power of 3.5 kW, a maximum discharge power of 3.2 kW, both charging and discharging efficiencies of 95%, an initial state of charge of 50%, and a state of charge range of 20-95%. Both the PV and ES adoption rates are fixed at 50%, and ownerships are randomly assigned independently of each other. In other words, each prosumer can have a PV system, or an ES system, or both, or neither. The energy import price follows a UK Economy 7 residential rate structure: £0.08/kWh for midnight–7am, and £0.18/kWh for 7am–midnight [29], and

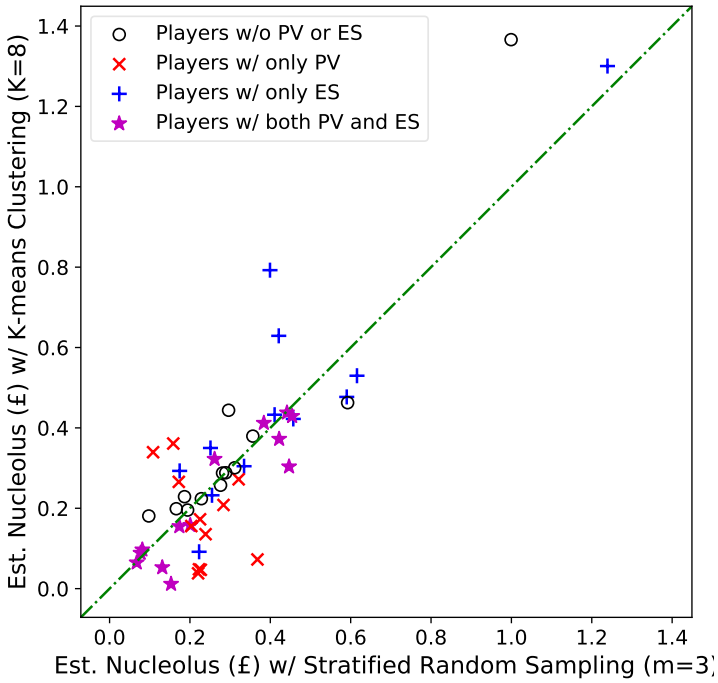


Figure 5.15: Comparison of nucleolus estimation methods for a 50-prosumer coalitional energy management scheme: simple K-means clustering ($K = 8$) vs. stratified random sampling ($m = 3$). The closer the data points are to the diagonal line, the more accurate the estimation is.

the energy export price is the UK feed-in tariff [28] fixed at £0.0379/kWh. For more detailed description of the model inputs, please refer to Appendix B.

5.4.1 Validation of Nucleolus Estimation

The computation time for the nucleolus estimation using the proposed full estimation method with clustering is both dependent on the number of prosumers, and the number of clusters K. Section 5.3.3.1 has demonstrated that the stratified random sampling method yields a good estimate of the nucleolus with a sample size of $m = 3$, and the method can be used to estimate the nucleolus of a 50-prosumer scheme within a reasonable amount of time (<10 hours). Therefore, setting $N = 50$, the results from stratified random sampling with $m = 3$ can be used as the benchmark to evaluate the estimation accuracy of the simple clustered player method (Section 5.3.1) and the proposed full estimation method (Section 5.3.2).

Figure 5.15 and Figure 5.16 both have the estimated nucleolus using stratified random

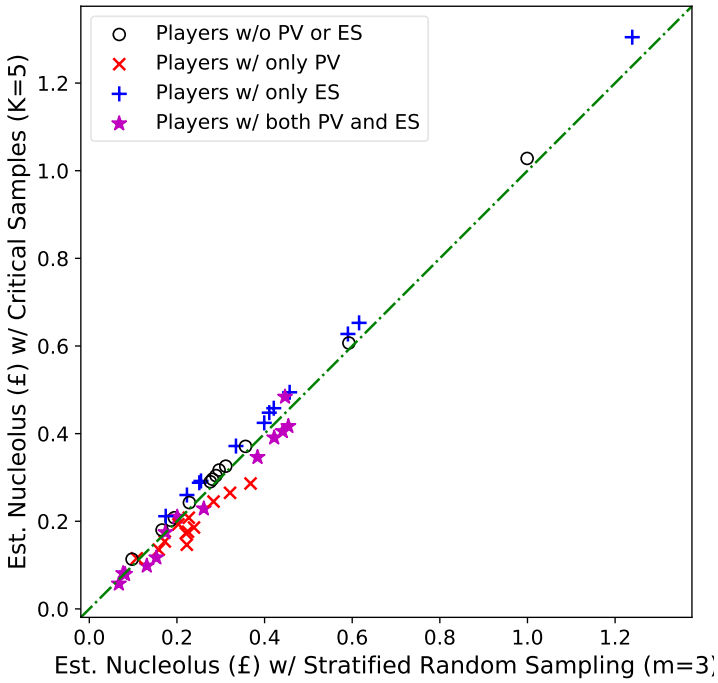


Figure 5.16: Comparison of nucleolus estimation methods for a 50-prosumer coalitional energy management scheme: full estimation method with K-means clustering ($K = 5$) vs. stratified random sampling ($m = 3$). The closer the data points are to the diagonal line, the more accurate the estimation is.

sampling with $m = 3$ on the x-axis, compared against the estimation results of the simple clustered player method (Section 5.3.1) and the proposed full estimation method (Section 5.3.2) respectively, both of which are using the K-means clustering technique. As shown in Section 5.3.3.1, the estimation results from stratified random sampling with $m = 3$ are likely to be close to the true nucleolus. Therefore, it is apparent that the estimation accuracy of the proposed full estimation method (Figure 5.16) is significantly higher than the simple clustered player method (Figure 5.15) despite the higher number of clusters ($K = 8$) for the latter.

Varying the number of clusters for the proposed full estimation method, Figure 5.17 shows a rapid decrease in the estimation error by increasing the K value from 3 to 5, by which point the estimation error continues to decrease as K increases, but at a lower rate. A higher estimation accuracy can be achieved by increasing the value of K, but according to (5.19), the sample size for $K = 8$ can be more than 8 times the sample

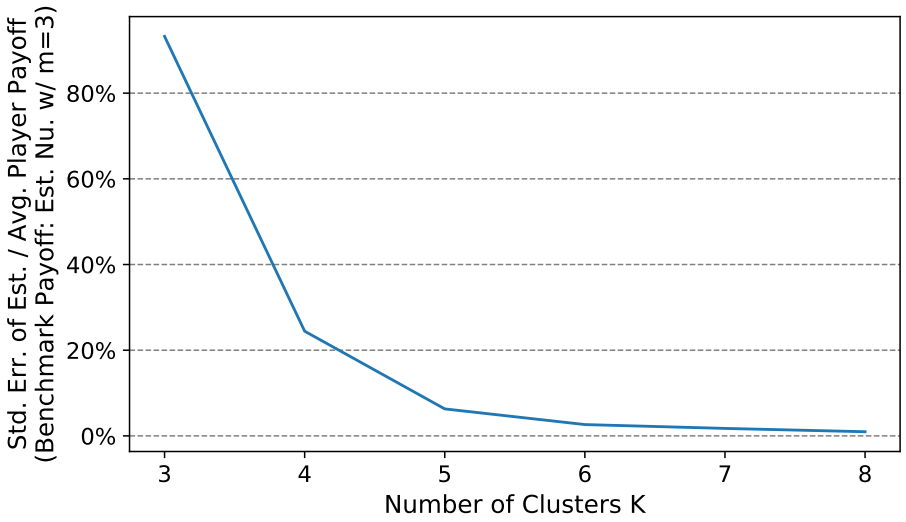


Figure 5.17: Estimation performance verification of the full estimation method with K-means clustering for a 50-prosumer coalitional energy management scheme using $m = 3$ as a benchmark: varying number of clusters ($K = 3, 4, 5, 6, 7$). Results are the average of 10 runs for each K value.

size for $K = 5$. Therefore, by varying K there is a trade-off between the estimation accuracy and the computation time.

5.4.2 Trade-Off between Estimation Accuracy and Computation Time

Taking K-means clustering as the benchmark clustering technique, using the grand coalition cooperative energy profile as the benchmark clustering features, the nucleolus computation time is plotted with an increasing number of prosumers in Figure 5.18. Compared to the full nucleolus computation, clustering is able to help scale up the coalitional energy management scheme to hundreds of prosumers while limiting the computation time within $10^4 s$, about 3 hours, which is reasonable for the day-ahead scheduling of prosumer operation. Depending on the chosen number of clusters, the upper limit of the size of the grand coalition varies. For example, For $K = 5, 6, 7, 8$, in order to keep the computation time within $10^4 s$, the proposed scheme can have as many as $400, 200, 100, 50$ prosumers. If the computation time is limited at around $10^3 s$ instead, the size of the scheme is limited to $200, 100, 50, 30$ prosumers.

Section 5.3.3.1 has demonstrated that the stratified random sampling method with a sample size $m = 3$ can be used as the benchmark for a scheme with up to 50 prosumers,

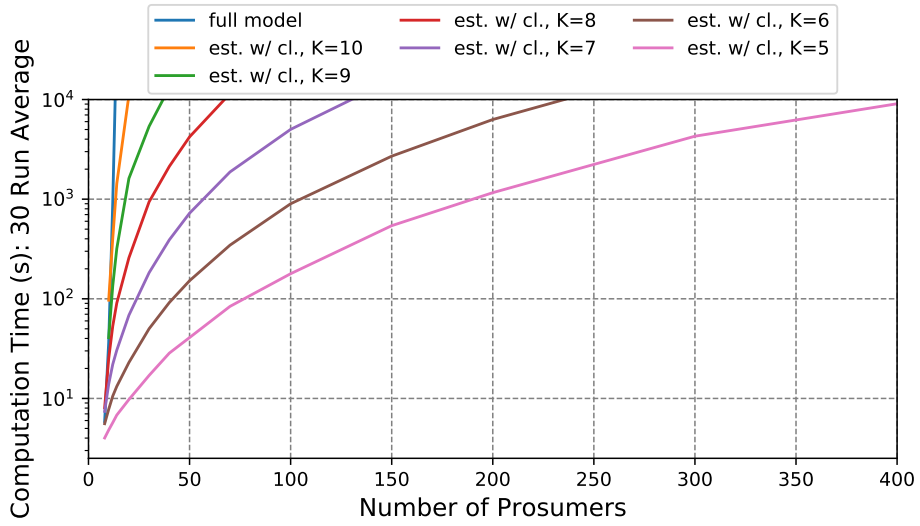


Figure 5.18: Comparison of the nucleolus estimation time for different number of clusters selected using K-means. The y-axis is on a log scale. The number of clusters K ranges from 5 to 10, and the numbers of prosumers N modeled to generate the plot are $10, 14, 20, 30, 50, 70, 100, 150, 200, 300, 400$.

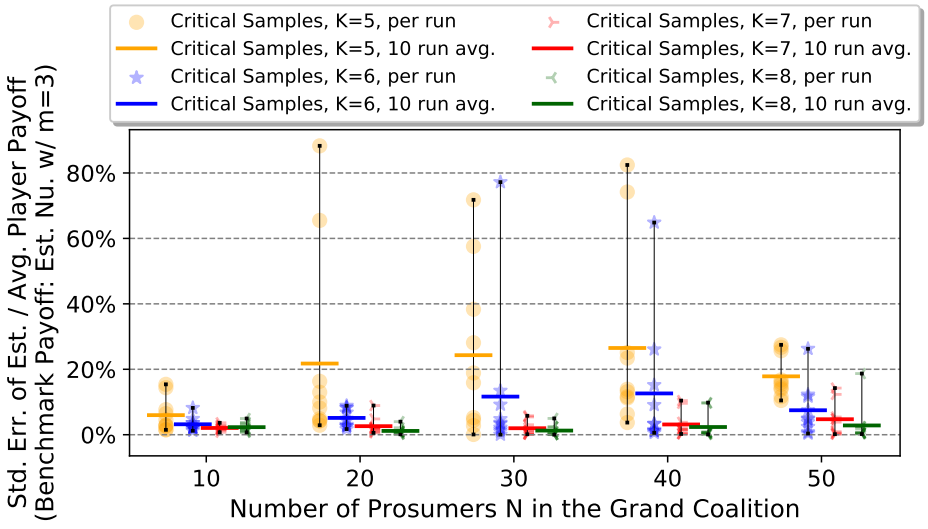


Figure 5.19: Nucleolus estimation performance of different techniques using the stratified random sampling method ($m = 3$ as baseline). Each data marker represents the estimation error of a model run, and the solid horizontal line represents the 10-run average.

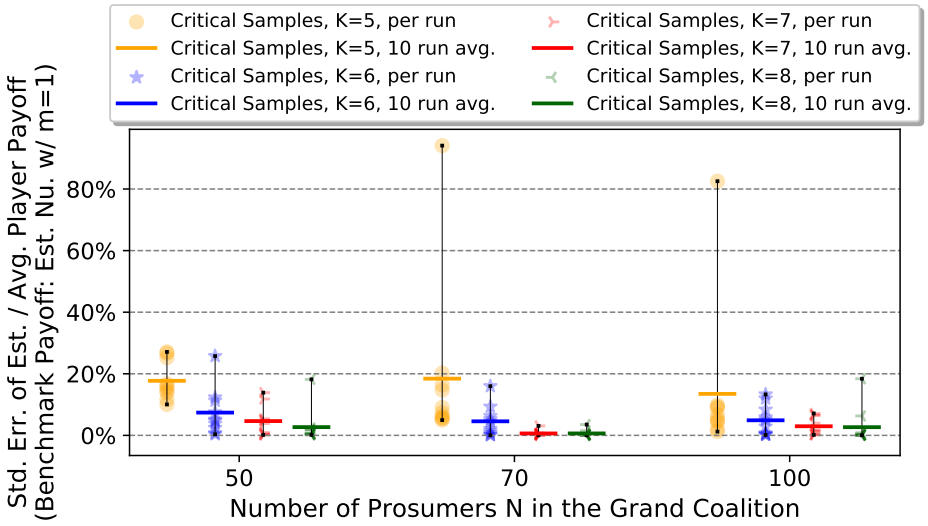


Figure 5.20: Nucleolus estimation performance of different techniques using the stratified random sampling method ($m = 1$ as baseline). Each data marker represents the estimation error of a model run, and the solid horizontal line represents the 10-run average.

and $m = 1$ for schemes with 50-100 prosumers. Figure 5.19 and Figure 5.20 compare the estimation performances of the proposed full estimation method with different K values. First, it can be seen that increasing K not only improves the average estimation accuracy, but also reduces the chance of under-sampling that produces highly inaccurate estimations, displayed in the figures as the outliers showing $> 50\%$ estimation errors. Second, it is noticed that for each K value, the estimation accuracy stays relatively consistent as N increases, which makes it promising to extend the application of the full estimation method beyond 100 prosumers without a benchmarking method.

Reviewing the results shown in Figures 5.18, 5.19, and 5.20, the tradeoff between the the computation time and the estimation accuracy is very clear. To balance between the two, a worst case standard error of the estimate is set to be 20% of the average prosumer payoff. In order to demonstrate the contrast of results between different clustering techniques, an aggressive computation time limit is set at $10^3 s$. If up to 100 prosumers is considered, based on Figure 5.18, the maximum K that guarantees the computation time being less than $10^3 s$ is selected. Therefore, for 14, 20, 30, 50, 100 prosumers, $K = 10, 9, 8, 7, 6$ respectively.

Table 5.2 shows the computation time for the clustering step only. Note that the

Table 5.2: Clustering Computation Time (s) Avg. over 10 Runs

No. prosumers (K)	**14 (10)**	**20 (9)**	**30 (8)**	**50 (7)**	**100 (6)**
K-means	5.3	4.0	3.5	4.1	5.1
Hierarchical-*ward*	0.6	0.5	0.5	0.6	0.5
GMM-*full*	2.2	1.8	1.7	2.1	1.7

computation time for K-means clustering and GMM is influenced by the number of runs repeated with different initial conditions. This number is set to be 1000 for K-means and 100 for GMM in this chapter. It can be seen that although different clustering techniques take slightly different times, the clustering computation time is negligible compared to the rest of the nucleolus estimation process. Therefore, Figure 5.18 is representative of all of the clustering techniques, and the K values selected based on the $10^3 s$ computation time criterion applies to them all as well.

5.4.3 Comparison of Clustering Techniques

In order to compare the performances of the different clustering techniques with different clustering features, it is no longer enough to use stratified random sampling as an estimation benchmark (Section 5.3.3.1) mainly for two reasons: 1) the performances of different clustering techniques can be very close, which are difficult to distinguish given the estimation error of the stratified random sampling method itself, and 2) it becomes computationally challenging to use the stratified sampling method for estimation for larger schemes. Therefore, in order to evaluate the estimation performance of the clustering techniques for more than 100 prosumers, the stratified random sampling method can only be used to verify the excesses of sampled coalitions (Section 5.3.3.2).

The estimation accuracy of different clustering techniques is evaluated by comparing the resulted percentages of coalitions that have *divisible excesses*. A divisible excess is a positive excess that, after being divided by the number of prosumers within the corresponding coalition, has a rounded value greater or equal to $£0.01$, the smallest monetary reward that can be materialized. The fewer coalitions with divisible excesses there are, the more likely it is for the estimated nucleolus to prevent defection. With prosumer profiles randomly selected from [182] in each run, the estimation evaluation

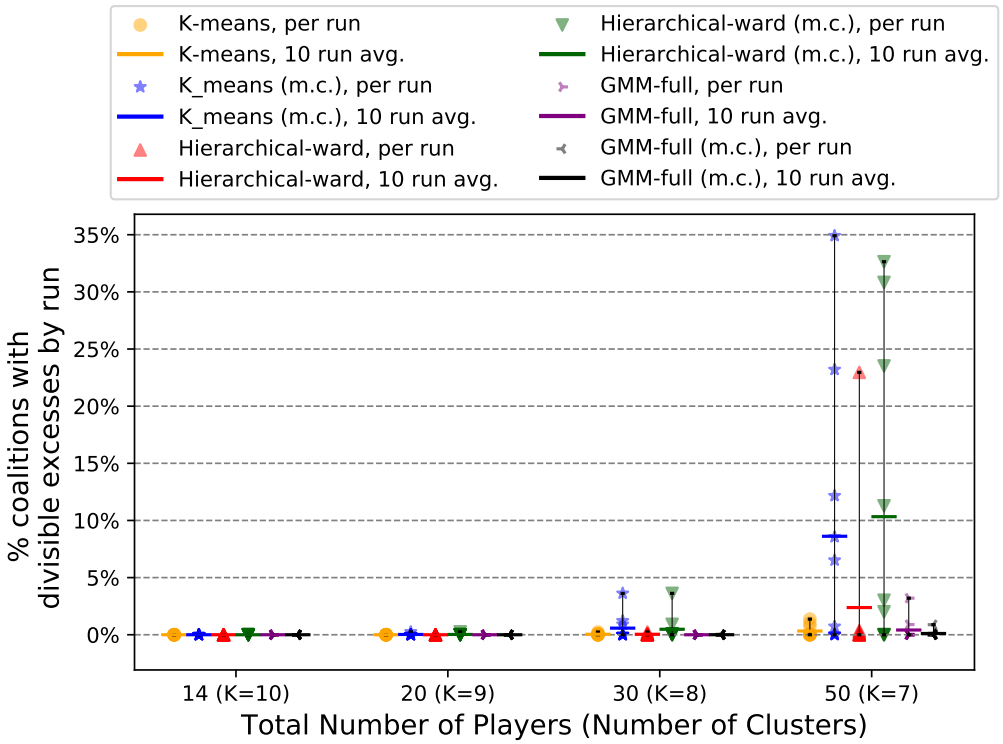


Figure 5.21: Nucleolus estimation performance with different clustering techniques validated using the stratified random sampling based approach ($m = 1$). The estimation performance is measured by the percentage of coalitions with divisible excesses, where a divisible excess is a positive excess that, after being divided by the number of prosumers within the corresponding coalition, has a rounded value greater or equal to $£0.01$. Coalitions that have positive divisible excesses can provide additional rewards for their member prosumers, and therefore have an incentive to defect from the grand coalition. Since the nucleolus guarantees that no coalitions have positive excesses, the estimation method that has the lowest number of coalitions with divisible excesses is considered to have the best estimation performance. Each data marker represents the estimation error of a model run, and the solid horizontal line represents the 10-run average.

results are plotted in Figure 5.21. It shows that GMM with the grand coalition marginal contribution profile as the clustering features has the lowest number of coalitions with divisible excesses, and therefore the best estimation performance. In fact, it is the only method that performs better than the estimation method of standard K-means with grand coalition cooperative energy profile as the clustering features. Even though there is a small number of coalitions (1%) that end up with positive excesses, giving them some incentive to defect on this particular day, it is unlikely that the same coalitions would consistently have a positive excess over time. An interesting area for future work is to analyze the level of incentive to defect based on the size of the coalition, the time duration

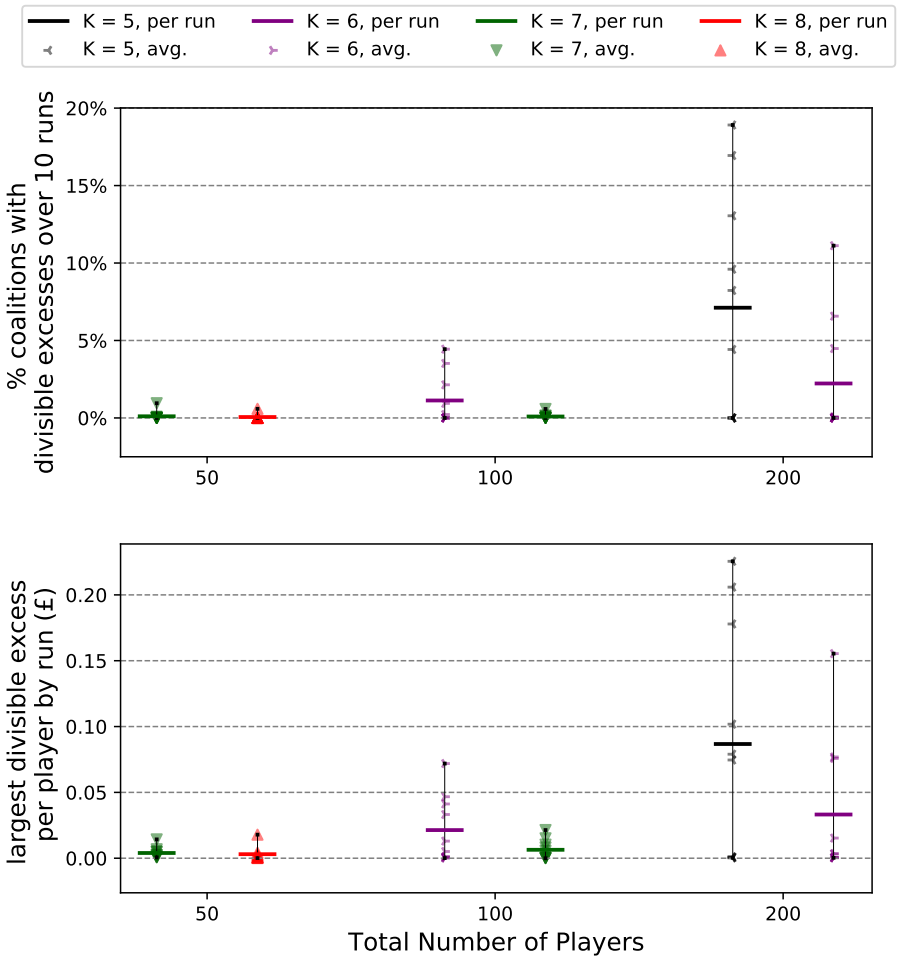


Figure 5.22: Nucleolus estimation performance with GMM-full with marginal contribution feature using different K values, validated using the stratified random sampling based approach ($m = 1$). The upper plot shows the percentage of coalitions with divisible excesses. The lower plot shows the largest divisible excess per prosumer, which is obtained by dividing all the divisible excesses by the number of prosumers in their corresponding coalitions and selecting the highest value. For both plots, the lower the value the better the estimation. Each data marker represents the estimation error of a model run, and the solid horizontal line represents the 10-run average.

of the presence of a positive excess, and the magnitude of the excess.

5.4.4 Scaling up with Relaxed Computation Time

Because the $10^3 s$ threshold is arbitrarily set, there is still room to increase the estimation accuracy by increasing the total number of clusters K. For example, if the time limit is set to $10^4 s$, $K = 7$ can be used for a 100-prosumer scheme, and $K = 5$ can be applied to up to 400 prosumers.

Since GMM with the grand coalition marginal contribution profile as the clustering feature is identified as the best nucleolus estimation method in Section 5.4.3, it is applied to compare the model performances with varying numbers of clusters. Two metrics are used: 1) percentage of coalitions that have divisible excesses, and 2) the largest divisible excess per prosumer, which is obtained by dividing all the divisible excesses by the number of prosumers in their corresponding coalitions and selecting the highest value. Both metrics are indications of the level of incentive for defection, which are plotted in Figure 5.22 for $K = 50, 100, 200$. It demonstrates that relaxing the allowed computation time can significantly improve the nucleolus estimation accuracy. For the 100-prosumer case, by having 7 clusters instead of 6, the number of coalitions with divisible excesses is reduced from a 10-run average of 1.1% to below 0.1%, and the maximum divisible excess per prosumer is reduced from $£0.07$ to $£0.02$, a very low incentive for defection. The investigation of what level of incentive can lead to defection in practice is out of the scope of this chapter.

5.5 Concluding remarks

In this chapter, a detailed description of the nucleolus computation is presented, from which it shows that the number of optimization problems that are required to solve for the nucleolus increases exponentially with the number of participants. To overcome this computational challenge, clustering is identified as a promising strategy.

After exploring different clustering techniques, this chapter first introduces the concept of clustered players, for which the nucleolus can be easily computed. However, the in-cluster allocation of the payoffs can induce significant estimation errors. To improve the method, a clustering based full estimation method is proposed. The computational complexity for an N-prosumer scheme is reduced from $\sim \mathcal{O}(2^N)$ to $\sim \mathcal{O}(2^K N)$ where K is the number of clusters that can be set between 5 and 10. With 15 prosumers as the maximum number of participants in a coalitional energy management scheme in existing studies, it is shown that the proposed method can scale it up to over 100 prosumers.

Varying the number of clusters, the trade-off between the computation time and the estimation accuracy is analyzed. To measure the nucleolus estimation accuracy, a stratified

sampling based technique is formulated. Among different clustering techniques and clustering features, the Gaussian mixture model coupled with the marginal contribution profiles is shown to have the highest estimation accuracy. For a 100-prosumer scheme, by grouping the prosumers into 7 clusters, it takes less than 3 hours to compute. It yields a payoff allocation that leaves an average of only 0.1% of the coalitions that will have divisive excesses, the highest of which is only $\pounds 0.02$ per person, amounting to a very low incentive for any prosumer to defect from the coalitional energy management scheme.

Although the world is full of suffering, it is also full of the overcoming of it.

— Helen Keller

6 Conclusions

Contents

6.1 Concluding Discussion 133
6.2 Future Work 139

6.1 Concluding Discussion

This thesis addresses the following research question:

As the adoption of DER continues to increase, how do we incentivize prosumers to manage these resources in an economically efficient and sustainable manner?

In order to answer this question, this thesis first proposes a cooperative game theoretic framework to construct the prosumer coalitional energy management scheme, and then tackles its scalability issues by developing novel computational strategies to bring this scheme up to scale.

Cooperative Game Theory Based Prosumer Energy Management

Due to the increasing amount of renewable generation and flexible resources in the distributed energy network, there has been a surge of research interest in recent years

in prosumer-centric energy markets. Facilitated by advanced ICT, these markets would allow prosumers to trade or share energy with each other.

In the review of the state-of-the-art literature on prosumer-centric energy markets, three major gaps are identified. First, the assumption that prosumers are purely social welfare driven is unrealistic, and the market schemes with objectives that are only focused on maximizing community benefits is unlikely to incentivize prosumers to participate. Second, in a competitive market environment, the potential risk of exercising market power to manipulate the trading outcome is often overlooked. Additionally, financial benefits are not guaranteed for every participant, introducing instability into the market. Third, even though cooperative game theoretic schemes are free of the first two caveats, they tend to be applied to simple tradings of excess generation, and therefore are unable to capture the benefit of flexible resources such as energy storage systems.

To address these research gaps, this thesis proposes a multi-prosumer multi-timestep coalitional energy management scheme, which 1) provides individual financial incentives to prosumers to encourage participation, 2) uses a cooperative game theory framework to ensure profits for all the participants, and 3) enables collective ES arbitrage and rewards prosumers who participate with ES units accordingly.

One of the fundamental motivations for establishing prosumer-centric local energy markets is that the price of importing energy from the grid is higher than the price of exporting energy to the grid. With this basic assumption, the proposed coalitional energy management scheme is constructed as follows (see Chapter 3).

1. An algorithm to jointly optimize the participants' DER to maximize self-consumption within prosumer coalitions is formulated.

2. For each prosumer coalition, the maximum achievable energy cost savings are calculated and defined as the coalition value. The value of the grand coalition is therefore equal to the total profit of the proposed scheme.

3. The grand coalition profit is divided among the participating prosumers using a chosen payoff allocation algorithm for cooperative game theory.

To further investigate the fairness and stability of the proposed scheme, this thesis conducts detailed analyses on the properties and computations of two popular payoff allocations in cooperative game theory: the Shapley value and the nucleolus.

The Shapley value is the most widely used payoff allocation in cooperative games. It is known for its *fairness*, because by definition a player's Shapley value is the mean of this player's marginal contribution to all the coalitions within the game. In the proposed coalitional energy management scheme, the Shapley value can be calculated analytically as long as all the coalition values are known.

The nucleolus is computed by lexicographically minimizing all the coalitions' excesses. The excess of a prosumer coalition in the proposed scheme is defined as the difference between the sum of payoffs for all the prosumers within the coalition and the coalition's maximum achievable energy cost savings. It is a metric to evaluate a prosumer coalition's dissatisfaction in the grand coalition. When the excess is above zero, it means the coalition is financially better off operating on its own.

Based on mathematical proofs in Section 3.3.2 and case studies in Section 3.5 using UK-based data inputs from Appendix B, the following list summarizes the main findings related to the proposed coalitional energy management scheme:

1. Collectively scheduling DER operation under the proposed scheme helps reduce the peak load and flattens the energy profile in the local distribution network. This is empirically demonstrated in a case study in Figure 3.1. It is worth noting that this is not a proof that cooperation always leads to less stress on the energy network.

2. The proposed scheme is a balanced cooperative game (see proof in Section 3.3.2.1), which means that there exist stabilizing payoff allocations to ensure that no prosumer coalitions can financially benefit from defecting from the grand coalition. This proof also guarantees that the nucleolus is a stabilizing payoff allocation.

3. Both the Shapley value and the nucleolus provide financial benefits for participating prosumers. In a counter example, Figure 3.4 demonstrates that the Shapley value is not guaranteed a stabilizing payoff allocation. Meanwhile, it provides empirical evidence to show that the nucleolus is a stabilizing payoff allocation.

4. Stochastic errors in predicting prosumer consumption and PV generation have a negative impact on the total profit of the scheme. As shown in Figure 3.8, 5% and 15% standard prediction errors for PV generation and prosumer consumption predictions respectively can lead to an average of around 2% total profit reduction, and 10% and 30% standard prediction errors for PV generation and prosumer consumption predictions respectively can lead to an average of around 4% total profit reduction.

A key limitation of the proposed coalitional energy management scheme is related to the computation of payoff allocations. Both the Shapley value and the nucleolus have a computational complexity of $\sim \mathcal{O}(2^N)$ where N is the number of prosumer participants. This leads to the second theme of the thesis:

Scalability of the Coalitional Energy Management Scheme

The two main payoff allocations of the proposed coalitional energy management scheme, the Shapley value and the nucleolus, are both computationally intractable as the number of prosumer participants increases. Limiting the computation time to 10 hours for a model with a 24-hour timespan, the maximum size of the grand coalition is 20 prosumers for using the Shapley value, and 14 prosumers for using the nucleolus. In order to bring this scheme to a real-world scale, this thesis proposes strategies to tackle these computational issues.

As the definition of the Shapley value is a participant's marginal contributions to all the coalitions and these contributions all have equal weights, random sampling is identified as a promising strategy for producing an estimate of the Shapley value. Sampling methods have indeed been investigated in literature in approximating the Shapley value, but it had never been implemented in an energy market context. This thesis proposes a novel multi-step sampling strategy to estimate the Shapley value of the coalitional energy management scheme (see Chapter 4). The following is a summary of the main contributions and findings related to this proposed strategy.

1. A coalitional stratified random sampling method is adapted to select a limited number of prosumer coalitions ($\sim \mathcal{O}(N^2)$) from all possible coalitions ($\sim \mathcal{O}(2^N)$). The values of the sampled coalitions are used to estimate the Shapley value by calculating the average marginal contribution of each prosumer to the sampled coalitions (see Algorithm 1). This method's main improvement over the basic random sampling method is its ability to select coalition samples of all sizes for each prosumer, a more accurate representation of the types of coalitions that a prosumer contributes to in a full Shapley value calculation. The proposed method is shown to reduce the computation time of a 20-prosumer scheme from $\sim 2 \times 10^4 s$ to $\sim 10^3 s$, and it is able to estimate the Shapley value of a 100-prosumer scheme in $\sim 2 \times 10^4 s$ (see Table 4.1). Its estimation accuracy is consistently higher than that of the basic random sampling method (see Figure 4.3).

2. A multi-step sampling strategy is proposed to divide the sampling process into incremental stages. During each stage the estimation results are verified and the sampling process is terminated when the estimation criteria are met (see Algorithm 2). This strategy further reduces the computation time of a 20-prosumer scheme to $\sim 3 \times 10^2 s$, and it can estimate the Shapley value for a 100-prosumer model in $\sim 3 \times 10^3 s$ and for a 200-prosumer model in $\sim 2 \times 10^4 s$ (see Table 4.1). As a benchmark, it is shown to be able to estimate the Shapley value for a scheme of 14 prosumers with a standard estimate error equal to approximately 2% of the average prosumer payoff (see Figure 4.3). To reach the same estimation accuracy, the basic random sampling method requires about 60% more time.

Compared to the Shapley value, the computation of the nucleolus is significantly more complicated because it requires an iterative process of minimizing coalitions' excesses. Different coalitions' values no longer have equal weights in the nucleolus, so sampling becomes a less effective method in estimating the nucleolus. It can, however, still be used as a benchmark to evaluate the performances of any estimation methods. It is then noticed that the nucleolus computation's iterative process effectively ranks coalitions from the most to the least likely to be dissatisfied. Using this as a basis, clustering,

an unsupervised machine learning technique, is identified as a promising strategy to reduce the computational complexity of the nucleolus. The rationale behind this is that clustering is known to group similar characteristics and separate dissimilar ones, which helps identify coalitions of prosumers that compensate each other well to create high values, and are thus likely to be dissatisfied if merged in a bigger coalition. Therefore, this thesis proposes a clustering based method to estimate the nucleolus of the coalitional energy management scheme (see Chapter 5). The following is a summary of the main contributions and findings related to this proposed method.

1. Using basic K-means clustering, prosumers are grouped into K clustered players, for whom the nucleolus can be calculated (see Section 5.3.1). The number of coalitions is reduced from 2^N to 2^K but the payoff allocation within each clustered player remains a challenge.

2. A full nucleolus estimation method is proposed, which minimizes the excesses of a selected number of coalitions, identified using clustering techniques as coalitions that are most likely to defect (see Section 5.3.2). Given K clusters ($K \leq 10$) and N prosumers, the number of coalitions selected with clustering no larger than $2^{K-1}(N+2)-1$, a significant reduction from 2^N when $N \geq 20$. As shown in Figure 5.18, for $K = 5, 6, 7, 8$, in order to keep the computation time within $10^4 s$, the proposed scheme can have as many as $400, 200, 100, 50$ prosumers. If the computation time is limited at around $10^3 s$ instead, the size of the scheme is limited to $200, 100, 50, 30$ prosumers.

3. A stratified sampling based technique is formulated to measure the nucleolus estimation accuracy (see Section 5.3.3.2). This technique selects coalition samples and calculates their excesses based on the estimated nucleolus as a metric to evaluate the performance of the nucleolus estimation method.

4. Different clustering techniques with different clustering features are tested (see Section 5.2), and the Gaussian mixture model as the clustering technique with the marginal contribution profile as the clustering feature is shown to have the highest

estimation accuracy. For a 100-prosumer scheme, it can produce a nucleolus estimate in less than 3 hours that guarantees that only 0.1% of the coalitions can operate independently to yield additional financial payoffs for its participants, and that the highest benefit a coalition can yield from defection is only £0.02 per person (see Figure 5.22).

With the proposed sampling and clustering methods, both the Shapley value and the nucleolus can be estimated for coalitional energy management schemes with 100 prosumers under $10^4 s$. There is a tradeoff between the computation time and the estimation accuracy. Depending on the size of the scheme, and the desired estimation accuracy, the number of samples or number of clusters can be adjusted accordingly to meet the set criteria.

To conclude, this thesis proposes a scalable prosumer energy management scheme based on cooperative game theory. With a rigorous mathematical foundation, it shows that the proposed scheme minimizes the joint energy cost of the grand coalition and guarantees profit for every participant. Depending on the social preference of whether a fair or stabilizing payoff allocation is desired, the Shapley value or the nucleolus can be highly accurately estimated for a coalitional energy management scheme of up to 100 prosumers in under $10^4 s$.

6.2 Future Work

Some interesting areas for future work include:

1. the technical and economic impact on the network. One of the most interesting areas for future work is to include the network operator into the 'game' so that the economic impact of prosumer cooperation on the network can be quantified and included in the value calculation. To achieve this, network constraints and line losses need to be built into the model, and the rate structure may need to become dynamic. Whether it is possible to include the line losses and network constraints into the scheme without undermining the convexity of the model is to

be determined. The usefulness of the convexity of the model is multifold: 1) it guarantees the optimal solution; 2) it limits the computation effort; 3) it ensures that the proofs in this thesis are valid. If the convexity of the model is indeed challenged, additional research effort is required to reconfigure the problem setup so that the cooperative game can be solved within a reasonable time while retaining the most important properties, such as fairness and stability, examined in this thesis.

2. battery degradation. Due to the short time frame (24 hours) used in the proposed scheme, and the fact that the battery also needs to operate in the baseline scenario, it was considered that the battery degradation cost would not have a huge impact on the proposed scheme. This needs to be further verified. If the additional battery degradation cost incurred by the proposed scheme is indeed significant, a linear degradation model can be incorporated into the model without compromising the model convexity.

3. impact of electricity pricing, ES model inputs, and DER penetrations on the final prosumer payoffs. As the market price for energy continues to evolve, ES technologies continue to mature, and DER penetrations continue to increase, it would be interesting to analyze what the scheme will look like under different future scenarios.

4. prosumer behavior and preference. It will be interesting to understand what level of estimation error for the Shapley value and nucleolus is socially acceptable. In other words, how much deviation from the Shapley value will spark questions on the fairness, and how much deviation from the nucleolus will push prosumers to defect from the grand coalition? Additionally, it is important to see if the financial benefits offered by the proposed scheme will inspire any behavioral change in prosumers.

The first extension requires the most extensive technical exploration as it may require a restructuring of the cooperative game and bring in nonconvexity into the computation, but it is also the most important piece that can push the proposed scheme closer to implementation. The second extension is a comparatively simpler task, but also an

important area of research in order to incorporate realistic battery models into the model. The third extension is an empirical exercise that can provide practical economic insights, but the model accuracy depends on how well the first two extensions are executed. The last extension is an interdisciplinary subject that requires expertise from behavioral science, economics, and engineering.

Appendices

We proved that we are still a people capable of doing big things and tackling our biggest challenges.

— Barack Obama

A Additional Proofs

Contents

A.1 Proof of Lemma 3.1 143
A.2 Proof of Lemma 3.2 144
A.3 Proof of Lemma 3.3 145
A.3.1 Proof of Positivity Homogeneity 145
A.3.2 Proof of Subadditivity 146
A.4 Proof of Theorem 3.4 147

A.1 Proof of Lemma 3.1

When $D(\mathfrak{B}^{\langle\mathcal{S}\rangle}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}}) = 1$, based on its definition, (3.27), (3.28) and (3.29) are satisfied. Noticing the linearity of these constraints, they can be scaled by a factor of $\lambda > 0$:

1) *Power constraint:*

$$\lambda\underline{\mathbf{B}}^{\mathcal{S}} \leq \lambda\mathfrak{B}^{\langle\mathcal{S}\rangle} \leq \lambda\overline{\mathbf{B}}^{\mathcal{S}}$$

2) *Energy constraint:*

$$\begin{aligned}\lambda\mathbf{E}^{\mathcal{S}}\underline{\mathbf{SoC}} \leq & \lambda\mathbf{E}^{\mathcal{S}}\mathbf{SoC}^{\mathbf{0}} + \mathbf{\Lambda}^{R}([\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}]_{+}\eta^{\mathbf{I}} + [\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}]_{-}\eta^{\mathbf{O}}) \\ \leq & \lambda\mathbf{E}^{\mathcal{S}}\overline{\mathbf{SoC}}\end{aligned}$$

3) *Cycle constraint:*

$$(\mathbf{1}^R)^\intercal([\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}]_+\eta^{\mathbf{I}} + [\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}]_-\eta^{\mathbf{O}}) = \mathbf{0}$$

Therefore, $D(\lambda\mathfrak{B}^{\langle\mathcal{S}\rangle}, \lambda\mathbf{E}^{\mathcal{S}}, \lambda\underline{\mathbf{B}}^{\mathcal{S}}, \lambda\overline{\mathbf{B}}^{\mathcal{S}}) = 1$

A.2 Proof of Lemma 3.2

For $\mathcal{S} \cap \mathcal{T} = \emptyset$, there is

$$\begin{aligned}\mathfrak{O}_{\mathcal{S}+\mathcal{T}} &= \{\mathfrak{B} \in \mathbb{R}^{R\times N} \mid D(\mathfrak{B}, \mathbf{E}^{\mathcal{S}+\mathcal{T}}, \underline{\mathbf{B}}^{\mathcal{S}+\mathcal{T}}, \overline{\mathbf{B}}^{\mathcal{S}+\mathcal{T}}) = 1\} \\ &\overset{(3.21)}{=} \{\mathfrak{B} \in \mathbb{R}^{R\times N} \mid D(\mathfrak{B}, \mathbf{E}^{\mathcal{S}\cup\mathcal{T}}, \underline{\mathbf{B}}^{\mathcal{S}\cup\mathcal{T}}, \overline{\mathbf{B}}^{\mathcal{S}\cup\mathcal{T}}) = 1\} \\ &= \mathfrak{O}_{\mathcal{S}\cup\mathcal{T}}\end{aligned} \tag{A.1}$$

$\mathcal{S}$ can be considered the union of all its players, and because the intersection of any two different players is $\emptyset$, there is

$$\begin{aligned}\mathbf{Q}^{\mathcal{S}} = \mathbf{Q}^{\bigcup_{i\in\mathcal{S}}\{i\}} &\overset{(3.21)}{=} \mathbf{Q}^{\sum_{i\in\mathcal{S}}\{i\}} = \mathbf{Q}^{\sum_{i\in\mathcal{N}} 1_i^{\mathcal{S}}\{i\}} \\ \mathfrak{O}_{\mathcal{S}} = \mathfrak{O}_{\bigcup_{i\in\mathcal{S}}\{i\}} &\overset{(A.1)}{=} \mathfrak{O}_{\sum_{i\in\mathcal{S}}\{i\}} = \mathfrak{O}_{\sum_{i\in\mathcal{N}} 1_i^{\mathcal{S}}\{i\}}\end{aligned} \tag{A.2}$$

where $\mathbf{Z} \in \mathbb{R}^{R\times N}$. Therefore,

$$\begin{aligned}\theta[\sum_{h=1}^{2^N}\lambda_h\mathcal{S}_h] &= \min_{\mathfrak{B}\in\mathfrak{O}_{\sum_{h=1}^{2^N}\lambda_h\mathcal{S}_h}} F(\mathfrak{B}, \mathbf{Q}^{\sum_{h=1}^{2^N}\lambda_h\mathcal{S}_h}) \\ &\overset{(3.20)(A.2)}{=} \min_{\mathfrak{B}\in\mathfrak{O}_{\sum_{h=1}^{2^N}\lambda_h\sum_{i\in\mathcal{N}} 1_i^{\mathcal{S}_h}\{i\}}} F(\mathfrak{B}, \mathbf{Q}^{\sum_{h=1}^{2^N}\lambda_h\sum_{i\in\mathcal{N}} 1_i^{\mathcal{S}_h}\{i\}}) \\ &= \theta[\sum_{h=1}^{2^N}\lambda_h\sum_{i\in\mathcal{N}} 1_i^{\mathcal{S}_h}\{i\}] = \theta[\sum_{i\in\mathcal{N}}\sum_{h=1}^{2^N}\lambda_h 1_i^{\mathcal{S}_h}\{i\}]\end{aligned}$$

A.3 Proof of Lemma 3.3

A.3.1 Proof of Positivity Homogeneity

When $\lambda = 0$,

$$\begin{aligned}
&\theta[\lambda\mathcal{S}] = \theta[0\mathcal{S}] \\
\overset{(3.30)}{=}& \min_{\mathfrak{B}\in\mathbb{R}^{R\times N}} F(\mathfrak{B}, \mathbf{Q}^{0\mathcal{S}}) \ \ s.t.\ D(\mathfrak{B}, \mathbf{E}^{0\mathcal{S}}, \underline{\mathbf{B}}^{0\mathcal{S}}, \overline{\mathbf{B}}^{0\mathcal{S}}) = 1 \\
\overset{(3.17)}{=}& \min_{\mathfrak{B}\in\mathbb{R}^{R\times N}} F(\mathfrak{B}, 0\mathbf{Q}^{\mathcal{S}}) \ \ s.t.\ D(\mathfrak{B}, 0\mathbf{E}^{\mathcal{S}}, 0\underline{\mathbf{B}}^{\mathcal{S}}, 0\overline{\mathbf{B}}^{\mathcal{S}}) = 1 \\
=& \ \ F(\mathbf{0}, \mathbf{0}) = 0 = 0\theta[\mathcal{S}] = \lambda\theta[\mathcal{S}]
\end{aligned}$$

When $\lambda > 0$, there is

$$C_{\lambda\mathcal{S}}(\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}) \leq C_{\lambda\mathcal{S}}(\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle}),\ \forall\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle} \in \mathfrak{O}_{\lambda\mathcal{S}}$$

Because $\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *} \in \mathfrak{O}_{\lambda\mathcal{S}}$, there is

$$D(\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}, \mathbf{E}^{\lambda\mathcal{S}}, \underline{\mathbf{B}}^{\lambda\mathcal{S}}, \overline{\mathbf{B}}^{\lambda\mathcal{S}}) = 1 \tag{A.3}$$

Define $\mathfrak{B}^{\langle\mathcal{S}\rangle'} = \frac{1}{\lambda}\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}$. Using Lemma 3.1, there is

$$D(\mathfrak{B}^{\langle\mathcal{S}\rangle'}, \frac{1}{\lambda}\mathbf{E}^{\lambda\mathcal{S}}, \frac{1}{\lambda}\underline{\mathbf{B}}^{\lambda\mathcal{S}}, \frac{1}{\lambda}\overline{\mathbf{B}}^{\lambda\mathcal{S}}) = 1$$

Using (3.17), there is

$$D(\mathfrak{B}^{\langle\mathcal{S}\rangle'}, \mathbf{E}^{\mathcal{S}}, \underline{\mathbf{B}}^{\mathcal{S}}, \overline{\mathbf{B}}^{\mathcal{S}}) = 1$$

In other words, $\mathfrak{B}^{\langle\mathcal{S}\rangle'} \in \mathfrak{O}_{\mathcal{S}}$. Therefore,

$$\begin{aligned}
\theta[\mathcal{S}] &\overset{(3.34)}{=} C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle *}) \\
&\overset{(3.35)}{\leq} C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle'}) \\
&\overset{(3.25)}{=} F(\frac{1}{\lambda}\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}, \mathbf{Q}^{\mathcal{S}}) \\
&= \ F(\frac{1}{\lambda}\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}, \frac{1}{\lambda}\mathbf{Q}^{\lambda\mathcal{S}}) \\
&\overset{(3.24)}{=} \frac{1}{\lambda}F(\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}, \mathbf{Q}^{\lambda\mathcal{S}}) \\
&\overset{(3.25)}{=} \frac{1}{\lambda}C_{\lambda\mathcal{S}}(\mathfrak{B}^{\langle\lambda\mathcal{S}\rangle *}) \\
&\overset{(3.34)}{=} \frac{1}{\lambda}\theta[\lambda\mathcal{S}]
\end{aligned}$$

After rearrangement, there is

$$\theta[\lambda\mathcal{S}] \geq \lambda\theta[\mathcal{S}], \forall\lambda > 0$$

Since λ can be any positive value, the following can be constructed

$$\theta[\mathcal{S}] = \theta[\frac{1}{\lambda}(\lambda\mathcal{S})]$$

Since $\frac{1}{\lambda} > 0$, there is

$$\theta[\mathcal{S}] = \theta[\frac{1}{\lambda}(\lambda\mathcal{S})] \geq \frac{1}{\lambda}\theta[\lambda\mathcal{S}]$$

Combining the equations above, there is

$$\theta[\lambda\mathcal{S}] \geq \lambda\theta[\mathcal{S}] \geq \lambda\frac{1}{\lambda}\theta[\lambda\mathcal{S}] = \theta[\lambda\mathcal{S}]$$

Therefore, it can only be the case that

$$\theta[\lambda\mathcal{S}] = \lambda\theta[\mathcal{S}]$$

A.3.2 Proof of Subadditivity

Rewriting (3.22) with variable $\mathbf{L} = \mathfrak{B} + \mathbf{Q}, \forall\mathfrak{B}, \mathbf{Q} \in \mathbb{R}^{R\times N}$:

$$F(\mathbf{L}) = \mathbf{r}^{im}[\mathbf{L}1^N]_+ + \mathbf{r}^{ex}[\mathbf{L}1^N]_-$$

For any $\mathbf{L}_1, \mathbf{L}_2 \in \mathbb{R}^{R\times N}$, $\mathbf{r}^{im} \geq \mathbf{r}^{ex}$, and $\lambda \geq 0$, there is

1. *(Positive Homogeneity)* $F(\lambda\mathbf{L}_1) = \lambda F(\mathbf{L}_1)$
2. *(Convexity)* $\frac{F(\mathbf{L}_1)+F(\mathbf{L}_2)}{2} \geq F(\frac{\mathbf{L}_1+\mathbf{L}_2}{2})$

Therefore,

$$\begin{aligned} & F(\mathbf{L}_1) + F(\mathbf{L}_2) \\ \geq\ & 2F(\frac{\mathbf{L}_1 + \mathbf{L}_2}{2}), \text{ by } \textit{convexity} \text{ of } F \\ =\ & F(\mathbf{L}_1 + \mathbf{L}_1), \text{ by } \textit{positive homogeneity} \text{ of } F \end{aligned} \tag{A.4}$$

Denoting $\mathbf{L}^{\langle\lambda\mathcal{S}\rangle*} = \mathfrak{B}^{\langle\lambda\mathcal{S}\rangle*} + \mathbf{Q}^{\lambda\mathcal{S}}, \forall\mathcal{S} \in 2^{\mathcal{N}}$, there is

$$\begin{aligned}\theta[\lambda_1\mathcal{S}] + \theta[\lambda_2\mathcal{T}] &= F(\mathbf{L}^{\langle\lambda_1\mathcal{S}\rangle*}) + F(\mathbf{L}^{\langle\lambda_2\mathcal{T}\rangle*}) \\ &\overset{(A.4)}{\geq} F(\mathbf{L}^{\langle\lambda_1\mathcal{S}\rangle*} + \mathbf{L}^{\langle\lambda_2\mathcal{T}\rangle*}) \\ &= F(\mathfrak{B}^{\langle\lambda_1\mathcal{S}\rangle*} + \mathfrak{B}^{\langle\lambda_2\mathcal{T}\rangle*}, \mathbf{Q}^{\lambda_1\mathcal{S}} + \mathbf{Q}^{\lambda_2\mathcal{T}}) \\ &\overset{(3.20)}{=} C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}\rangle*} + \mathfrak{B}^{\langle\lambda_2\mathcal{T}\rangle*}) \end{aligned} \tag{A.5}$$

Based on (A.3), $\begin{cases} D(\mathfrak{B}^{\langle\lambda_1\mathcal{S}\rangle*}, \mathbf{E}^{\lambda_1\mathcal{S}}, \underline{\mathbf{B}}^{\lambda_1\mathcal{S}}, \overline{\mathbf{B}}^{\lambda_1\mathcal{S}}) = 1 \\ D(\mathfrak{B}^{\langle\lambda_2\mathcal{T}\rangle*}, \mathbf{E}^{\lambda_2\mathcal{T}}, \underline{\mathbf{B}}^{\lambda_2\mathcal{T}}, \overline{\mathbf{B}}^{\lambda_2\mathcal{T}}) = 1 \end{cases}$

Because $\eta^{\mathbf{I}}, \eta^{\mathbf{O}} \leq \mathbf{1}$, by combining the ES *energy capacities*, *charge limits* and *discharge limits* using (3.20), $\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle}$ can be constructed such that

$$\begin{aligned} C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}\rangle*} + \mathfrak{B}^{\langle\lambda_2\mathcal{T}\rangle*}) \geq C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle}), \\ s.t.\ D(\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle}, \mathbf{E}^{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}, \underline{\mathbf{B}}^{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}, \overline{\mathbf{B}}^{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}) = 1 \end{aligned} \tag{A.6}$$

Based on (3.35), there is

$$C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle}) \geq C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle*}) \tag{A.7}$$

Combining (A.5), (A.6) and (A.7), it is proven that

$$\theta[\lambda_1\mathcal{S}] + \theta[\lambda_2\mathcal{T}] \geq C_{\lambda_1\mathcal{S}+\lambda_2\mathcal{T}}(\mathfrak{B}^{\langle\lambda_1\mathcal{S}+\lambda_2\mathcal{T}\rangle*}) = \theta[\lambda_1\mathcal{S} + \lambda_2\mathcal{T}]$$

A.4 Proof of Theorem 3.4

Using (3.26), the left-hand side of (3.8) can be rewritten as

$$\begin{aligned} &\sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) v(\mathcal{S}) \qquad (A.8) \\ &= \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) [\sum_{i\in\mathcal{S}} C_{\{i\}}(\mathfrak{B}^{\langle\{i\}\rangle*}) - C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle*})] \\ &= \sum_{i\in\mathcal{N}} \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) 1_i^{\mathcal{S}} C_{\{i\}}(\mathfrak{B}^{\langle\{i\}\rangle*}) - \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle*}) \\ &= \sum_{i\in\mathcal{N}} C_{\{i\}}(\mathfrak{B}^{\langle\{i\}\rangle*}) - \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle*}) \qquad (A.9) \end{aligned}$$

and the right-hand side of (3.8):

$$v(\mathcal{N}) = \sum_{i\in\mathcal{N}} C_{\{i\}}(\mathfrak{B}^{\langle\{i\}\rangle*}) - C_{\mathcal{N}}(\mathfrak{B}^{\langle\mathcal{N}\rangle*}) \tag{A.10}$$

Combining (A.9) and (A.10), it means that this proposed *cooperative game* is balanced as long as for any *balanced map* γ,

$$C_{\mathcal{N}}(\mathfrak{B}^{\langle\mathcal{N}\rangle*}) \leq \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S}) C_{\mathcal{S}}(\mathfrak{B}^{\langle\mathcal{S}\rangle*}) \tag{A.11}$$

Rewriting (A.11) using (3.34), the proof becomes

$$\theta[\mathcal{N}] \leq \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S})\theta[\mathcal{S}], \forall\gamma : 2^{\mathcal{N}} \rightarrow [0,1] \tag{A.12}$$

The following can then be derived:

$$\begin{aligned}
&\sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S})\theta[\mathcal{S}] \\
&= \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \theta[\gamma(\mathcal{S})\mathcal{S}], \quad \text{by } \textit{positive homogeneity} \text{ of } \theta \\
&\geq \theta\,[\sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S})\mathcal{S}], \quad \text{by } \textit{subadditivity} \text{ of } \theta \\
&= \theta\,[\sum_{i\in\mathcal{N}} \sum_{\mathcal{S}\in 2^{\mathcal{N}}} \gamma(\mathcal{S})\, 1_i^{\mathcal{S}}\{i\}], \quad \text{by } \textit{Lemma } 3.2 \\
&= \theta\,[\sum_{i\in\mathcal{N}} \{i\}] = \theta[\mathcal{N}], \quad \text{by } \gamma(\mathcal{S}) \text{ definition, } \textit{Lemma } 3.2
\end{aligned}$$

Hence (A.12) is proven, further proving the game is balanced.

Maybe stories are just data with a soul.

— Brene Brown

B
Case Studies Inputs

The inputs for the case studies in this thesis are largely UK based. However, the coalitional energy management scheme is not location specific, and can be implemented anywhere with the relevant supporting infrastructure.

The following list describes the main inputs for those case studies.

1. Customer consumption: domestic energy consumption data measured in the Customer-Led Network Revolution trials [182], a UK smart grid demonstration project led by Northern Powergrid. The smart meter data used in this thesis are from Test Cell 5, collected for 155 residential UK customers at 1-minute resolution from 01/06/2012 to 31/03/2014. The data are post-processed to 15-min resolution and truncated to the time span of 01/01/2013-31/12/2013 before being applied to the case studies. Profiles that have missing entries are discarded, and a total of 102 profiles are selected for usage in this thesis.

2. PV generation: solar PV generation profile simulated in PVWatts [183] using a standard 4kW, fixed 20 degree tilt residential PV system, and the London Gatwick solar data accessed in June 2019.

3. ES model: a standard battery with an *energy capacity* of 7 kWh, a maximum charge power of 3.5 kW, a maximum discharge power of 3.2 kW, both charge and discharge

efficiencies of 95%, an *initial state of charge* of 50%, and a *state of charge* range of 20-95%.

4. Energy import price: UK Economy 7 residential rate structure as of June 2019: £0.08/kWh for midnight–7am, and £0.18/kWh for 7am–midnight [29]

5. Energy export price: June 2019 level of UK feed-in tariff [28] fixed at £0.0379/kWh.

The models in this thesis are run on a computer equipped with 16 GB RAM and a 2.8 GHz Quad-Core Intel Core i5 processor.

History has shown us that courage can be contagious, and hope can take on a life of its own.

— Michelle Obama

Bibliography

[1] Y. Parag and B. K. Sovacool, "Electricity market design for the prosumer era," *Nature Energy*, vol. 1, no. 4, p. 16032, 2016.

[2] S. Hunt and G. Shuttleworth, *Competition and Choice in Electricity*. Wiley, 1996.

[3] a. K. David, "Transmission planning and investment under competitive electricity market environment," *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, vol. 3, no. C, pp. 1725–1730, 2001.

[4] B. C. Farhar, "Willingness to Pay for Electricity from Renewable Resources : A Review of Utility Market Research Office of Power Technologies," *Renewable Energy*, no. July, pp. VI+20, 1999.

[5] M. Bazilian, I. Onyeji, M. Liebreich, I. MacGill, J. Chase, J. Shah, D. Gielen, D. Arent, D. Landfear, and S. Zhengrong, "Re-considering the economics of photovoltaic power," *Renewable Energy*, vol. 53, pp. 329–338, 2013.

[6] J. Skea, D. Anderson, T. Green, R. Gross, P. Heptonstall, and M. Leach, "Intermittent renewable generation and maintaining power system reliability," *IET Generation, Transmission Distribution*, vol. 2, no. 1, pp. 82–89, January 2008.

[7] J. von Appen, M. Braun, T. Stetz, K. Diwold, and D. Geibel, "Time in the Sun," *IEEE Power & Energy Magazine*, no. February, pp. 55–64, 2013.

[8] A. Pyrgou, A. Kylili, and P. A. Fokaides, "The future of the feed-in tariff (fit) scheme in europe: The case of photovoltaics," *Energy Policy*, vol. 95, pp. 94 – 102, 2016.

[9] S. Griffiths, "Renewable energy policy trends and recommendations for GCC countries," *Energy Transitions*, vol. 1, no. 1, p. 3, 2017.

[10] K. C. Divya and J. Østergaard, "Battery energy storage technology for power systems-An overview," *Electric Power Systems Research*, vol. 79, no. 4, pp. 511–520, 2009.

[11] J. A. Lopes, N. Hatziargyriou, J. Mutale, P. Djapic, and N. Jenkins, "Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities," *Electric Power Systems Research*, vol. 77, no. 9, pp. 1189–1203, 2007.

[12] S. Vazquez, S. M. Lukic, E. Galvan, L. G. Franquelo, and J. M. Carrasco, "Energy Storage Systems for Transport and Grid Applications," *IEEE Trans. Ind. Electron.*, vol. 57, no. 12, pp. 3881–3895, 2010.

[13] D. J. Hill, T. Liu, and G. Verbič, "Smart grids as distributed learning control," *2012 IEEE Power and Energy Society General Meeting*, pp. 1–8, 2012.

[14] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, 2010.

[15] S. Bull, "Renewable energy today and tomorrow," *Proceedings of the IEEE*, vol. 89, no. 8, pp. 1216–1226, 2001.

[16] F. Beck and E. Martinot, "Renewable Energy Policies and Barriers," pp. 365–383, 2004.

[17] G. Chicco and P. Mancarella, "Distributed multi-generation: A comprehensive view," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 3, pp. 535–551, 2009.

[18] C. Masters, "Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines," *Power Engineering Journal*, vol. 16, no. 1, pp. 5–12, 2002.

[19] M. GUARNIERI, M. Liserre, T. Sauter, and J. Y. Hung, "Future energy systems: Integrating renewable energy sources into the smart power grid through industrial electronics," *IEEE Industrial Electronics Magazine*, vol. 4, no. 1, pp. 18–37, 2010.

[20] P. Menanteau, D. Finon, and M. L. Lamy, "Prices versus quantities: Choosing policies for promoting the development of renewable energy," *Energy Policy*, vol. 31, no. 8, pp. 799–812, 2003.

[21] Y. Yamamoto, "Pricing electricity from residential photovoltaic systems: A comparison of feed-in tariffs, net metering, and net purchase and sale," *Solar Energy*, vol. 86, no. 9, pp. 2678–2685, 2012.

[22] Z. Abdmouleh, R. A. M. Alammari, and A. Gastli, "Review of policies encouraging renewable energy integration & best practices," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 249–262, 2015.

[23] J. T. Wenders, "Peak Load Pricing in the Electric Utility Industry," *The Bell Journal of Economics*, vol. 7, no. 1, pp. 232–241, apr 1976. [Online]. Available: http://www.jstor.org/stable/3003198

[24] D. S. Callaway and I. A. Hiskens, "Achieving Controllability of Electric Loads," *Proceedings of the IEEE*, vol. 99, no. 1, 2010.

[25] K. Moslehi and R. Kumar, "A reliability perspective of the smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 57–64, 2010.

[26] H. Klinge Jacobsen and S. T. Schroder, "Curtailment of renewable generation: Economic optimality and incentives," *Energy Policy*, vol. 49, pp. 663–675, 2012.

[27] T. Couture and Y. Gagnon, "An analysis of feed-in tariff remuneration models: Implications for renewable energy investment," *Energy Policy*, vol. 38, no. 2, pp. 955–965, 2010.

[28] GOV.UK. (2019, Feb.) Feed-in tariffs: get money for generating your own electricity. [Online]. Available: https://www.ofgem.gov.uk/environmental-programmes/fit/fit-tariff-rates

[29] GOV.UK. (2019, June) Domestic energy bills. [Online]. Available: https://www.gov.uk/government/collections/domestic-energy-prices

[30] H. Wirth, "Recent facts about photovoltaics in Germany," *Fraunhofer ISE*, vol. 1, no. 100, p. 92, 2017. [Online]. Available: http://pschuetzenduebe.webclient5.de/wp-content/uploads/130912-Recent-Facts-PV-Germany.pdf

[31] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52–62, 2009.

[32] W. A. Braff, J. M. Mueller, and J. E. Trancik, "Value of storage technologies for wind and solar energy," *Nature Climate Change*, vol. 6, no. 10, pp. 964–969, 2016.

[33] L. Barth, N. Ludwig, E. Mengelkamp, and P. Staudt, "A comprehensive modelling framework for demand side flexibility in smart grids," *Computer Science - Research and Development*, vol. 33, no. 1-2, pp. 13–23, 2018. [Online]. Available: https://doi.org/10.1007/s00450-017-0343-x

[34] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, 2011.

[35] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 82–88, 2010.

[36] M. Nick, R. Cherkaoui, and M. Paolone, "Optimal allocation of dispersed energy storage systems in active distribution networks for energy balance and grid support," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2300–2310, 2014.

[37] M. S. Whittingham, "History, Evolution, and Future Status of Energy Storage," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1518–1534, 2012.

[38] T. Morstyn, B. Hredzak, and V. G. Agelidis, "Control strategies for microgrids with distributed energy storage systems: An overview," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016.

[39] K. Worthmann, C. M. Kellett, P. Braun, L. Grüne, and S. R. Weller, "Distributed and Decentralized Control of Residential Energy Systems Incorporating Battery Storage," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1914–1923, 2015.

[40] Y. Gu, X. Xiang, W. Li, and X. He, "Mode-adaptive decentralized control for renewable dc microgrid with enhanced reliability and flexibility," *IEEE Transactions on Power Electronics*, vol. 29, no. 9, pp. 5072–5080, Sept 2014.

[41] Y. S. F. Eddy, H. B. Gooi, and S. X. Chen, "Multi-agent system for distributed management of microgrids," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 24–34, Jan 2015.

[42] T. Morstyn, B. Hredzak, and V. G. Agelidis, "Distributed cooperative control of microgrid storage," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2780–2789, Sept 2015.

[43] M. Liu, P. K. Phanivong, and D. S. Callaway, "Electric vehicle charging control in residential distribution network: A decentralized event-driven realization," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 214–219.

[44] T. Morstyn, B. Hredzak, R. P. Aguilera, and V. G. Agelidis, "Model predictive control for distributed microgrid battery energy storage systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 3, pp. 1107–1114, May 2018.

[45] K. Zhang, Y. Mao, S. Leng, S. Maharjan, Y. Zhang, A. Vinel, and M. Jonsson, "Incentive-Driven Energy Trading in the Smart Grid," *IEEE Access*, vol. 4, pp. 1243–1257, 2016.

[46] G. Andersson and M. Geidl, "Optimal Power Flow of Multiple Energy Carriers," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 145–155, 2007.

[47] T. J. Lui., W. Stirling, and H. O. Marcy, "Get Smart," *IEEE Power and Energy Magazine*, vol. 8, no. 3, pp. 66–78, 2010.

[48] K. Hamilton and N. Gulhar, "Taking demand response to the next level," *IEEE Power and Energy Magazine*, vol. 8, no. 3, pp. 60–65, 2010.

[49] P. Mazza, "The Smart Energy Network: Electricity's Third Great Revolution," *Climate Solutions.org*, pp. 1–22, 2003.

[50] H. Beitollahi and G. Deconinck, "Peer-to-Peer Networks Applied to Power Grid," *International Conference on Risks and Security of Internet and Systems (CRiSIS)*, no. November, p. 8, 2007.

[51] T. Sousa, T. Soares, P. Pinson, F. Moret, T. Baroche, and E. Sorin, "Peer-to-peer and community-based markets: A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 104, pp. 367 – 378, 2019.

[52] D. Fudenberg and J. Tirole, "Dominated Strategies," *Game Theory, MIT Press Books*, vol. 2008, no. 14th August, pp. xxiv+579, 1984.

[53] R. Serrano, "Four Lectures on Cooperative Game Theory Delivered at the University of Sao Paulo International Workshop of the Game Theory Society Lecture 1 : Introduction and Cooperative Bargain- ing," no. July, 2014.

[54] T. L. Turocy, B. V. Stengel, and B. von Stengel, "Game Theory," *Encyclopedia of Information Systems*, pp. 403 – 420, 2003.

[55] P. Berenbrink, T. Friedetzky, L. A. Goldberg, P. W. Goldberg, Z. Hu, and R. Martin, "Distributed selfish load balancing," *SIAM Journal on Computing*, vol. 37, no. 4, pp. 1163–1181, 2007.

[56] L. Jia and L. Tong, "Dynamic pricing and distributed energy management for demand response," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1128–1136, March 2016.

[57] N. Liu, X. Yu, C. Wang, and J. Wang, "Energy sharing management for microgrids with pv prosumers: A stackelberg game approach," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1088–1098, June 2017.

[58] M. Yazdani-Damavandi, N. Neyestani, M. Shafie-khah, J. Contreras, and J. P. S. Catalao, "Strategic behavior of multi-energy players in electricity markets as aggregators of demand side resources using a bi-level approach," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 397–411, Jan 2018.

[59] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, Sept 2012.

[60] W. Saad, Z. Han, and H. V. Poor, "Coalitional game theory for cooperative microgrid distribution networks," *IEEE International Conference on Communications*, pp. 6–10, 2011.

[61] E. Baeyens, E. Y. Bitar, P. P. Khargonekar, and K. Poolla, "Coalitional aggregation of wind power," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3774–3784, 2013.

[62] E. Mengelkamp, J. Gärttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets: A case study: The Brooklyn Microgrid," *Applied Energy*, vol. 210, pp. 870–880, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.06.054

[63] GOV.UK. (2019) Oxfordshire will receive £40 million of funding for a project to trial a new smart local energy system ? or ?smart grid? [Online]. Available: https://www.oxford.gov.uk/news/article/1050/oxfordshire_to_trial_40m_industry_first_local_energy_system_project

[64] EU-Project-Horizon-2020. (2020) Open virtual neighbourhood network to connect iot infrastructures and smart objects. [Online]. Available: https://vicinity2020.eu/vicinity/content/summary

[65] EU-Project-Horizon-2020. (2020) Storage capacity over virtual neighbourhoods of energy ecosystems. [Online]. Available: http://www.sharqproject.eu/the-project

[66] U. TCP. (2020) Global observatory on peer-to-peer, community self-consumption and transactive energy models. [Online]. Available: https://userstcp.org/annex/peer-to-peer-energy-trading/

[67] I. A. Sajjad, G. Chicco, and R. Napoli, "Definitions of demand flexibility for aggregate residential loads," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2633–2643, Nov 2016.

[68] N. Martin and J. Rice, "Solar feed-in tariffs: Examining fair and reasonable retail rates using cost avoidance estimates," *Energy Policy*, vol. 112, pp. 19 – 28, 2018.

[69] L. Einav, C. Farronato, and J. Levin, "Peer-to-Peer Markets," *Annual Review of Economics*, vol. 8, no. 1, pp. 615–635, 2016.

[70] Y. Zhou, J. Wu, and C. Long, "Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework," *Applied Energy*, vol. 222, no. February, pp. 993–1022, 2018.

[71] R. Zafar, A. Mahmood, S. Razzaq, W. Ali, U. Naeem, and K. Shehzad, "Prosumer based energy management and sharing in smart grid," *Renewable and Sustainable Energy Reviews*, vol. 82, no. April 2017, pp. 1675–1684, 2018.

[72] A. Pouttu, J. Haapola, P. Ahokangas, Y. Xu, M. Kopsakangas-Savolainen, E. Porras, J. Matamoros, C. Kalalas, J. Alonso-Zarate, F. D. Gallego, J. M. Martin, G. Deconinck, H. Almasalma, S. Clayes, J. Wu, M. Cheng, F. Li, Z. Zhang, D. Rivas, and S. Casado, "P2P model for distributed energy trading, grid control and ICT for local smart grids," *EuCNC 2017 - European Conference on Networks and Communications*, 2017.

[73] J. Guerrero, A. C. Chapman, and G. Verbic, "Decentralized P2P Energy Trading under Network Constraints in a Low-Voltage Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5163–5173, 2018.

[74] T. Baroche, P. Pinson, R. L. G. Latimier, and H. Ben Ahmed, "Exogenous Cost Allocation in Peer-to-Peer Electricity Markets," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2553–2564, 2019.

[75] J. Kim and Y. Dvorkin, "A P2P-dominant Distribution System Architecture," *IEEE Transactions on Power Systems*, pp. 1–1, 2019.

[76] C. Park and T. Yong, "Comparative review and discussion on P2P electricity trading," *Energy Procedia*, vol. 128, pp. 3–9, 2017. [Online]. Available: https://doi.org/10.1016/j.egypro.2017.09.003

[77] F. Sossan, E. Namor, R. Cherkaoui, and M. Paolone, "Achieving the Dispatchability of Distribution Feeders Through Prosumers Data Driven Forecasting and Model Predictive Control of Electrochemical Storage," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1762–1777, 2016.

[78] E. Sorin, L. Bobo, and P. Pinson, "Consensus-Based Approach to Peer-to-Peer Electricity Markets with Product Differentiation," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 994–1004, 2019.

[79] E. Munsing, J. Mather, and S. Moura, "Blockchains for decentralized optimization of energy resources in microgrid networks," *1st Annual IEEE Conference on Control Technology and Applications, CCTA 2017*, vol. 2017-Janua, pp. 2164–2171, 2017.

[80] D. Vangulick, B. Cornelusse, and D. Ernst, "Blockchain for peer-to-peer energy exchanges: Design and recommendations," *20th Power Systems Computation Conference, PSCC 2018*, 2018.

[81] M. Mihaylov, R. Radulescu, I. Razo-Zapata, S. Jurado, L. Arco, N. Avellana, and A. Nowé, "Comparing stakeholder incentives across state-of-the-art renewable support mechanisms," *Renewable Energy*, vol. 131, pp. 689–699, 2019.

[82] M. Andoni, V. Robu, D. Flynn, S. Abram, D. Geach, D. Jenkins, P. McCallum, and A. Peacock, "Blockchain technology in the energy sector: A systematic review of challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 100, no. November 2018, pp. 143–174, 2019.

[83] C. Zhang, J. Wu, Y. Zhou, M. Cheng, and C. Long, "Peer-to-Peer energy trading in a Microgrid," *Applied Energy*, vol. 220, no. December 2017, pp. 1–12, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2018.03.010

[84] T. Morstyn and M. McCulloch, "Multi-Class Energy Management for Peer-to-Peer Energy Trading Driven by Prosumer Preferences," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 4005–4014, 2018.

[85] W. Tushar, C. Yuen, H. Mohsenian-Rad, T. Saha, H. V. Poor, and K. L. Wood, "Transforming energy networks via peer-to-peer energy trading: The potential of game-theoretic approaches," *IEEE Signal Processing Magazine*, vol. 35, no. 4, pp. 90–111, 2018.

[86] A. J. D. Rathnayaka, V. M. Potdar, T. Dillon, O. Hussain, and S. Kuruppu, "Goal-oriented prosumer community groups for the smart grid," *IEEE Technology and Society Magazine*, vol. 33, no. 1, pp. 41–48, 2014.

[87] P. Asmus, "Microgrids, Virtual Power Plants and Our Distributed Energy Future," *The Electricity Journal*, vol. 23, no. 10, pp. 72–82, 2010.

[88] D. E. Olivares, C. A. CaÃ±izares, and M. Kazerani, "A centralized energy management system for isolated microgrids," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1864–1875, July 2014.

[89] H. Dagdougui, L. Dessaint, and A. Ouammi, "Optimal power exchanges in an interconnected power microgrids based on model predictive control," *IEEE Power and Energy Society General Meeting*, vol. 2014-Octob, no. October, pp. 1–5, 2014.

[90] Y. Xu, W. Zhang, G. Hug, S. Kar, and Z. Li, "Cooperative control of distributed energy storage systems in a microgrid," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 238–248, Jan 2015.

[91] M. Amini and M. Almassalkhi, "Trading off robustness and performance in receding horizon control with uncertain energy resources," *20th Power Systems Computation Conference, PSCC 2018*, 2018.

[92] O. Okpako, H. S. Rajamani, P. Pillai, U. Anuebunwa, and K. S. Swarup, "Investigation of an optimized energy resource allocation algorithm for a community based virtual power plant," in *2016 IEEE PES PowerAfrica*, June 2016, pp. 153–157.

[93] T. Morstyn, N. Farrell, S. J. Darby, and M. D. McCulloch, "Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants," *Nature Energy*, vol. 3, no. 2, pp. 94–101, 2018.

[94] F. Moret, P. Pinson, and S. Member, "Energy Collectives : a Collaborative Approach to Future Consumer-Centric Electricity Markets," *IEEE Transactions on Power Systems*, pp. 1–2, 2018.

[95] C. Long, J. Wu, C. Zhang, L. Thomas, M. Cheng, and N. Jenkins, "Peer-to-peer energy trading in a community microgrid," *IEEE Power and Energy Society General Meeting*, vol. 2018-Janua, pp. 1–5, 2018.

[96] N. Liu, X. Yu, C. Wang, C. Li, L. Ma, and J. Lei, "Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3569–3583, Sept 2017.

[97] Y. Wu, X. Tan, L. Qian, D. H. Tsang, W. Z. Song, and L. Yu, "Optimal pricing and energy scheduling for hybrid energy trading market in future smart grid," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1585–1596, 2015.

[98] W. El-Baz, P. Tzscheutschler, and U. Wagner, "Integration of energy markets in microgrids: A double-sided auction with device-oriented bidding strategies," *Applied Energy*, vol. 241, no. March, pp. 625–639, 2019.

[99] T. Chen, W. Su, and Y. S. Chen, "An innovative localized retail electricity market based on energy broker and search theory," *2017 North American Power Symposium, NAPS 2017*, 2017.

[100] Y. Cai, T. Huang, E. Bompard, Y. Cao, and Y. Li, "Self-sustainable community of electricity prosumers in the emerging distribution system," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2207–2216, 2017.

[101] T. Morstyn, A. Teytelboym, and M. D. Mcculloch, "Bilateral contract networks for peer-to-peer energy trading," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2026–2035, March 2019.

[102] W. Lee, L. Xiang, R. Schober, and V. W. S. Wong, "Direct electricity trading in smart grid: A coalitional game analysis," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 7, pp. 1398–1411, 2014.

[103] S. Ramyar, A. L. Liu, and Y. Chen, "A Power Market Model in Presence of Strategic Prosumers," *IEEE Transactions on Power Systems*, vol. PP, pp. 1–1, 2019.

[104] E. Mengelkamp, P. Staudt, J. Garttner, and C. Weinhardt, "Trading on local energy markets: A comparison of market designs and bidding strategies," *International Conference on the European Energy Market, EEM*, 2017.

[105] T. Chen, Q. Alsafasfeh, H. Pourbabak, and W. Su, "The next-generation U.S. retail electricity market with customers and prosumers-A bibliographical survey," *Energies*, vol. 11, no. 1, 2018.

[106] J. Lee, J. Guo, J. K. Choi, and M. Zukerman, "Distributed energy trading in microgrids: A game-theoretic model and its equilibrium analysis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3524–3533, 2015.

[107] P. Samadi, V. W. Wong, and R. Schober, "Load Scheduling and Power Trading in Systems with High Penetration of Renewable Energy Resources," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1802–1812, 2016.

[108] N. Zhang, Y. Yan, and W. Su, "A game-theoretic economic operation of residential distribution system with high participation of distributed electricity prosumers," *Applied Energy*, vol. 154, pp. 471–479, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.apenergy.2015.05.011

[109] H. Wang and J. Huang, "Incentivizing Energy Trading for Interconnected Microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2647–2657, 2018.

[110] S. Cui, Y.-W. Wang, and J.-W. Xiao, "Peer-to-Peer Energy Sharing among Smart Energy Buildings by Distributed Transaction," *IEEE Transactions on Smart Grid*, vol. PP, no. c, pp. 1–1, 2019.

[111] K. Anoh, S. Maharjan, A. Ikpehai, Y. Zhang, and B. Adebisi, "Energy Peer-to-Peer Trading in Virtual Microgrids in Smart Grids: A Game-Theoretic Approach," *IEEE Transactions on Smart Grid*, vol. PP, no. c, pp. 1–1, 2019.

[112] W. Tushar, B. Chai, C. Yuen, D. B. Smith, K. L. Wood, Z. Yang, and H. V. Poor, "Three-Party Energy Management with Distributed Energy Resources in Smart Grid," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2487–2498, 2015.

[113] M. Ludkovski and R. Sircar, "Game Theoretic Models for Energy Production," pp. 1–18, 2015.

[114] E. J. Anderson and T. D. Cau, "Implicit collusion and individual market power in electricity markets," *European Journal of Operational Research*, vol. 211, no. 2, pp. 403–414, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.ejor.2010.12.016

[115] W. McNamara, *The California Energy Crisis: Lessons for a Deregulating Industry*. PennWell, 2002.

[116] D. M. Lemoine, D. M. Kammen, and a. E. Farrell, "An innovation and policy agenda for commercially competitive plug-in hybrid electric vehicles," *Environmental Research Letters*, vol. 3, no. 1, p. 014003, 2008.

[117] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain, "Enabling Localized Peer-to-Peer Electricity Trading among Plug-in Hybrid Electric Vehicles Using Consortium Blockchains," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3154–3164, 2017.

[118] L. Deori, K. Margellos, and M. Prandini, "On the connection between nash equilibria and social optima in electric vehicle charging control games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 320 – 14 325, 2017, 20th IFAC World Congress.

[119] Z. J. Ma, D. S. Callaway, and I. A. Hiskens, "Decentralized Charging Control of Large Populations of Plug-in Electric Vehicles," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 67–78, 2013.

[120] B. G. Kim, S. Ren, M. Van Der Schaar, and J. W. Lee, "Bidirectional energy trading for residential load scheduling and electric vehicles," *Proceedings - IEEE INFOCOM*, vol. 31, no. 7, pp. 595–599, 2013.

[121] A. Ghosh, V. Aggarwal, and H. Wan, "Strategic Prosumers: How to set the prices Dynamically in a Tiered Market?" *IEEE Transactions on Industrial Informatics*, vol. PP, no. Section III, p. 1, 2018.

[122] W. Tushar, T. K. Saha, C. Yuen, T. Morstyn, Nahid-Al-Masood, H. V. Poor, and R. Bean, "Grid Influenced Peer-to-Peer Energy Trading," *IEEE Transactions on Smart Grid*, vol. PP, no. c, p. 1, 2019. [Online]. Available: http://arxiv.org/abs/1908.09449

[123] H. Wang, T. Huang, X. Liao, H. Abu-Rub, and G. Chen, "Reinforcement Learning in Energy Trading Game among Smart Microgrids," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 5109–5119, 2016.

[124] A. Paudel, K. Chaudhari, C. Long, and H. B. Gooi, "Peer-to-peer energy trading in a prosumer-based community microgrid: A game-theoretic model," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 8, pp. 6087–6097, 2019.

[125] H. Moulin, "Equal or proportional division of a surplus, and other methods," *International Journal of Game Theory*, vol. 16, no. 3, pp. 161–186, 1987.

[126] L. S. Shapley, "Cores of convex games," *International Journal of Game Theory*, vol. 1, no. 1, pp. 11–26, 1971.

[127] L. Shapley, *On Balanced Sets and Cores*, ser. Memorandum (Rand Corporation). Rand Corporation, 1965.

[128] A. Okada, "The Nash bargaining solution in general n-person cooperative games," *Journal of Economic Theory*, vol. 145, no. 6, pp. 2356–2379, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.jet.2010.07.001

[129] J. F. J. Nash, "Two-Person Cooperative Games Author (s): John Nash Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1906951 REFERENCES Linked references are available on JSTOR for this article : You may need to log in to JSTOR to access t," *Journal of Econometric Society*, vol. 21, no. 1, pp. 128–140, 1953.

[130] D. Schmeidler, "The nucleolus of a characteristic function game," *SIAM Journal on Applied Mathematics*, vol. 17, no. 6, pp. 1163–1170, 1969. [Online]. Available: https://doi.org/10.1137/0117107

[131] M. Maschler, B. Peleg, and L. S. Shapley, "Geometric properties of the kernel, nucleolus, and related solution concepts," *Mathematics of Operations Research*, vol. 4, no. 4, pp. 303–338, 1979. [Online]. Available: https://doi.org/10.1287/moor.4.4.303

[132] C. S. K. Yeung and A. S. Y. Poon, "Game theoretical multi-agent modelling of coalition formation for multilateral trades," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 929–934, 1999.

[133] W. Tushar, T. K. Saha, C. Yuen, P. Liddell, R. Bean, and H. V. Poor, "Peer-to-Peer Energy Trading With Sustainable User Participation: A Game Theoretic Approach," *IEEE Access*, vol. 6, pp. 62 932–62 943, 2018.

[134] L. Han, T. Morstyn, and M. McCulloch, "Constructing prosumer coalitions for energy cost savings using cooperative game theory," in *2018 Power Systems Computation Conference (PSCC)*, June 2018, pp. 1–7.

[135] L. Han, T. Morstyn, and M. McCulloch, "Incentivizing prosumer coalitions with energy management using cooperative game theory," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 303–313, Jan 2019.

[136] P. Chakraborty, E. Baeyens, K. Poolla, P. P. Khargonekar, and P. Varaiya, "Sharing Storage in a Smart Grid: A Coalitional Game Approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4379–4390, 2019.

[137] C. Feng, F. Wen, S. You, Z. Li, F. Shahnia, and M. Shahidehpour, "Coalitional Game-Based Transactive Energy Management in Local Energy Communities," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1729–1740, 2020.

[138] Q. Zhou, M. Shahidehpour, T. Sun, D. Feng, and M. Yan, "Cooperative Game for Carbon Obligation Allocation among Distribution System Operators to Incentivize the Proliferation of Renewable Energy," *IEEE Transactions on Smart Grid*, vol. PP, no. c, pp. 1–1, 2019.

[139] A. C. Chapman, S. Mhanna, and G. Verbič, "Cooperative Game Theory for Non-linear Pricing of Load-side Distribution Network Support," *The 3rd IJCAI Algorithmic Game Theory Workshop*, 2017.

[140] J. Castro, D. Gómez, and J. Tejada, "Polynomial calculation of the Shapley value based on sampling," *Computers and Operations Research*, vol. 36, no. 5, pp. 1726–1730, 2009.

[141] J. Castro, D. Gómez, E. Molina, and J. Tejada, "Improving polynomial estimation of the Shapley value by stratified random sampling with optimum allocation," *Computers and Operations Research*, vol. 82, pp. 180–188, 2017.

[142] J. K. Sankaran, "On finding the nucleolus of an n-person cooperative game," *International Journal of Game Theory*, vol. 19, no. 4, pp. 329–338, Dec 1991.

[143] A. K. Jain, R. P. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[144] G. Chicco, "Overview and performance assessment of the clustering methods for electrical load pattern grouping," *Energy*, vol. 42, no. 1, pp. 68–80, 2012.

[145] G. Chicco, O. M. Ionel, and R. Porumb, "Formation of load pattern clusters exploiting ant colony clustering principles," *IEEE EuroCon 2013*, no. July, pp. 1460–1467, 2013.

[146] P. Stephenson and M. Paun, "Consumer advantages from half-hourly metering and load profiles in the UK competitive electricity market," in *DRPT2000. International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. Proceedings (Cat. No.00EX382)*, apr 2000, pp. 35–40.

[147] C. S. Chen, J. C. Hwang, and C. W. Huang, "Application of load survey systems to proper tariff design," *IEEE Transactions on Power Systems*, vol. 12, no. 4, pp. 1746–1751, nov 1997.

[148] C. Flath, D. Nicolay, T. Conte, C. Van Dinther, and L. Filipova-Neumann, "Cluster analysis of smart metering data: An implementation in practice," *Business and Information Systems Engineering*, vol. 4, no. 1, pp. 31–39, 2012.

[149] J. Nazarko and Z. A. Styczynski, "Application of statistical and neural approaches to the daily load profiles modelling in power distribution systems," in *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, vol. 1, apr 1999, pp. 320–325.

[150] B. D. Pitt and D. S. Kirschen, "Application of data mining techniques to load profiling," *IEEE Power Industry Computer Applications Conference*, pp. 131–136, 1999.

[151] P. Cichosz, *Data mining algorithms: explained using R*. John Wiley & Sons, 2014.

[152] G. Chicco, R. Napoli, and F. Piglione, "Comparisons among clustering techniques for electricity customer classification," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933–940, May 2006.

[153] Y. Li and P. C. Flynn, "Electricity deregulation, spot price patterns and demand-side management," *Energy*, vol. 31, no. 6-7, pp. 908–922, 2006.

[154] T. Chen, K. Qian, A. Mutanen, B. Schuller, P. Jarventausta, and W. Su, "Classification of electricity customer groups towards individualized price scheme design," *2017 North American Power Symposium, NAPS 2017*, no. 1, pp. 4–7, 2017.

[155] Y. Cheng and Y. Li, "Research of classification of electricity consumers based on principal component analysis," *6th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2009*, vol. 5, pp. 201–206, 2009.

[156] Y. Wang, Q. Chen, C. Kang, Q. Xia, and M. Luo, "Sparse and Redundant Representation-Based Smart Meter Data Compression and Pattern Extraction," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2142–2151, 2017.

[157] Y. Lu, T. Zhang, and Z. Zeng, "Adaptive weighted fuzzy clustering algorithm for load profiling of smart grid customers," *2016 IEEE/CIC International Conference on Communications in China, ICCC 2016*, 2016.

[158] R. Li, F. Li, and N. D. Smith, "Multi-Resolution Load Profile Clustering for Smart Metering Data," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4473–4482, 2016.

[159] A. Rajabi, M. Eskandari, M. J. Ghadi, L. Li, J. Zhang, and P. Siano, "A comparative study of clustering techniques for electrical load pattern segmentation," *Renewable and Sustainable Energy Reviews*, no. August, 2019.

[160] Š. Raudys and V. Pikelis, "On dimensionality, sample size, classification error, and complexity of classification algorithm in," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-2, no. 3, pp. 242–252, 1980.

[161] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, and C. Toader, "Customer Characterization Options for Improving the Tariff Offer," *IEEE Power Engineering Review*, vol. 22, no. 11, p. 60, 2002.

[162] S. Haben, C. Singleton, and P. Grindrod, "Analysis and clustering of residential customers energy behavioral demand using smart meter data," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 136–144, 2016.

[163] I. B. Sánchez, I. D. Espinós, L. M. Sarrión, A. Q. López, and I. N. Burgos, "Clients segmentation according to their domestic energy consumption by the use of self-organizing maps," *2009 6th International Conference on the European Energy Market, EEM 2009*, pp. 1–6, 2009.

[164] F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognition*, vol. 44, no. 3, pp. 678–693, 2011.

[165] E. Carpaneto, G. Chicco, R. Napoli, and M. Scutariu, "Electricity customer classification using frequency-domain load pattern data," *International Journal of Electrical Power and Energy Systems*, vol. 28, no. 1, pp. 13–20, 2006.

[166] J. E. Jackson, *PCA With More Than Two Variables*. John Wiley & Sons, 1991, pp. 26–62.

[167] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, 1984. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0098300484900207

[168] M. A. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381–396, 2002.

[169] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, "Development of Low Voltage Network Templates - Part I: Substation Clustering and Classification," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3036–3044, 2015.

[170] D. Müllner, "Fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python," *Journal of Statistical Software*, vol. 53, no. 9, pp. 1–18, 2013.

[171] T. Kohonen, *Self-Organization and Associative Memory*. Springer Berlin Heidelberg, 1989. [Online]. Available: https://doi.org/10.1007%2F978-3-642-88163-3

[172] V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, "An electric energy consumer characterization framework based on data mining techniques," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 596–602, 2005.

[173] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*. Siam, 2007, vol. 20.

[174] T. S. Ferguson, "Part IV. Games in Coalitional Form," *Game Theory*, pp. 1–25.

[175] T. S. H. Driessen, *Cooperative Games, Solutions and Applications*, ser. Theory and decision library: Series C, Game theory, mathematical programming, and operations research. Springer, 1988.

[176] M. Leng and M. Parlar, "Analytic solution for the nucleolus of a three-player cooperative game," *Naval Research Logistics*, vol. 57, no. 7, pp. 667–672, 2010.

[177] H. Wirth, "Recent facts about photovoltaics in Germany," *Institute for Solar Energy Systems*, 2015, download from www.pv-fakten.de, current version June 14, 2018.

[178] E. Perez, H. Beltran, N. Aparicio, and P. Rodriguez, "Predictive power control for PV plants with energy storage," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 482–490, 2013.

[179] M. Gulin, J. Matusko, and M. Vasak, "Stochastic model predictive control for optimal economic operation of a residential DC microgrid," *IEEE International Conference on Industrial Technology (ICIT)*, pp. 505–510, 2015.

[180] E. G. L. Parisio A.; Rikos, "A Model Predictive Control Approach to Microgrid Operation Optimization," *Control Systems Technology, IEEE Transactions*, vol. 1, no. 4, pp. 1813–1827, 2014.

[181] E. Faria, L. A. Barroso, R. Kelman, S. Granville, and M. V. Pereira, "Allocation of firm-energy rights among hydro plants: An Aumann-Shapley approach," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 541–551, 2009.

[182] Network-Revolution. (2014) Enhanced profiling of domestic customers with solar photovoltaics (pv). [Online]. Available: http://www.networkrevolution.co.uk/project-library/dataset-tc5-enhanced-profiling-solar-photovoltaic-pv-users/

[183] Pvwatts.nrel.gov. (2019, June) Pvwatts calculator. [Online]. Available: http://pvwatts.nrel.gov/pvwatts.php

[184] J. P. Christophersen, C. D. Ho, C. G. Motloch, D. Howell, and H. L. Hess, "Effects of Reference Performance Testing during Aging Using Commercial Lithium-Ion Cells," *Journal of The Electrochemical Society*, vol. 153, no. 7, p. A1406, 2006.

[185] J. Kim, H. Chun, M. Kim, J. Yu, K. Kim, T. Kim, and S. Han, "Data-Driven State of Health Estimation of Li-Ion Batteries with RPT-Reduced Experimental Data," *IEEE Access*, vol. 7, pp. 106 987–106 997, 2019.

[186] R. B. Myerson, "Graphs and cooperation in games," *Mathematics of Operations Research*, vol. 2, no. 3, pp. 225–229, 1977.

[187] R. Van Den Brink, "An axiomatization of the Shapley value using a fairness property," *International Journal of Game Theory*, vol. 30, no. 3, pp. 309–319, 2002.

[188] R. Weber, "Probabilistic values for games," Cowles Foundation for Research in Economics, Yale University, Cowles Foundation Discussion Papers 471R, 1977.

[189] S. Han, S. Han, and H. Aki, "A practical battery wear model for electric vehicle charging applications," *Applied Energy*, vol. 113, pp. 1100–1108, 2014.

[190] BloombergNEF. (2019, Dec.) Battery pack prices fall as market ramps up with market average at $156/kwh in 2019. [Online]. Available: https://about.bnef.com/blog/battery-pack-prices-fall-as-market-ramps-up-with-market-average-at-156-kwh-in-2019/

[191] C. Zhou, K. Qian, M. Allan, and W. Zhou, "Modeling of the cost of EV battery wear due to V2G application in power systems," *IEEE Transactions on Energy Conversion*, vol. 26, no. 4, pp. 1041–1050, 2011.