UNIVERSITÄT PASSAU
Fakultät für Informatik und Mathematik

Faculty of Computer Science and Mathematics
University of Passau, Germany

Towards Quality of Service and Fairness in Smart Grid Applications

Dominik Danner

A thesis submitted for
Doctoral Degree (Dr. rer. nat.)
March, 2023

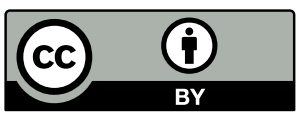
1. Reviewer Prof. Dr.-Ing. Hermann de Meer
Chair of Computer Networks and Computer Communications
University of Passau

2. Reviewer Prof. Dr. Omid Ardakanian
Department of Computer Science
University of Alberta

Dominik Danner

Towards Quality of Service and Fairness in Smart Grid Applications
Ph.D. thesis, March, 2023
Defense, November 9, 2023
Reviewers: Prof. Dr.-Ing. Hermann de Meer and Prof. Dr. Omid Ardakanian

University of Passau, Germany
Faculty of Computer Science and Mathematics
Innstraße 43
94032 Passau



Abstract

Due to the increasing amount of distributed renewable energy generation and the emerging high demand at consumer connection points, *e. g.*, electric vehicles, the power distribution grid will reach its capacity limit at peak load times if it is not expensively enhanced. Alternatively, smart flexibility management that controls user assets can help to better utilize the existing power grid infrastructure for example by sharing available grid capacity among connected electric vehicles or by disaggregating flexibility requests to hybrid photovoltaic battery energy storage systems in households. Besides maintaining an acceptable state of the power distribution grid, these smart grid applications also need to ensure a certain quality of service and provide fairness between the individual participants, both of which are not extensively discussed in the literature. This thesis investigates two smart grid applications, namely electric vehicle *charging-as-a-service* and *flexibility-provision-as-a-service* from distributed energy storage systems in private households.

The electric vehicle charging service allocation is modeled with distributed queuing-based allocation mechanisms which are compared to new probabilistic algorithms. Both integrate user constraints (arrival time, departure time, and energy required) to manage the quality of service and fairness. In the queuing-based allocation mechanisms, electric vehicle charging requests are packetized into logical charging current packets, representing the smallest controllable size of the charging process. These packets are queued at hierarchically distributed schedulers, which allocate the available charging capacity using the time and frequency division multiplexing technique known from the networking domain. This allows multiple electric vehicles to be charged simultaneously with variable charging currents. To achieve high quality of service and fairness among electric vehicle charging processes, dynamic weights are introduced into a *weighted fair queuing scheduler* that considers electric vehicle departure time and required energy for prioritization. The distributed probabilistic algorithms are inspired by medium access protocols from computer networking, such as *binary exponential backoff*, and control the quality of service and fairness by adjusting sampling windows and waiting periods based on user requirements.

The second smart grid application under investigation aims to provide *flexibility provision-as-a-service* that disaggregates power flexibility requests to distributed battery energy storage systems in private households. Commonly, the main purpose of stationary energy storage is to store energy from a local photovoltaic system for later use, *e. g.*, for overnight charging of an electric vehicle. This is optimized locally by a home energy management system, which also allows the scheduling of external flexibility requests defined by the deviation from the optimal power profile at the grid connection point, for example, to perform peak shaving at the transformer. This thesis discusses a *linear heuristic* and a *meta heuristic* to disaggregate a flexibility request to the single participating energy management systems that are grouped into a flexibility pool. Thereby, the linear heuristic iteratively assigns portions of the power flexibility to the most appropriate energy management system for one time slot after another, minimizing the total flexibility cost or maximizing the probability of flexibility delivery. In addition, a multi-objective *genetic algorithm* is proposed that also takes into account power grid aspects, quality of service, and fairness among par-

ticipating households. The genetic operators are tailored to the flexibility disaggregation search space, taking into account flexibility and energy management system constraints, and enable power-optimized buffering of fitness values.

Both smart grid applications are validated on a realistic power distribution grid with real driving patterns and energy profiles for photovoltaic generation and household consumption. The results of all proposed algorithms are analyzed with respect to a set of newly defined metrics on quality of service, fairness, efficiency, and utilization of the power distribution grid. One of the main findings is that none of the tested algorithms outperforms the others in all quality of service metrics, however, integration of user expectations improves the service quality compared to simpler approaches. Furthermore, smart grid control that incorporates users and their flexibility allows the integration of high-load applications such as electric vehicle charging and flexibility aggregation from distributed energy storage systems into the existing electricity distribution infrastructure. However, there is a trade-off between power grid aspects, *e. g.*, grid losses and voltage values, and the quality of service provided. Whenever active user interaction is required, means of controlling the quality of service of users' smart grid applications are necessary to ensure user satisfaction with the services provided.

Acknowledgement

Over the past six years of working on this thesis, I have received great support and assistance from numerous people. Therefore, I would like to express my sincere gratitude to everyone who has supported me during this time.

First and foremost, I would like to thank my supervisor Prof. Dr.-Ing. Hermann de Meer for his support, guidance, and intense discussions in which I learned all the relevant tools for scientific research. Moreover, he also provided me with the opportunity to work on interdisciplinary European and nationwide research projects, which allowed me to collaborate with partners from industry and academia and also brought new ideas to my thesis topic. I want to cordially thank Prof. Dr. Omid Ardakanian, one of the leading experts in the field of smart grid control, for kindly agreeing to serve as a second reviewer.

While working on my thesis, I had the pleasure to work with many talented colleagues, and I want to thank all my former and current colleagues at the *Chair of Computer Networks and Computer Communications*. I am especially grateful to Philipp Danner, my twin brother, for the very fruitful daily discussions during commuting to Passau, which mostly evolved into new impulses for my research. I also thank Dr. Ammar Alyousef for guiding me into the scientific landscape at the beginning of my work. In addition to the members of my research group, I would like to thank all my students whose bachelor's and master's theses opened up new fields in the area of my thesis.

Last but definitely not least, I want to express my gratitude to my beloved wife Tamara for her unconditional belief in me and my work and for her endless love. I also want to thank my family, my parents, and my brothers for their continued encouragement and unconditional support.

Contents

List of Acronyms

AC	*Alternating Current* 4, 11, 22, 24, 25, 29, 34, 35, 89
ADMM	*Alternating Direction Method of Multipliers* 26
aFRR	*automatic Frequency Restoration Reserves* 16, 17, 87, 141–143
AIFS	*Arbitration Inter-Frame Spacing* 52–55
AIMD	*Additive Increase Multiplicative Decrease* 27
API	*Application Programming Interface* 140
AS	*Ancillary Service* 2, 14–16, 23, 30, 35, 87, 89, 138, 141
BDEW	*German Association of Energy and Water Industries* 137
BEB	*Binary Exponential Backoff* 49, 51, 85
BESS	*Battery Energy Storage System* 1, 2, 4–6, 8, 9, 14, 16, 29, 30
BEV	*Battery Electric Vehicle* 3, 21, 22
BMS	*Battery Management System* 23, 25
BRP	*Balance Responsible Party* 138, 140
CCCV	*Constant-Current Constant-Voltage* 23
CDF	*Cumulative Distribution Function* 98, 99, 102, 129
CoS	*Class of Service* 17, 52
CP	*Control Pilot* 24
CPU	*Central Processing Unit* 46, 116, 119
CSMA/CA	*Carrier Sense Multiple Access / Collision Avoidance* 49, 85
CSMA/CD	*Carrier Sense Multiple Access / Collision Detection* 49
CW	*Contention Window* 49–54, 70
DC	*Direct Current* 2, 11, 22, 29, 34
DCF	*Distributed Coordination Function* 49, 51–53, 55, 62, 63, 65–70, 72, 73, 75, 76, 148
DR	*Demand Response* 1, 29, 45
DRES	*Distributed Renewable Energy Source* 2, 12, 14, 83, 84
DSO	*Distribution System Operator* 2, 5, 12, 16, 31, 38, 68, 79, 88, 94, 118, 135, 138, 139, 141
DWFQ	*Dynamically Weighted Fair Queuing* 48, 53, 55, 62, 63, 65–77, 84, 85, 148
EDCA	*Enhanced Distributed Channel Access* 52, 54, 55, 62, 63, 65–68, 70, 72, 73, 75, 76, 148, 151
EDF	*Earliest-Departure-First* 27, 44, 46, 62, 63, 65–68, 70–77, 148
EMS	*Energy Management System* 4–6, 8, 9, 29–35, 87–90, 92–111, 113–123, 126, 129–131, 133–145, 147–150

ENTSO-E *European Network of Transmission System Operators for Electricity* 15
ESS *Energy Storage System* 15, 16, 26, 29–33, 35, 87–98, 100, 103, 117, 118, 120, 136, 137, 140–145, 148, 149
EV *Electric Vehicle* 1–9, 14, 16, 18, 19, 21–35, 37–49, 51–63, 67–75, 77, 79–85, 87–91, 93–98, 100, 117, 119, 120, 122, 137, 138, 140, 141, 143, 147–150
EXI *Efficient XML Interchange* 24

FCFS *First-Come-First-Served* 27, 46, 62, 63, 65, 66, 68, 70, 72, 73, 75, 76, 148
FCR *Frequency Containment Reserves* 16, 17

G2V *Grid to Vehicle* 22, 23, 34
GA *Genetic Algorithm* 8, 31, 87, 104, 108–110, 112, 115, 116, 119–121, 123, 124, 126–129, 131–133, 135–138, 140, 145, 149
GPS *Generalized Processor Sharing* 20, 47

HEV *Hybrid Electric Vehicle* 21, 22
HH *Household* 91

ICT *Information and Communication Technology* 6, 38, 45, 46, 79
IP *Internet Protocol* 24
ISP *Internet Service Provider* 5
ITU *International Telecommunication Union* 17, 18

LLF *Least-Laxity-First* 27, 28, 44, 47, 48, 62, 63, 65–68, 70–77, 148
LP *Linear Programming* 30, 32
LSTM *Long-Short Term Memory* 117, 120

MAC *Medium Access Control* 5, 44, 49, 52, 68, 85, 148
MDP *Markov Decision Process* 31
mFRR *manual Frequency Restoration Reserves* 16, 17
MILP *Mixed Integer Linear Programming* 8, 9, 26, 30, 32, 87, 89, 90, 92, 93, 95, 96, 100–103, 106, 108, 111, 115–117, 119–121, 124, 139–141, 145, 148, 150
MINLP *Mixed Integer Non-Linear Programming* 26
MOS *Mean Opinion Score* 18
MPC *Model Predictive Control* 27

NN *Neural Network* 27, 31

OCPP *Open Charge Point Protocol* 3
OLTC *On Load Tap Changer* 2, 32
OSI *Open Systems Interconnection* 24

PE *Protective Earth* 24
PEV *Plugin Electric Vehicle* 22, 35
PHEV *Plugin Hybrid Electric Vehicle* 4, 14, 22
PLC *Power-Line Communication* 24
PP *Proximity-Pilot* 24
PROP *Proportional* 47, 55, 62, 63, 65–67, 70–73, 75, 76, 148, 149, 154–156
PSO *Particle Swarm Optimization* 27
PV *Photovoltaic* 1–3, 5, 8, 21, 27, 29, 30, 32, 38, 79, 87, 89–94, 96–100, 103, 117–120, 126, 135, 136, 138, 140, 143, 148
PWM *Pulse-Width Modulation* 4, 24, 34, 81

QoE *Quality of Experience* 5–9, 11, 18, 19, 21, 28, 35, 37, 40, 42–44, 46, 56, 62, 64, 66, 84, 85, 100, 102–104, 116, 134, 145, 147–149, 152
QoS *Quality of Service* 2, 5–9, 11, 14, 17–19, 21, 22, 25–28, 31, 33, 35, 37, 38, 40–44, 46, 48, 52–56, 61–66, 68, 71, 74, 75, 81, 84, 85, 87, 100–102, 104, 116, 134, 145, 147–150, 152

RR *Reserve Replacement* 16, 17

SAIDI *System Average Interruption Duration Index* 12
SLP *Standard Load Profile* 1, 136–138
SMGW *Smart Meter Gateway* 38, 83
SoC *State of Charge* 7, 23, 25–29, 34, 39, 42, 43, 46, 48, 63, 65, 70–72, 74, 75, 77, 80, 85, 88, 91, 94, 103, 137, 142, 143, 147, 150
SoH *State of Health* 23, 27, 29, 94
SU *Scheduling Unit* 44–49, 55, 56, 60, 61, 68, 72, 79, 80, 83, 84, 150

TCP *Transmission Control Protocol* 28, 29
TSO *Transmission System Operator* 141

V2G *Vehicle to Grid* 22, 24, 26, 32, 34, 39, 40, 89
V2H *Vehicle to Home* 22, 34
V2L *Vehicle to Load* 22, 34
V2X *Vehicle to X* 22, 34, 150
VPP *Virtual Power Plant* 16, 87, 99, 142

WFQ *Weighted Fair Queuing* 44, 47, 48, 62, 63, 65, 68, 70–73, 75–77, 148

XML *Extensible Markup Language* 24

List of Algorithms

List of Figures

List of Tables

CHAPTER 1

Introduction

Addressing climate change has become the most important goal for the world population in the 21st century, and reducing greenhouse gas emissions is one key factor in achieving this goal. With more than 70 % of the global greenhouse gas emissions coming from the energy sector (energy used in buildings, industry, and transportation) [141], many governments are granting incentives for new renewable energy sources, energy storage, energy-efficient heating, and carbon-neutral transportation, focusing primarily on fossil fuel reduction. Consequently, many distributed *Photovoltaic* (*PV*) installations, *Battery Energy Storage Systems* (*BESSs*), and new electrical loads for heating and transportation such as heat pumps and *Electric Vehicles* (*EVs*) are emerging on the customer side. This chapter provides a short motivation for the two smart grid applications under investigation in this thesis, states the problem formulation, and gives a summary of the contributions.

1.1 Motivation

The power grid is the largest machine built by humankind and typically consists of generators (*producers*), loads (*consumers*), and grid assets (cables, transformers, *etc.*) that connect the former two. Because a large amount of electrical power is difficult to store, the traditional power grid is operated with the paradigm *supply follows demand*, hence centralized (fossil-fueled) generators produce as much power as required to serve all connected loads. With the shift towards renewable energy sources such as wind or solar, which depend on weather conditions and therefore can only be partially controlled, generation can no longer follow demand, and demand-side integration is required for stable power system operation. In addition, generation capacity in the form of *PV* systems is installed at residential loads, turning them into *prosumers*, who consume and produce at different times. Forecasting the future demand and generation of these prosumers is much more complicated than simply using *Standard Load Profiles* (*SLPs*) like it is done in the traditional power system [132]. This paradigm shift requires more flexible loads that participate in market-based *Demand Response* (*DR*) programs or can be controlled remotely by the power system operator or any third party offering control services. Furthermore, the rapid development of *BESS* technologies enables new operation scenarios, where surplus energy can be stored for later self-consumption and these *BESSs* additionally provide flexibility potential for advanced power system operation. However, interaction with flexible loads and storage systems requires additional stakeholders to be involved in the power grid operation, *e. g.*,

end customers and *BESS* owners, and the possible limitations and conditions of the interactions with the customer must be defined by service contracts.[1] Because people tend to not participate in services if they are not satisfied with the results, providing service quality and fairness to the customers becomes more and more important. This shifts the original customer relationship of the power grid operators by 180 degrees, whereby the customer no longer only expects a stable grid connection from the power grid operator, but conversely the power grid operator is dependent on control services from the customer side.

To compensate for demand and supply mismatches or power hardware faults, stable power grid operation continuously requires *Ancillary Service* (*AS*), *e. g.*, inertia, frequency response, reactive power compensation, and voltage regulation, all of which are provided by synchronous generators nowadays. On January 8, 2021, for example, the European transmission system was separated into two parts due to outages of several transmission system assets [49, 138, 139]. This separation caused an increased power system frequency in the South-East area due to oversupply and reduced frequency in the North-West part of Europe due to power deficit. The initial frequency drop was recovered by automatic countermeasures including inertia, additional supportive power supply (frequency response), and ultimately shutoff of interruptible loads in France and Italy. To reduce the frequency in the South-East, generation was reduced until the frequency stabilized within the control limits. This example emphasizes the importance of *AS* for stable power grid operation. Moving from centralized synchronous generators, which can be found in large (fossil-fueled) power plants, to smaller inverter-based renewable generation units results in a reduction of available *ASs*. Of course, special power grid assets such as flywheels, huge *BESS*, *Direct Current* (*DC*) links, capacitor banks, and *On Load Tap Changers* (*OLTCs*) can provide *ASs*, but utilizing existing technical equipment like wallboxes for *EV* charging or stationary *BESS* at households is more ecological and will help to reduce required investment costs. Literature has demonstrated that *EVs* [116, 120, 148] and *BESSs* [63, 118, 172] can contribute to frequency response on a large scale level. However, because some *AS* require the adaption of user behavior or result in reduced power grid feed-in of *PV* systems, the provision of these *ASs* should be done in a way to impact the comfort of the users in the lowest possible way. *Quality of Service* (*QoS*) and fairness for the user[2] are important aspects that should be considered during provision in addition to technical requirements.

The integration of *Distributed Renewable Energy Sources* (*DRESs*) such as *PV* installations and additional relatively high loads such as *EV* charging on the customer side can put tremendous pressure on today's power distribution grids, which initially have been built to only supply loads from mainly higher power grid levels. Especially, keeping power quality standards and voltage bands in acceptable ranges as defined by EN 50160, while not overloading grid assets, is a challenging task for the *Distribution System Operators* (*DSOs*), who typically operate their power distribution grid with little to no measurement devices. Although grid connection codes already require certain passive behavior from connected devices, *e. g.*, reactive power supply for voltage regulation at *PV* inverters or predefined feed-in curtailments, *DSOs* lack the capability for fine-granular active regulation in their

[1] Note that current energy contracts guarantee a 100 % service rate, and therefore participating in power grid operation services can be seen as potential degradation.

[2] Here it is referred to the service that the flexible device provides to the user.

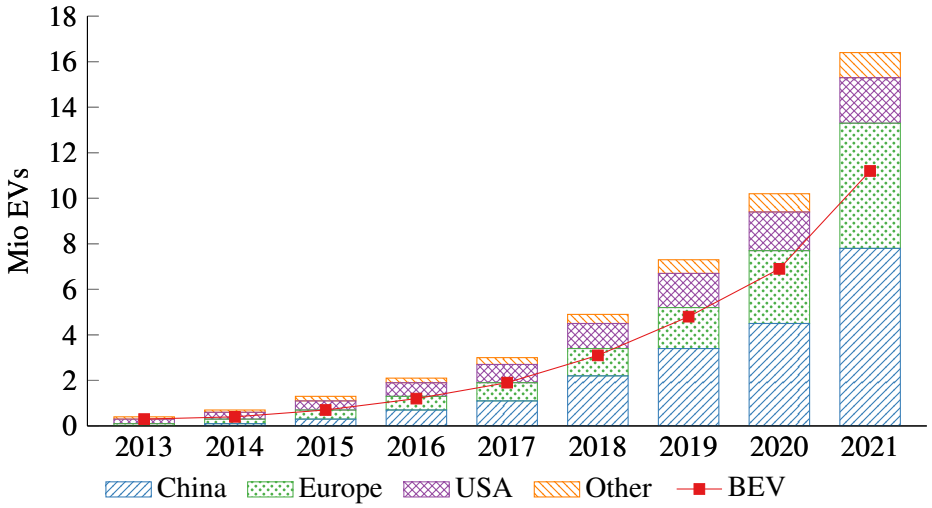


Figure 1.1: Global *EV* stock (battery and plugin hybrid *EVs*) according to data from [82]. The line plot further shows the ratio of *Battery Electric Vehicles* (*BEVs*).

power distribution grids so far. Therefore, power distribution grids are designed and built for estimated peak situations using historical load profiles and assumed simultaneity factors. However, *EV* charging facilities for households and *PV* rooftop installations can have a rated power of more than 20 kW, which exceeds the assumed average household consumption of around 4 kW. Hence, integrating numerous *EVs* and *PV* systems will require expensive grid reinforcements when upgrading the power grid to estimated peak loads. In 2019, E.ON, one of the biggest electric utility operators in Europe, estimated the cost of grid reinforcement to cover 100 % *EV* penetration (approximately 6.5 million cars in their service area) at € 2.5 billion over the next 25 years [45]. With intelligent solutions and incentives that target charging flexibility to shift *EV* charging from peak to low demand times, *e. g.*, from 18:00 evening peak to night hours, the required investment can be halved.

1.1.1 Electric Vehicle Home Charging

The global automotive industry is undergoing a transformation from internal combustion engines to plugin-in *EVs* based on the needs of private sustainable-minded people and policymakers. Thereby, even new technology companies, such as Tesla, enter the market. Figure 1.1 shows that the worldwide sales figures of *EVs* are increasing over the past years leading to a stock of over 16 million *EVs* in 2021, including battery electric and plugin hybrid *EVs*, which both can be charged from the electrical power grid. All these *EVs* need public and private charging points, for which European policymakers started funding campaigns, *e. g.*, Germany invested more than € 500 million into private wallboxes until mid of 2021 [93].

Most already installed charging equipment can be remotely controlled via the *Open Charge Point Protocol* (*OCPP*), which starting from version 1.6 supports charging capacity limitations [127]. Similarly, the communication between a charging point and an *EV* allows con-

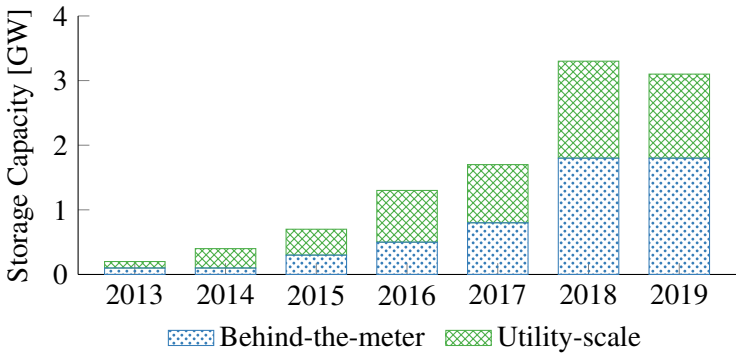


Figure 1.2: Annual global *BESS* behind-the-meter and utility-scale deployment according to data from [81].

trol signals. For example, the legacy Type 2 *Alternating Current* (*AC*) charging, which is widely used in Europe, can be controlled via basic *Pulse-Width Modulation* (*PWM*) signals according to IEC 61851-1 or bidirectional high-level communication using ISO 15118-20. More details on *EV* charging control are given in Section 2.3.1.

Although *EVs* are high power loads, their actual energy requirement depends on the driving behavior of the users. According to [44], *EV* charging processes in domestic charging have certain charging flexibility that reaches a peak in the evening, which aligns with the expected peak load of household loads. This gives the possibility to dynamically change the actual charging time as long as the mobility need of the drivers is satisfied. This charging flexibility makes *EVs* in private households a perfect flexible unit to contribute to power grid operation purposes.

1.1.2 Home Energy Management Systems

With the increasing number of stationary battery storage systems in private households, *e. g.*, between 2013 and 2018 more than 85 000 were installed in Germany [90], more and more combined photovoltaic battery storage systems can provide higher negative and even positive flexibility to the power grid by decreasing or increasing the power feed-in. The potential is reflected by the yearly installed storage capacity depicted in Figure 1.2, which was drastically increasing over the past years. This development is mainly driven by a decrease of more than 50 percent in retail prices of *lithium-ion* batteries during that time, which is reinforced by the rapid development of storage technologies for *EVs* [90].

Furthermore, the trend towards *Plugin Hybrid Electric Vehicles* (*PHEVs*) [80] and heat pumps [83] enables additional flexibility at the household level. These high-power devices, which can be shifted in time, are typically controlled by home *Energy Management Systems* (*EMSs*) optimizing for customers' objectives, *e. g.*, maximize self-consumption of local photovoltaic generation or the degree of autarky, avoid feed-in limits at the power grid connection point, or minimize total energy costs [60]. Nevertheless, to achieve a successful energy transition towards 100 % renewable energy, the flexibility potential of distributed

households must be utilized from a more global viewpoint, while customers' perspectives should not be ignored to keep users satisfied with the provided service. Because *BESSs* are typically used to store energy from local *PV* installations for later use, metrics like autarky and self-consumption can be used to quantify the service quality of the *BESS* toward the user. On the other hand, aspects of flexibility delivery probability might be of importance for flexibility aggregators. Research projects [1, 61] have technically demonstrated how local flexibility markets may look like. However, their focus is more on the technical realization of flexibility than service quality of aggregated flexibility or users' recognition of their system in use.

1.2 Problem Formulation

Newly emerging high-power devices such as *EVs* and stationary *BESSs* must smoothly be integrated into the power distribution grid. Therefore, to avoid unnecessary grid expansion, their flexibility can be used to adapt their power profile to fit the *DSO*'s needs. However, most control strategies in literature do not care much about the service quality of connected smart grid applications. In our opinion, *Quality of Service*, *Quality of Experience* (*QoE*), and fairness aspects are mandatory to discuss, because finally the customers are involved when intervening with their devices, which transforms the technical smart grid applications into a socio-technical challenge.

A similar trend can be seen in orchestrating communication networks, where some *Internet Service Providers* (*ISPs*) use traffic shapers to guarantee a certain *QoS*, *e. g.*, video quality for Internet-based television and streaming, or to restrict bandwidth usage of certain applications, *e. g.*, peer-to-peer file sharing. Network resource allocation mechanisms are also used in data centers to guarantee service level agreements to the customers. Among others, queuing techniques such as fair queuing or packet prioritization are used to optimize the utilization of the shared network, typically in a central or distributed hierarchical manner. A fully distributed resource sharing can be found in *Medium Access Control* (*MAC*) protocols, where several communication endpoints share the same physical medium, *e. g.*, a cable or the air. Thereby, they need to coordinate who is allowed to use the medium and at which time. Among others, probabilistic methods or master-slave protocols are used, whereby priority access to the medium can be used to guarantee *QoS* for a certain application.

In the future, this might require new supply agreements between power grid operators, aggregators, and customers that include service level agreements in terms of provided service on the application level, *e. g.*, *charging-as-a-service* instead of flat-rate connection service. Similar to *ISPs* that offer a connection with up to a certain bandwidth. These contracts can adapt to specific application fields, *e. g.*, flexible charging power for *EV* charging. This thesis investigates two specific smart grid applications, namely *EV charging-as-a-service* and *flexibility disaggregation service* to distributed *EMSs*. Actively controlling these smart grid applications requires consideration of the effect of the control signal on the power grid and the service degradation that can be expected at each customer. The following requirements target technical applicability, interoperability with controllable devices, and the acceptance of the user.

- **Legacy Equipment**: The rollout of technical equipment for *EV* charging and *EMSs* as well as stationary *BESSs* already started and with evolving time, the capabilities of these devices improve. However, a control solution should work with a default set of control options to be wide-range applicable and backward compatible.
- **Power grid utilization**: Control strategies should be designed in a way that they can operate with the existing power grid infrastructure to avoid grid reinforcements and, therefore, save material and emissions. Besides overloading situations, voltage violations must be avoided as well. Finally, control strategies should efficiently utilize the power grid infrastructure to reduce power losses.
- **Reliability and safety**: *Information and Communication Technology* (*ICT*)-based control strategies should be able to operate in case of communication failures and by design consider principles of functional safety in critical infrastructures.
- **Scalability**: Given the target number of *EVs* from political decisions and the expected number of distributed *EMSs* with *BESS*, the control mechanisms for smart grid applications must efficiently scale with the number of participants and the underlying power grid infrastructure.
- **Responsiveness**: Energy markets work in 15 min time intervals; hence, all set-point planning should be possible in this time horizon. However, power systems have faster dynamics and a reaction to threshold violations should be possible in a timescale of seconds to minutes to avoid harming the power grid.
- **Quality of Service/Quality of Experience**: To keep participating customers happy, the delivered *QoS* of the smart grid application must be continuously high. Otherwise, customers might not be attracted by the smart grid application service and switch to another flexibility aggregator or request a non-controlled grid connection from the power grid operator.
- **Fairness**: Besides high *QoS* and *QoE*, fairness among participants is similarly important for smart grid applications. If customers are served unfairly, their trust in smart controller logic is damaged, and they might try to trick the system.

The discussed control strategies in this thesis will target all the aforementioned requirements, mainly focusing on the customers' perspective as well as grid-related challenges.

1.3 Contributions

We study the *Quality of Service* (*QoS*), *Quality of Experience* (*QoE*), and fairness aspects of smart grid applications in the two main areas of (1) *charging-as-a-service* for distributed *EV* home charging, and (2) *flexibility provision-as-a-service* to distributed home *EMS*, including *EV* charging processes and stationary *BESS*. While mainly focusing on the service of the smart grid applications, we additionally discuss means of controlling voltage and power grid losses from a grid operators' perspective.

Given the scenario of *EV* home charging, where certain service parameters are known in advance, we answer the following research question.

C1 *How to quantify the* QoS *and* QoE *of* EV *charging-as-a-service?* To answer this question, we propose a comprehensive set of *QoS* and *QoE* metrics in Sections 3.3.1 and 3.3.2. Among others, these metrics consider the ratio of charged energy to the required energy, the continuity of charging power rate, the battery *State of Charge* (*SoC*) at departure, and the ability to reach the destination of an upcoming trip. Compared to the literature, the proposed metrics cover the full charging process as a service, instead of *EV* charging as a grid participant.

C2 *How to quantify experienced fairness between charging processes?* We discuss fairness notions in Section 2.2 and apply the fairness index by Hoßfeld et al. to *EV charging-as-a-service*, whereby we not only examine the fairness of the quality metrics at the very end of a charging process but also throughout the whole charging process duration.

C3 *How can charging-as-a-service be realized with high* QoS*,* QoE*, and fairness in a hierarchical and distributed probabilistic way?* We propose an efficient and hierarchically scalable packet queuing allocation mechanism in Section 3.4.1 that takes the residual charging time and the current *SoC* into account and ensures fairness among charging services. The provided models include not only temporal charging slot allocation (with fixed charging power rates), but also distribute the available charging capacity during each time slot while respecting charging hardware limitations and control protocol capabilities. In Section 3.4.2, we further propose a novel distributed probabilistic control strategy inspired by communication network protocols to allocate available charging power capacity. Thereby, charging service parameters are used to control *QoS* aspects by configuring additional waiting time and random backoff windows in case of power grid threshold violations.

C4 *How well do queuing-based and probabilistic solutions perform concerning to* QoS*,* QoE*, and fairness?* In Section 3.5, our proposed solutions are evaluated on the IEEE 906 low-voltage test feeder with real user driving profiles extracted from a mobility survey. In contrast to simpler queuing policies, the proposed *dynamically weighted fair queuing* approach achieves both high *QoS* results and good fairness indices throughout the whole charging service. The probabilistic approaches reach lower service quality due to a lack of global information; however, scale better with the number of participating *EVs*.

C5 *How can voltage levels be maintained by charging-as-a-service and what impact does it have on the service quality and fairness?* In the simplest case, each *EV* will need to react to voltage violations locally as defined in Section 3.5.1.2, *e. g.*, using $Q(U)$ and $P(U)$ droop controllers according to VDE-AR-N 4100. However, this may lead to service quality and fairness degradation as shown in experiments in Section 3.5.3.3. Furthermore, the simultaneous reaction of *EVs* might cause undesirable high power grid losses. To reduce this, we further discuss a decentralized mechanism to orchestrate reactive power provision based on the max-consensus protocol

enriched with domain knowledge of the power distribution grid in Section 3.6.4 as an extension to the proposed charging power allocation mechanisms.

Given the scenario of home *EMS* in private households, where day-ahead load and generation profiles are known in advance, we answer the following research question.

C6 *How to utilize local flexibility from* EVs *and* BESS *for optimal day-ahead operation and external flexibility requests?* In Section 4.2, a detailed *Mixed Integer Linear Programming* (*MILP*) formulation is introduced that models the day-ahead planning of an *EMS* at a private household. This includes scheduling *EV* charging services as well as optimally charging and discharging a local *BESS*. In addition to local optimization, the solution space of the *MILP* yields the flexibility potential for external flexibility requests from the flexibility aggregator.

C7 *How to quantify the* QoS *and* QoE *of flexibility provision-as-a-service?* To answer this research question, we propose a comprehensive set of *QoS* and *QoE* metrics in Sections 4.3.1 and 4.3.2 that consider quality aspects relevant for the resource provider as well as the resource aggregator. Among others, these metrics target the ratio of service participation of an individual *EMS*, the reached level of self-consumption and autarky, as well as the total cost for the resource aggregator.

C8 *How to disaggregate flexibility requests to a pool of* EMSs *efficiently to optimize for the flexibility cost, delivery probability, power grid losses, as well as* QoS/QoE *and fairness among the* EMSs*?* We propose a linear heuristic for flexibility disaggregation in Section 4.4.1 that targets minimization cost or maximization of flexibility delivery probability. This method however is limited to a single convex objective. For a more generalized setup, we present a meta heuristic to include *QoS*/*QoE* and fairness aspects, as well as power grid losses in the multi-objective optimization in Section 4.4.2. Thereby the proposed *Genetic Algorithm* (*GA*) does not implement the classical procedure but applies either the crossover or the mutation operator to each chromosome in one generation. This improves the performance of the *GA* in two ways. First, the chance of obtaining an invalid chromosome is reduced, and, second, due to efficient fitness value caching, the number of processed chromosomes is increased within the same time limitation.

C9 *How well do linear and meta heuristic perform concerning to* QoS*,* QoE*, and fairness?* In Section 4.5, the proposed linear and meta heuristic are evaluated on their performance with regard to the defined quality metrics. The tested flexibility aggregator consists of 55 households that operate their own *EMS* with recorded household and *PV* profiles from real systems and are connected to the same IEEE 906 low-voltage test feeder.

1.4 Chapter Summary and Thesis Structure

Fighting against global warming is the most important task of modern society. This can only be achieved through a massive reduction of greenhouse gas emissions, leading to

drastic changes in global energy management. The increasing number of *EVs* and *BESS* behind-the-meter will pose new challenges to today's electrical power system, but also offers the opportunity to take advantage of their flexibility. However, the transition from flat-rate tariffs to service-based smart grid application contracts involves interference with the normal usage pattern of the assets, thus putting the user into the control loop. To ensure the convenience of the user and to keep them participating with their flexibility in the power system operation, automatic control methods must take into account *QoS* and fairness of the smart grid applications. In addition, other requirements such as support for legacy equipment, optimal power grid asset utilization, and scalability must be ensured. This problem formulation leads to a number of contributions in this thesis, which are organized as follows.

First, Chapter 2 gives an overview of the European power system, flexibility in the power domain, and ancillary services that are required to operate the power grid. The chapter further discusses the notion of *QoS*, *QoE*, and fairness, which come from the networking domain. Furthermore, required background knowledge on the two smart grid applications under investigation is provided, including control capabilities and typically used communication protocols. A literature review provides an overview on the domain of *EV* charging control and power flexibility provision to distributed *EMS*, whereby different notions of *QoS*, *QoE*, and fairness in the domains are highlighted. The chapter concludes by providing the general assumptions that are valid for both smart grid applications under investigation.

In the context of online charging control, *EV charging-as-a-service* is discussed in Chapter 3. First additional assumptions are given, before the charging service model and its *QoS* and *QoE* metrics are defined. The chapter further describes the two proposed charging service allocation mechanisms, one using queuing networks and one based on probabilistic network protocols. The evaluation is carried out on the IEEE 906 low-voltage test feeder in co-simulation with realistic *EV* driving demands extracted from a mobility survey in Germany. Afterward, possible application scopes of the proposed methods are discussed as well as remaining technical and legal challenges.

Chapter 4 investigates the smart grid application for *flexibility provision-as-a-service* to distributed *EMSs* as an offline planning problem. First, the local optimization of an *EMS* within a private household is modeled by a *MILP*, which is capable of scheduling external flexibility requests. This flexibility potential is aggregated to a flexibility pool. For disaggregation of a flexibility request to the single *EMSs*, a linear heuristic with a single objective function and a multi-objective meta heuristic are proposed. Both methods are evaluated in terms of their performance, service quality, and fairness. Afterward, a possible application scope is discussed together with technical challenges.

Finally, the main contributions of this thesis are summarized in Chapter 5 with a discussion of the limitations of the thesis results.

CHAPTER 2

Background and Related Work

This chapter provides fundamental background on power systems in Section 2.1, including briefly the history, common power grid topologies, and main aspects of power grid operation. Second, Section 2.2 discusses the terms *Quality of Service*, *Quality of Experience*, and different concepts of fairness, mainly coming from the computer-networking domain. In the following, the smart grid applications under investigation are analyzed in Sections 2.3 and 2.4, including the technical constraints and controllability, as well as a literature overview on related work in these two domains. Afterwards, the main assumptions for the thesis are stated in Section 2.5, before the chapter concludes in Section 2.6

2.1 Electrical Power System

Since the end of the 19th century, electricity is used worldwide for commercial purposes. However, electricity generation does not always reside at the location of its consumption, therefore the power grid infrastructure started to grow. In the beginning, single generators were serving nearby loads using short *DC* power distribution systems, which nowadays would be called (*DC*-) microgrids. With increasing demand, more current is transmitted over the *DC* lines, which in turn causes higher power losses according to Ohm's law. As consequence, system engineers switched to *AC* power systems that by design decouple the power distribution from the power transmission system, which is operated at much higher voltage levels to reduce the current flow and therefore transmission losses. *AC*-based power systems became the world standard, typically operated with a frequency of either 50 Hz or 60 Hz. With an increasing number of customers and their geographical distribution, smaller *AC* microgrids merged to form synchronous interconnected power systems. The Western European power system for instance spans from Portugal to Turkey, and from Algeria to Denmark. Other European regions, such as the Island of Ireland, the United Kingdom, Scandinavia, and the Baltic States operate their own synchronized power system, which is interconnected with other power systems usually via *DC* lines to decouple the frequency. To keep the frequency stable, power supply and demand must match in the synchronized interconnected power system. Nowadays, this is ensured by energy market mechanisms with different time scales from long-term contracts (several years) to day-ahead and intraday markets, where energy is traded in products of 15-minute duration. Power reserve markets allow for balancing the power grid, in case the traded energy and the actual energy flow mismatch at the moment of delivery. Including all the involved

stakeholders, the market design, and control mechanisms, the power system is a complex system that requires many entities to interact with each other.

The current power system is composed of power transmission and distribution grids as shown in Figure 2.1. Typically, huge power generators (*e. g.*, nuclear power plants or large offshore wind farms) are connected to extra-high voltage levels at the transmission system, which transports the power over long distances to the consumers. The power distribution grid is operated at high voltage (for big customers with high demand), medium voltage (for smaller industries), and low voltage levels (for household connections). The trend towards *DRES* leads to many small generators at low and medium voltage levels being connected to the power distribution grid, which originally was built to transport electricity from large generators to distributed consumers. In some situations, this even can cause a reverse power flow from the lower voltage distribution levels to higher ones, which makes operating the not-fully monitored power distribution grid a more complex task. In Europe, the essential characteristics of the supply voltage at grid connection points are given in EN 50160, which defines the frequency, amplitude, waveform, and symmetry of three-phase voltages. The frequency must reside 99.5 % of a week within 49.5 Hz and 50.5 Hz and is not allowed to pass 47 Hz or 52 Hz. Similarly, 10-minute average voltage values must reside within ±10 % of the nominal voltage. The standard further specifies supply voltage unbalance lower than 2 % and total harmonic distortion limits for system frequencies of higher-order harmonics. Besides these power quality aspects, the power grid service for the customer in terms of reliability is quantitatively measured by the *System Average Interruption Duration Index* (*SAIDI*) which asses the mean power system interruption per customer connection over one year. The *SAIDI* and other reliability metrics that capture the interruption frequency and load-normalized interruption measurements are defined in IEEE 1366.

Different power grid topologies exist, which differ in terms of structure, reliability, and operation characteristics.

- Within **radial power networks**, one substation supplies all connected customers via a radial tree topology, where the substation is the root node of the tree. Hence, there exists only one single path from the supplying node (transformer) to each customer. This topology is simple and cheap to build, and faults can be detected easily. However, in case of a line failure, all customers connected via this line are affected.

- In a **ring network**, always two paths between the supplying substation and each customer exist, whereby the connections are organized as a ring. Such a ring network can also span different voltage levels, *e. g.*, where the ring is closed via a second substation that interconnects sub-grids via the upper voltage level. Reliability drastically improves with the ring topology, because in case of a single line or substation failure, all customers are still served, potentially with reduced power quality. For fault-tolerance reasons, *DSOs* typically operate their low-voltage power grids following the $n-1$ criterion, where every customer is supplied even in case one substation or a particular line fails. Therefore, low-voltage networks are often built with ring topology but are operated with an open connection like a radial network. In this way, faults can be identified fast and supply can be re-established via this redundancy.

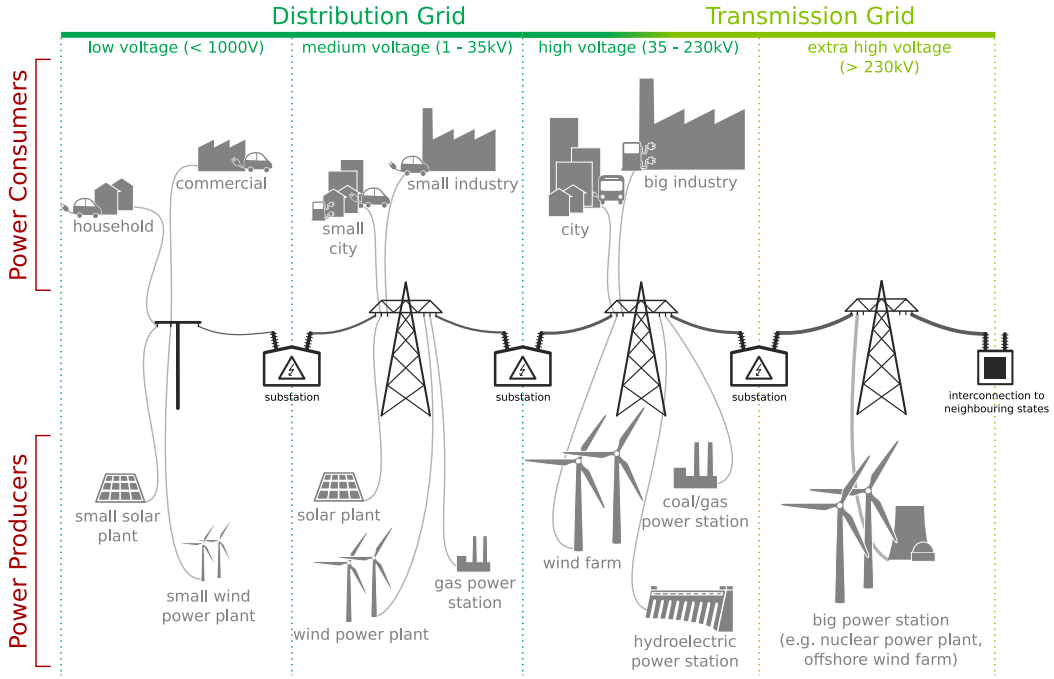


Figure 2.1: Power system from an architectural perspective including the power transmission and distribution grid. (Figure made by Philipp Danner CC BY SA 4.0)

- **Meshed networks** are the most general classification, where at least one connection between the customer and the supplying substation must exist, but the network can also contain several cross-connections to improve reliability. This topology is the most expensive one, and it is hard to track power system asset failures. Therefore, meshed networks are typically only used for well-monitored medium to extra high voltage networks, where failure safety is required.

2.1.1 Flexibility in Power Systems

The power grid is historically operated with the paradigm *supply follows demand*, where central generators produce as much power as loads consume at any moment in time. However, with increasing volatile renewable generation, this is not possible anymore and the demand side needs to adapt to the generation. This leads to a paradigm shift towards *demand follows supply*, which involves *demand-side management* mechanisms. In this context flexibility on the customer and generation side plays an increasingly important role. In literature, there exist various, slightly different definitions of operational flexibility in power systems depending on the involved stakeholders [69, 111, 168, 178], however mostly the term flexibility relates to the variability and uncertainty of power generation and demand. Degefa et al. [40] express the idea behind flexibility in power systems very well in an abstract manner:

> *"The ability of power system operation, power system assets, loads, energy storage assets and generators, to change or modify their routine operation for a limited duration, and responding to external service request signals[, without inducing unplanned disruptions]."*

Thereby, the authors explicitly exclude load shedding, *e. g.*, reducing peak load without shifting the energy consumption to a different time, and *DRES* curtailment from being a flexibility, which from a technical point of view is quite limiting. For example, this excludes flexibility originating from sector coupling, where a heating system can switch the energy source from electrical power to gas in order to provide flexibility in the power system, or a *PHEV* uses the combustion engine to cover the next trip distance when the battery is not sufficiently charged. However, the idea of excluding flexibility measures that disrupt routine operations is in line with the focus of this thesis to provide service quality for smart grid applications. Therefore, this thesis does not fully follow their definition and neglects the aspect of *inducing unplanned disruptions*. Nevertheless, the proposed methods still aim to not cause unplanned disruptions to guarantee high *QoS* while providing flexibility to the power grid.

A *flexibility resource* is an asset that can provide flexibility. Flexibility resources can be further classified according to their availability [168]:

- *Potential flexibility resource* is a flexibility resource that exists physically but is neither monitored nor controllable, *e. g.*, nuclear fusion, which is physically known but not yet under control.
- *Actual flexibility resource* is a potential flexibility resource that is monitored and controllable, and therefore can technically be used, *e. g.*, public *EV* charging, which is usually monitored and controllable, but charging operators focus on charging and not on the flexibility potential from the *EVs*.
- *Flexibility reserve* is an actual flexibility resource that can be utilized from an economical point of view, *e. g.*, a *BESS* or *EV* charging at private households, whose main purpose is to provide flexibility for local optimization.
- *Market-available flexibility reserve* is a flexibility reserve that can be procured from power and *AS* markets, *e. g.*, a gas power plant that adopts to market signals.

Only *market-available flexibility reserve* and *flexibility reserve* are interesting because power grid operations can only interact via the market and smart grid applications will only utilize flexibility if it is economically beneficial for the power grid operator and the flexibility provider.

Flexibility can be quantified by the so-called *flexibility trinity* that specifies the *ramping capacity*, *power capacity*, and *energy capacity* [113, 168]. Thereby, the energy capacity is the integral of the power capacity, which in turn is the integral of the ramping capacity. The sign of the amplitude of the power capacity defines the *direction* of the flexibility, which can be either *positive* or *negative*. *Positive* flexibility refers to decreasing from the routine

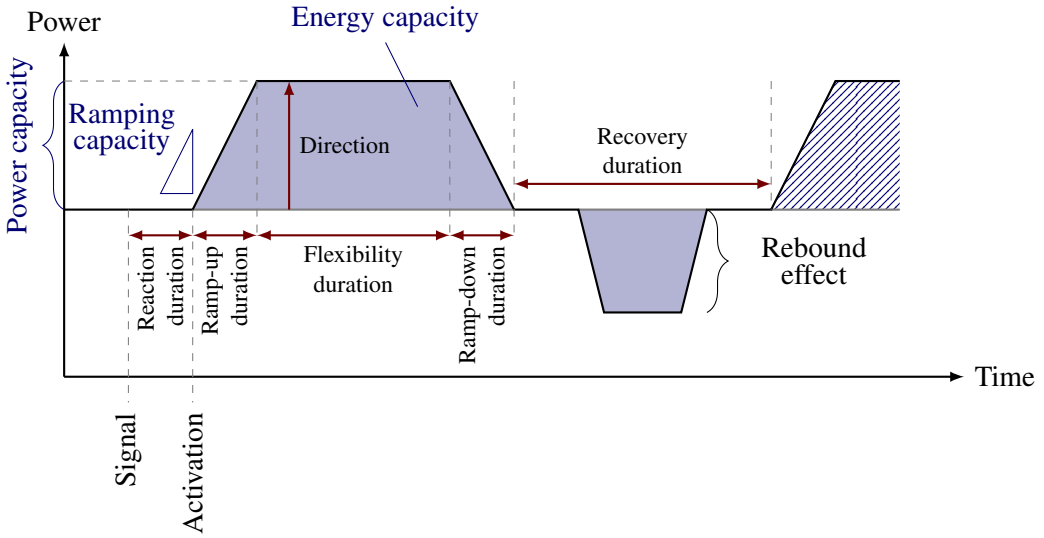


Figure 2.2: Important characteristics of flexibility resources (power capacity, ramping capacity, energy capacity, rebound effect and timing aspects) according to [40].

consumption profile, *e. g.*, by discharging an *Energy Storage System* (*ESS*), whereas *negative* flexibility refers to increasing from the routine consumption profile, *e. g.*, by charging an *ESS*. Note that power capacity, and therefore ramping and energy capacity, is not limited to real power P in W (Wh^{-1} and Wh), but can also be reactive power Q in VAr. In this case, energy capacity can be neglected, because reactive power is supplied by inductive and capacitive effects and does not involve any energy storage.

In addition to the main three characteristics defined by the flexibility trinity, a flexibility resource can have additional characteristics depicted in Figure 2.2. After receiving an activation signal, there might be a *reaction duration* in which the flexibility resource stays with the routine profile. After that, the *ramp-up duration* starts, where the power capacity increases up to a maximum of the power capacity, limited by the ramping capacity of the flexibility resource. Note, that this ramp rate is not necessarily linear as depicted in Figure 2.2. The energy capacity limits the *flexibility duration*, after which the *ramp-down duration* starts. After flexibility provision, the flexibility resource might require a *recovery duration* before providing its service again. During this phase, energy bounded flexibility can cause a *rebound effect*, *e. g.*, recharge an empty battery for the next use. Note that with reactive power, no rebound effect is expected, because its energy capacity is not bounded.

2.1.2 Ancillary Services

The *European Network of Transmission System Operators for Electricity* (*ENTSO-E*) defines *ASs* to refer *"to a range of functions which TSOs contract so that they can guarantee system security. These include black start capability (the ability to restart a grid following a blackout); frequency response (to maintain system frequency with automatic and very fast responses); fast reserve (which can provide additional energy when needed); the provision*

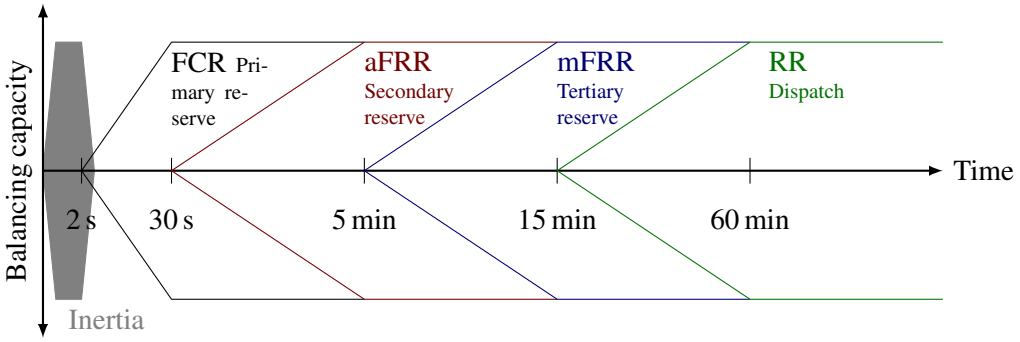


Figure 2.3: Balancing reserves and their respective time horizons: Inertia, *FCR*, *aFRR*, *mFRR*, and *RR*.

of reactive power and various other services" [48]. However, from *DSOs*' perspective, there are additional *ASs* that are required for the operation of the power distribution grid, *e. g.*, congestion management, voltage control, and harmonic mitigation. These *ASs* can be provided with flexibility from (distributed) flexibility resources supplied by smart converters in the future, as investigated in the European research project EASY-RES [129], where ramping, power, and energy capacity are defined according to the power grid requirements.

Balancing services that offer *frequency response* and *fast reserves* are in general categorized by the time horizon of their availability, as depicted in Figure 2.3.

- *Inertial response*, also known as *frequency response reserve*, is traditionally provided by rotating masses, where a fast dip or increase of the power systems' frequency results in a direct response of the rotating masses, which slow down or speed up to reduce the rate of frequency change. In future power systems without big mechanical generation units, virtual inertia must be provided by converter-based renewable generations that are equipped with fast *ESS*, *e. g.*, super-capacitors [126].
- *Frequency Containment Reserves* (*FCR*), which is also known as *primary reserve*, must be available at full capacity latest 30 seconds after the frequency event occurs and must cover a duration of up to 15 minutes. The task of this balancing service is to keep the power system's frequency within the desired bandwidth. This service may be provided by *EV* charging in the future [114, 116, 120].
- The *automatic Frequency Restoration Reserves* (*aFRR*), also known as *secondary reserve*, is automatically activated and replaces the *FCR* gradually after 30 seconds. Its full capacity is required to be provided at the latest five minutes after the frequency event. Traditionally, hydro-storage and gas power plants are the main contributors to this reserve, however, *Virtual Power Plants* (*VPPs*) consisting of bio-gas power plants, combined heat and power plants, and *BESSs* can provide this service in the future [63, 117, 118, 172].
- In contrast to *aFRR*, the *manual Frequency Restoration Reserves* (*mFRR*), also known as *tertiary reserve*, is manually or semi-automatically requested by the system operators. Its full capacity must be available latest after 15 minutes.

- Finally, *Reserve Replacement* (*RR*), also known as *dispatch*, is required to take over from the former frequency reserves at the latest after one hour and re-schedules generation units to meet the new demand-supply balance. In the future, this can also be achieved by re-planning customers' flexibility.

FCR, *aFRR*, and *mFRR* are traded on the European energy markets with both positive and negative quantities, defined similarly to positive and negative flexibility. In Germany, the trading window of *aFRR* is equal to four hours, in which the traded energy reserve must be provided. In most countries, there are two different types of remuneration for the balancing service. First, for keeping the capacity available (capacity price) and, second, for actually activating the balancing service (energy price) in case of a frequency event.

2.2 Notions of Quality of Service, Quality of Experience, and Fairness

Quality of Service (*QoS*) is the measurement of the overall performance of a service and was initially introduced for telecommunication services by *International Telecommunication Union* (*ITU*) in 1994. In the current valid recommendation ITU E.800, *QoS* is defined as the *"Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service"* [86]. This definition implies that characteristics of the service, which can be measured quantitatively or qualitatively, need to match the users' requirements towards the service, and hence involve the user to decide about the service quality. The *QoS* criteria can be differentiated between network performance (intrinsic *QoS*) and non-network performance, where the former measures network parameters, such as packet loss rate, bit error rate, average bit rate, throughput, transmission delay, availability, and *jitter*. Non-network performance can include service mean repair time, provisioning time, and billing process. The *ITU* further specifies four viewpoints of *QoS*, also depicted in Figure 2.4:

(i) The *customer's* QoS *requirements* can be specified in non-technical terms and describe the resulting end-to-end service quality a customer may expect from the service, without knowledge of the internal network design.

(ii) QoS *offered by the service provider* expresses the values of the *QoS* parameter that the service provider plans to deliver. This may include non-technical definitions, as well as technical measures that relate to intrinsic *QoS* parameters that can be part of service level agreements.

(iii) QoS *achieved by the service provider* is the statement of what *QoS* parameters are finally delivered to the customer measured in the offered *QoS* parameters.

(iv) QoS *perceived by the customer* finally is a statement of the customer's satisfaction with the service typically measured in non-technical terms, such as customer surveys.

To control the *QoS* parameter, network standards define *QoS* classes, sometimes also called *Class of Service* (*CoS*), where *"services belonging to the same class are described by a*

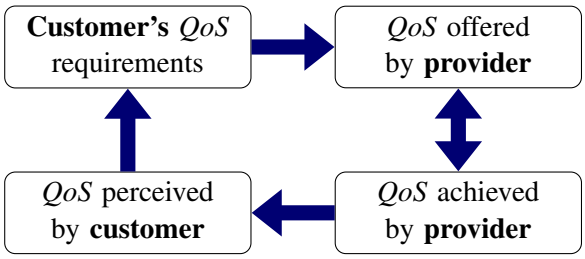


Figure 2.4: The four viewpoints of *QoS* according to ITU G.1000. On the provider side, (non-)achieved *QoS* parameter may directly influence the *QoS* offers. Furthermore, the perceived *QoS* by the customer may influence their requirements, *e. g.*, perceived *QoS* may become expectations as part of new requirements.

specific set of parameters, which can be expressed qualitatively or quantitatively" [158]. These parameters need to be met for high performance of the application, and relate to traffic requirements and also intrinsic network parameters. Typically, service classes in networking are categorized by *elastic/non-interactive* (*e. g.*, file download), *elastic/interactive* (*e. g.*, web browsing), *non-elastic/non-interactive* (*e. g.*, video streaming), and *non-elastic/interactive* (*e. g.*, voice or video conferences). Some applications may additionally specify the required network bandwidth and symmetry of communication. The different standardization entities (*ITU*, *IETF*, *IEEE*, and *3GPP*) define slightly different *QoS* classes for their protocols that arise from specific traffic requirements and characteristics. These definitions do not always map straight forward to each other, however, there exist efforts for mapping between them to improve interoperability [158].

Quality of Experience (*QoE*) is defined as *"The degree of delight or annoyance of the user of an application or service"* [24]. *QoE* (assessed *QoS*) relates (not necessarily linear) to *QoS* parameters and additionally integrates the personality and the current state of the user. According to [74], a generic *QoE* model Q, as in Equation (2.1), can be defined as a function of *QoS* parameters and other influencing factors x to the *QoE* value y.

$$Q : x \rightarrow y = Q(x) \in [L;H] \tag{2.1}$$

L and H are the lower and upper boundary of experience values, respectively, *e. g.*, $L = 1$ (bad quality) and $H = 5$ (high quality), which is typically used in the 5-point *Mean Opinion Score* (*MOS*). These *QoE* models are often derived from user surveys, but can also be metrics on the *QoS* parameters, *e. g.*, median or percentiles.

The following two examples will help to understand the difference between *QoS* and *QoE* in the *EV* charging domain.

- An *EV* that can fully recharge its battery during the available charging time, between arrival and departure, receives a high *QoS* because the charging service succeeds. Logically, if a charging process cannot finish before departure, the *QoS* is lower. Hence, the perceived *QoS* is coupled with the energy requirement of the charging process, which relates to intrinsic *QoS* metrics that measure the network performance, such as the charging rate. Now, considering the context of the *EV* user, the *QoE*

may differ with respect to his next trip. If the *EV* has recharged enough energy to reach the next destination (independent of whether the charging process has finished or not), the *QoE* is obviously higher than if an additional charging stop is required during the next trip. Although the driving distance of the next (unplanned) trip does not affect the perceived *QoS* of the charging process, it may influence the *QoE*.

- In a second example, two *EVs* connect to charging stations that share the same power grid resource, where both *EVs* have equal charging requirements (required amount of energy and available charging time) and hence are in the same *QoS* class. In case the charging rates need to be reduced due to power grid limitations, both charging processes receive the same lower charging rate, which equally reduces the intrinsic *QoS* parameter for both. Note, with this approach proportional *QoS* fairness on the charging rate is achieved. However, if the two *EV* users have different grid connection capacities (or maximum charging capacity specifications of their wallboxes), their *QoE* may differ. Assuming the charging power is limited to 11 kW per *EV*, one user with 22 kW peak charging power reaches only 50 % of utilization, whereas the other user may charge with his full peak charging power of 11 kW. Obviously, only the charging process of the first user lags behind expectation, which results in lower *QoE*, regardless the fact whether the actual charging service is finished in time or not (perceived *QoS*). This example shows that a fair *QoS* distribution must not necessarily result in fair *QoE*.

Both *QoS* and *QoE* respect the overall measure of user satisfaction but do not express how fair the achieved metric scores are distributed among the participating users. Especially if the maximum quality level cannot be achieved, *e. g.*, due to power grid limitations in the case of *EV* charging, every user should receive a fair share of the limited service. It is important to note that fairness can be defined on all different levels of *QoS*. However, a proportionally fair intrinsic *QoS* parameter may not result in fair perceived *QoS*, and *vice versa*. On the other hand, increasing fairness may reduce the overall achieved mean *QoS* and *QoE* scores. Hence, resource allocation algorithms need to balance between *QoS* and *QoE* scores, and fairness.

In communication networks, the term *fairness* relates to fair access to a shared communication medium, such as the air in wireless communication. Access to a shared medium is typically modeled with queues and a scheduler. In General, application systems for queues can be classified into *flow-based* and *service-based* systems [105]. Flow-based queues handle a stream of jobs for each customer and fairness guarantees that competing flows, that share the same resource, are served equally fair at any point in time. For example, network routers only see flows of packets assigned to different customers, and fairness is defined on the network access level because routers do not have insight into the purpose of the packet stream (hence do not know to which service they belong). On the contrary, service-based systems only deal with a single job for each customer and, hence, the fairness objective is limited to the completion of that single job, *e. g.*, in terms of service delay. These queues typically apply to the application layer, *e. g.*, the download of a remote file via the Internet.

Many definitions for fairness of resource allocations exist for the networking and computer science context, where the most prominent are outlined in the following.

- One of the first attempts for quantifying fairness for server load sharing was introduced by Yung-Terng Wang and Morris [173] as the *Q-factor* (quality of load sharing factor). This factor combines the overall efficiency and fairness of allocation algorithms and compares the mean job response time against the mean response time under a global *first come first served* scheduler, which implements both their defined goals of not discriminating jobs from particular sources and avoiding servers to be idle.

- The *max-min fairness* is a well-known principle in computer networking [87]. A resource allocation is said to be max-min fair if the bit rate for one flow cannot be increased without decreasing the bit rate of another flow with a smaller bit rate. The *progressive filling algorithm*, also called the *water-filling algorithm*, uniquely defines the max-min fairness. Max-min fairness is also known as *Generalized Processor Sharing* (*GPS*), which is known to be unimplementable in the networking domain, because in packet switching, the minimum size of a packet is fixed and, therefore, only approximations of the fully equal allocation exits [21, 42, 59, 152].

- The *relative and absolute fairness bound*, used and analyzed among others in [59, 62, 92, 174], provide an estimation of how fair a system resource allocation is distributed among different flows. Similar to the Q-factor, the absolute fairness bound compares the resource allocation of a scheduling policy P with a baseline scheduler, here the *GPS*. Since *GPS* is hard to estimate, researchers tend to use the relative fairness, which simply compares different flows. The absolute and relative fairness bound finally calculates the maximum of the absolute and relative fairness of all combinations of flows and time intervals.

- *Proportional fairness* [91] is one of the most commonly used fairness criteria in literature, where the assigned resource ratio to each flow is inversely proportional to its required capacity. Note, even though proportional fairness was initially designed for flow-based systems it can also be applied to service-based queuing applications if the bit rate is not allocated based on the number of flows or their bit rate requirements, but on other service aware parameters, *e. g.*, the file size of a file to be downloaded.

- The *index of fairness* [88] (also known as *Jain's index*) quantifies fairness of any resource allocation to n flows and ranges from $\frac{1}{n}$ (unfair) to 1 (most fair). The proposed fairness index for an allocation $B = \{b_1, \ldots, b_n\}$ is calculated as in Equation (2.2), where $c(B)$ is the *relative standard deviation* $c = \frac{\sigma}{\mu}$ of all values $b_1, \ldots, b_n$.

$$\mathcal{J}(b_1, \ldots, b_n) = \frac{(\sum_{i=1}^{n} b_i)^2}{n \cdot \sum_{i=1}^{n} b_i^2} = \frac{\mu(B)^2}{\mu(B^2)} = \frac{1}{1 + c(B)} \quad , \tag{2.2}$$

  It can be seen that Jain's fairness index depends on the input value level [74], hence allocations with the same standard deviation have a better fairness value when its mean value is greater. This fact must be considered when analyzing fairness independent of the allocation values. Nevertheless, the goal is to reach both high absolute allocation values and high fairness.

- The *fairness index* by Hoßfeld et al. [73, 74] defines the fairness decoupled from the achieved allocation values and is originally designed to measure the *QoE* fairness. Its calculation is given in Equation (2.3), where H and L are the upper and lower bound of allocation values.

$$F(b_1,\ldots,b_n) = 1 - \frac{\sigma(b_1,\ldots,b_n)}{\sigma_{\max}} = 1 - \frac{2\sigma(b_1,\ldots,b_n)}{H-L} \tag{2.3}$$

  Note that due to the decoupling from the allocation values, a fair allocation does not imply overall good allocation values, *e. g.*, even low *QoS* parameters can yield a fair allocation if their standard deviation is small.

In addition to the aforementioned fairness definitions, several other notions of fairness exist, *e. g.*, *balanced fairness* [22] or *worst-case fairness* [21]. In the field of game theory, fairness was introduced as part of the utility function to model human fairness treatment and social welfare [134].

The difference between *flow-based* and *service-based* fairness in the energy domain can be seen in the example of *PV* curtailment. On the one hand, flow-based fairness defines that all *PVs* are curtailed according to the currently injecting power at any point in time [57, 58, 95, 162], whereas service-based fairness would consider the total losses due to curtailed *PV* generation throughout a year [160]. Furthermore, literature uses different definitions of fairness, among others egalitarian (equal absolute curtailment for all *PV* systems) [57], proportional fairness [57, 58, 95, 162], or min-max fairness [57].

In a different application area, authors of [27] propose a fair power allocation for air conditioners in the smart grid, where the power consumption is indirectly controlled by allocating thermostat settings in each time slot. In this way, ambient temperature and the amount of power required for the same temperature reduction is decoupled and fairness is defined on the *QoS* level of air conditioning (flow-based).

2.3 Electric Vehicle Charging

The general term *Electric Vehicle* (*EV*) refers to vehicles that use electric repulsion for driving. However, several sub-categories with varying configurations exist.

- *Battery Electric Vehicles* (*BEVs*) are *pure* or *all-electric vehicles*, which contain battery storage that can only be charged via a charging outlet from an external electricity source. The battery capacity and maximum charging rate are therefore quite large to be able to drive long distances and re-charge the battery in a fast manner.
- *Hybrid Electric Vehicles* (*HEVs*) are not limited to electricity as the energy source for repulsion. Some variants of *EVs* additionally contain an internal combustion engine and the electric motor is mainly used for fuel saving, *e. g.*, the electrical motor is used for acceleration from stand while the combustion engine is operated at a more efficient rotational-speed range during driving. Others have only one electrical motor, but use a combustion engine or hydrogen fuel cells for electricity generation and a

Table 2.1: Charging modes and the corresponding maximum charging current and charging power according to IEC 61851-1

Mode	Type	Current [A]	Power [kW]	
			single-phase	three-phase
1 (slow charging)	*AC*	16	3.7	11
2 (slow charging)	*AC*	32	7.4	22
3 (slow and quick charging)	*AC*	63	14.5	43.5
4 (fast charging)	*DC*	200	80	

small battery or supercapacitor as an energy buffer. Sometimes, the internal battery is only charged by recuperative braking.

- *Plugin Electric Vehicles* (*PEVs*) have external charging outlets to charge their internal batteries. This category includes all *BEVs* and some variants of *HEV*, called *Plugin Hybrid Electric Vehicles* (*PHEVs*).

Because the focus of the smart grid applications is on power grid-connected *EVs*, this thesis refers to *PEV* when using the acronym *EV* in the remainder of this document. Furthermore, the thesis targets *QoS* and fairness for *EV* charging, where user involvement is required, hence *EVs* are limited to electrical passenger cars. Nevertheless, the proposed methods might also apply to other types of *EVs*, like trucks, buses, or bikes.

2.3.1 Charging Modes and Communication Protocols

All *PEVs* can be charged from the electrical power grid at charging stations that support different charging modes, also called *Grid to Vehicle* (*G2V*). Some modern *EVs* support *Vehicle to X* (*V2X*) capabilities, including *Vehicle to Grid* (*V2G*), *Vehicle to Home* (*V2H*), and *Vehicle to Load* (*V2L*). This thesis however focuses on *charging-as-a-service*, hence discharging of an *EV* is not discussed. Table 2.1 shows the charging specifications defined in IEC 61851-1, where charging modes 1 and 2 apply to regular *AC* electricity sockets (mode 2 with additional monitoring and protection function). As most common for *AC* charging, mode 3 requires a permanently installed specific charging socket, which usually can be found at public charging stations or private wallboxes. Mode 4 uses a *DC* charging socket and requires an *AC*/*DC* converter at the charging station, which is uncommon for private wallboxes at home. In addition to the charging modes defined by IEC 61851-1, upcoming *ultra-fast* charging stations with *ChaoJi* connectors support up to 600 A and 1500 V on the *DC* side, reaching charging powers of more than 900 kW [159].

The charging needs of *EVs* are typically divided into *home charging*, *workplace charging*, and *public charging*, where the latter mostly relates to *DC* (ultra-)fast charging. Charging facilities at home locations and workplaces are mostly *AC* charging stations with limited charging power, typically not more than 22 kW. *Home charging* refers to the use case where people plug their *EVs* after arriving home in the evening and require a fully charged battery until departure, usually the next morning. This offers large flexibility at which

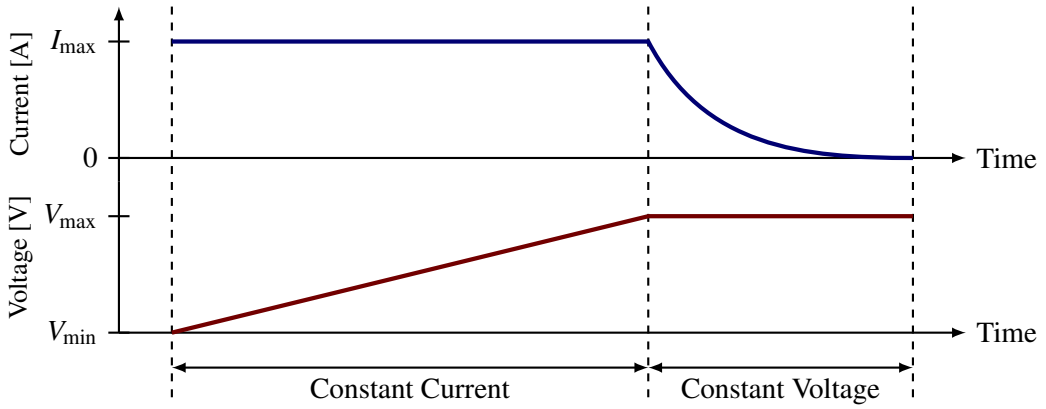


Figure 2.5: Typical *CCCV* charging behavior of *lithium-ion* batteries.

time the *EV* can be charged. However, trips that occur shortly after arriving home require immediate charging, which must be respected by smart charging control algorithms to keep *EV* drivers happy. Similar flexibility is offered by *workplace* charging, where *EVs* park for approximately eight hours at parking lots during the day. Because charging can be seen as an extra service to the employees with potentially reduced charging fees, there is no strict requirement to fully charge the connected *EVs*, hence charging control has more flexibility and can prioritize for example the company fleet. *Public charging* goes along with high charging power and energy demand because *EV* drivers want to resume their trip with maximum charged energy during a typically short stay. The public (ultra-)fast charging has the lowest flexibility in shifting charging processes. However, due to the rated high charging power, *AS* support can be realized, *e. g.*, reactive power supply or voltage regulation. Uni-directional *EV* charging (only *G2V*) can participate in the *AS* market to provide positive and negative reserve power by reserving aggregated charging flexibility from an *EV* fleet, as proposed in [148, 149].

Besides charging equipment characteristics, the *EV* itself can have additional charging limitations originating from its *Battery Management System* (*BMS*) or the battery chemistry. *Lithium-ion* batteries for example, which are usually used in *EVs*, are typically charged with the so-called *Constant-Current Constant-Voltage* (*CCCV*) charging pattern, as depicted in Figure 2.5. First, the battery charger supplies constant current, while the internal cell voltage continuously increases during charging until the maximum cell voltage is reached. For *lithium-ion* batteries in *EVs*, this is usually above 80 % of the battery *SoC* in case of low charging power of around 22 kW [147], and even above 90 % in case of 11 kW charging power [68]. Afterward, the voltage is kept constant and the supply current decreases until the battery is fully charged. This phase is called the *saturation phase* of the battery. Additionally, the *BMS* may restrict the usable battery capacity between 10 % to 90 % of the factory storage capacity to avoid deep discharge and overcharging, which both can harm the battery *State of Health* (*SoH*). Ambient temperature, cooling technology, and battery core temperature may further influence the actual charging power of an *EV*. Because this thesis aims to provide *charging-as-a-service* from the charging station and power grid point of view, *EV*-specific internal *EV* restrictions are not considered. The used charging service

Table 2.2: *PWM* signal and corresponding maximum charging current per phase according to IEC 61851-1.

	PWM [%]				I_{max} [A]
		x	$<$	3	No charging allowed
3	$\leq$	x	$\leq$	7	High-level communication, *e. g.*, ISO 15118
7	$<$	x	$<$	8	No charging allowed
8	$\leq$	x	$<$	10	$I_{max} = 6\,\text{A}$
10	$\leq$	x	$\leq$	85	$I_{max} = x \cdot 100 \cdot 0.6\,\text{A}$
85	$<$	x	$\leq$	96	$I_{max} = (x - 64\,\%) \cdot 100 \cdot 2.5\,\text{A}$
96	$<$	x	$\leq$	97	$I_{max} = 80\,\text{A}$
97	$<$	x			No charging allowed

parameters include *arrival time*, *departure time*, and *required energy* to be charged, and are explained in more detail in Section 3.2.

Standards such as the IEC 62196 specify charging plugs and IEC 61851-1 specifies the communication with the *EVs* via these connectors. The *AC* Type 2 charging connector, specified in IEC 62196-2, applies for charging modes 1 to 3 and is the most widespread *AC* charging plug in Europe. The connector has one or three-phase conductors, a neutral and *Protective Earth* (*PE*) conductor, a *Proximity-Pilot* (*PP*), and a *Control Pilot* (*CP*), which together are used to communicate with the *EV*. In *basic communication* mode, the charging station provides a voltage of +12 V between *CP* and *PE*. If an *EV* is connected, a *PWM* signal of 1 kHz is provided by the charging station, which specifies the maximum charging current for the *EV* according to Table 2.2. On the other side, the *EV* can modify the resistance between *CP* and *PE* to share its status with the charging station. Different resistances change the voltage amplitude and can encode the following six states: *standby* (+12 V), *vehicle detected* (+9 ± 1 V), *ready for charging* (+6 ± 1 V), *with ventilation* (+3 ± 1 V), *no power* (0 V) and *error* (−12 V). A cable can further limit the maximum current capacity by defining a constant resistance between *PP* and *PE*, which should reflect its conductor diameter.

On top of the *basic communication* between the charging station and the *EV* via *PWM*, there exist more advanced protocols, *e. g.*, bi-directional *high-level communication* specified by ISO 15118. This standard supports advanced features that support use cases to identify, authenticate and authorize *EVs*, start, stop, and control charging processes, and provide reactive power compensation or *V2G* support. The bi-directional communication is based on the *Open Systems Interconnection* (*OSI*) model, where the *Internet Protocol* (*IP*) based communication is either wired (via *Power-Line Communication* (*PLC*) over the *CP* conductor according to *HomePlug Green PHY*) or wireless (IEEE 802.11n). Messages are exchanged in *Extensible Markup Language* (*XML*) format with *Efficient XML Interchange* (*EXI*) encoding. For controlling an *EV* charging process the most important feature in ISO 15118-2 is the *charging profile* negotiation mechanism between the *EV* and the charging station. Thereby, the *EV* request possible *charging parameters* that define charging limitations, *e. g.*, from the power grid, and selects a suitable *charging profile* via a power

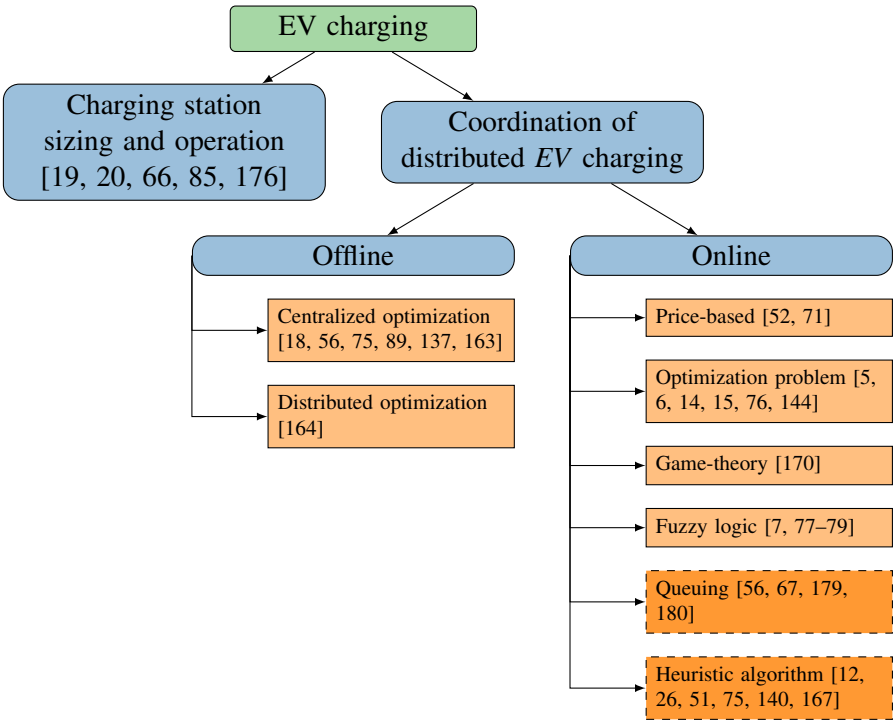


Figure 2.6: Categorization of literature in the domain of *EV* charging that considers *QoS* and fairness aspects. The proposed methods in Chapter 3 fit to the two highlighted online *EV* coordination categories.

delivery request. In case the limitations at the charging station change, a re-negotiation is initiated. Furthermore, the *EV* can transmit its planned departure time and energy requirement along with the charging parameter request, as well as *BMS*-specific parameters, *e. g.*, at which *SoC* fast charging is not possible anymore. During charging, the charging station can control the *EV* charging process by modifying the *target power* values. This may also include reactive power in the case of *AC* charging, as introduced in the extension of ISO 15118-20. Overall, this allows power grid operators to request reactive power from the *EV* charging processes to support the power grid in terms of voltage control and reactive power balance.

2.3.2 Literature Overview

Many publications in literature deal with smart charging of *EVs* targeting the mitigation of power grid issues as their main objective [8, 9, 11, 28, 29, 38, 41, 94, 109, 115, 142]. However, the main focus of this thesis is to highlight the impact of smart *EV* charging on the users' expectations, hence the main objective is to make *EV* drivers happy under given power grid constraints. Literature on *EV* charging problems that optimize for or at least discusses obtained results with respect to *QoS* and fairness can be categorized as given in Figure 2.6.

One research direction, which considers service quality aspects for the *EV* driver, targets charging station sizing and operation of smart charging stations with auxiliary *ESSs*. Thereby, the main concern is the probability that an *EV* is blocked at a charging station [19, 20, 176]. In particular, authors of [20, 176] try to improve the *QoS* by introducing different *QoS* classes, where higher service levels correspond to higher payment by the *EV* driver, whereas authors of [19] optimize the revenue of the charging station operator by extracting the optimal size of the charging station and its local battery storage using a continuous time Markov chain that is used to determine the blocking probability. A different definition of *QoS* for public smart charging stations is considered in [66], where *QoS* includes the continuity of power supply, the delay of *EVs* queuing for the charging service, and the overall charging time. Again, the main objective is battery sizing of local storage systems but additionally, all *EVs* can supply energy back to the charging station. In [85], photovoltaic and battery sizes are optimized for business charging, where *EV* availability highly correlates with photovoltaic generation. Their optimization considers *QoS* as the ratio of charging energy delivered to charging energy demanded, where a *fairness factor* influences the charging rate of each *EV* based on its *SoC*.

The second research direction deals with the coordination of distributed *EV* charging, which can be divided into offline planning problems and (real-time) online algorithms. The main assumption of offline planning, sometimes called *day-ahead scheduling*, is that future charging processes are known in advance or at least can be forecast to a reasonable extent. Most of these papers perform day-ahead optimization using an optimization problem, *e. g.*, *MILP* [56, 89], *Mixed Integer Non-Linear Programming* (*MINLP*) [75], or distributed optimization with *Alternating Direction Method of Multipliers* (*ADMM*) [164]. In particular, authors of [56] use a *MILP* for day-ahead planning of charging processes at business parking lots where the optimality is determined by a weighted linear objective function that considers the fair share (social welfare on the final *SoC*), total charging cost, peak power at the grid connection, and three-phase grid imbalance. Thereby, the *QoS* for the *EV* user is measured only based on the *SoC* of the battery. Additionally, they propose a priority queuing algorithm for the online control that utilizes knowledge from the day-ahead schedule and prioritizes with respect to the same four objectives. The *MILP* problem in [89] uses *V2G* capability to plan the charging and discharging of *EVs* to flatten the duck curve. Thereby, battery degradation as the main quality metric is minimized. The *MINLP* problem in [75] aims for planning *EV* charging processes with respect to active and reactive power supply for voltage regulation, which in turn is performed online via a parallel consensus mechanism. Thereby, the goal is to optimize the profit of each *EV* with fair pricing based on the involvement level. Authors of [164] propose a distributed version of their optimization problem using *ADMM* while aiming for proportionally fair allocation of charging power considering voltage drops at the feeders.

Many methodologies are applied for online *EV* charging coordination. Some of them are discussed in more detail since Chapter 3 proposes a queuing and a probability-based algorithm as highlighted in Figure 2.6. The *QoS* in [52] is measured by a utility function on the charging rate and the demand-response charging control is adapted from congestion pricing from the networking context using a *willingness-to-pay* factor. The bidding model proposed in [71] measures the comfort of *EV* drivers only by the *SoC* at departure.

Multiple papers consider online optimization problems, sometimes with distributed solving methods [6, 14, 15] or decentralized formulations [84]. Authors of [14, 15] describe the proportionally fair distribution of available power capacity as an optimization problem that can be solved in a distributed manner via a decomposition of its dual problem. The satisfaction of an *EV* user is defined depending on the current charging power. Contrarily, the fair charging capacity allocation in [6] considers the laxity of a charging process as weight in an optimization problem. To reduce the impact of users that cheat with their departure time, they propose to integrate the user's reputation into the weights, such that people who estimate their departure time similarly well obtain the same fair share of power capacity. The optimization problem in [144] measures the charging service performance by the charged energy and additional aims to avoid battery wear as one of its objective functions. The authors further propose a heuristic to reduce the number of integer variables in the problem formulation to increase solving speed required for online operation. Besides providing real *EV* charging data [100], authors of [101, 102] propose a central *Model Predictive Control* (*MPC*) optimization problem for *EV* charging in parking garages. The objective function of their framework is configurable and contains among others fair charging rate distribution and prioritization to finish charging all *EVs*. Their approach is compared to *Earliest-Departure-First* (*EDF*), *Least-Laxity-First* (*LLF*), and a *round-robin* method, all of which perform very similarly to the analysis in Section 3.5.

Other authors apply game-theoretic concepts to minimize the cost of the charging processes per *EV* in a non-cooperative game framework [170]. They further propose a distributed generalized Nash equilibrium-seeking algorithm. Authors of [77–79] propose fuzzy logic to solve the charging power allocation problem while focusing on the *SoC* as the main *QoS* metric for the *EV* driver. The multi-objective fuzzy controller in [7] targets user requirements by the waiting time, dynamic energy prices, the battery *SoH*, and required *SoC* for the next trip, as well as local renewable generation from *PV*. Their fuzzy controller is supported by *Particle Swarm Optimization* (*PSO*) and a *Neural Network* (*NN*) to find a global optimal solution and cope with the inherent complexity of the controller.

The approach in [179] shows how available power capacity can be shared fairly using a weighted fair queuing scheduler. Their approach is based on packetization of the charging process of *EVs*, where each packet represents the permission to charge for the next time slot. These packets are queued by the single charging processes to the proposed scheduler, which computes the packet assignment weight based on the *SoC* at arrival. Demand and supply mismatch defines the available power in their solution, not considering the congestion of the underlying power grid topology. The *QoS* is measured by the ratio of delayed charging processes. In [180], the same authors compare different queuing policies from the networking context for distributing the available power to charging processes, *e. g.*, *round-robin*, *First-Come-First-Served* (*FCFS*), or *EDF*. *QoS* is again evaluated by the charging delay. More recently, authors of [67] apply weighted fair queuing to any type of deferrable loads in power systems. Thereby, they calculate the weights based on the required wattage and remaining waiting time.

In [51], an online *QoS*-aware admission control algorithm is introduced that makes use of different *QoS* classes, where *QoS* relates to the overall charging time and finishing charging faster implies a higher *QoS*. A different approach in [167] applies the *Additive Increase*

Table 2.3: Overview of *QoS* parameters for *EV* charging processes in the literature.

Reference	Blocking probability or charging delay	Charging time	(Continuous) charging power	Charged energy or *SoC*	Success rate (target *SoC* reached)	Battery degradation	Charging cost
[19, 20, 176, 179, 180]	✓						
[66]	✓	✓	✓				
[67]	✓		✓				
[7]	✓			✓			✓
[177]	✓				✓		
[51]		✓					
[84]		✓	✓				
[6, 14, 15, 52, 140, 164, 167]			✓				
[101, 102]			✓	✓	✓		
[5, 12, 56, 85, 137]				✓			
[26, 78]				✓	✓		
[144]				✓		✓	
[71]				✓			✓
[89]						✓	
[18, 75, 170]							✓

Multiplicative Decrease (*AIMD*) principle to *EV* charging, where the quality of power service is defined such that the voltage drop must be kept stable, but proportional fair charging rates must be provided to all *EVs*, regardless of their location in the grid. As a result, *QoS* is defined as instantaneous power delivery. The focus of the approach in [12] is the power quality of the power grid but implicitly *QoS* is defined in terms of energy delivered to the *EV* without differentiating the actual energy requirement of the user. Their *Transmission Control Protocol* (*TCP*)-inspired approach results in a fairer distribution than a pure power quality-aware algorithm from [11]. A *smoothed least laxity first* algorithm is introduced by [26], which improves the classical *LLF* method such that potentially oscillating charging signals are smoothed. Thereby, the algorithm aims for proportional fair one-step-ahead laxity and is evaluated with respect to the success rate of the single charging processes.

Table 2.3 summarizes the different definitions of *QoS* and *QoE* metrics that are used for charging power allocation or algorithm evaluation in the literature. Most research focuses only on one single quality metric and no paper provides a broader overview of how their charging power allocation algorithm impacts the users' expectations. Therefore, Chapter 3 introduces a comprehensive set of *QoS* and *QoE* metrics that not only measure the quality of the charging process but also how users may perceive the service. Note that in this thesis charging costs are not considered as a charging service quality metric, since *EV*

charging power allocation is considered to be a power grid service, for which no variable compensation is provided to the users. Furthermore, battery degradation is out of scope because with only uni-directional *EV* charging no large additional impact on the *SoH* can be expected apart from the anyhow planned charging process of the user.

Literature applies many notions of fairness, among others (weighted) proportional fairness on *one-step ahead laxity* [26] or *charging power* [14, 15, 67, 144, 164, 167]. Others refer to Jain's fairness index [6, 76], define a geographic fairness measure [67], or aim for social welfare [56]. This thesis measures fairness as the deviation of each quality metric between the *EV* charging processes to ensure that all receive the same fair service.

2.4 Home Energy Management Systems

Home *Energy Management Systems* (*EMSs*),[3] in the following referred to as *EMS*, are used when multiple flexible appliances are managed behind-the-meter. In most cases, the goal is to optimally utilize local flexibility provided by controllable devices to reduce the total energy cost or to automatically react to *DR* signals from the power grid. Home *EMSs* are sometimes integrated with *smart home* and *building automation*, which among others provide control of flexible appliances, occupancy detection, and sensor data for energy optimization. Furthermore, *EMSs* usually operate their control loops in the range of seconds, *e. g.*, the open source software *OpenEMS* [54] uses a default cycle loop of 1 s.

2.4.1 Flexibility and Controllability

Electric Vehicle (*EV*) charging, as discussed in the last section, is one of the appliances with the largest flexibility in private households due to comparably high energy consumption, high flexibility in time, and controllability. Flexibility characteristics and relevant protocols are already discussed in the last section.

Furthermore, *BESSs*, in the following referred to as *ESSs*, are usually installed at private households to store energy from local *PV* installations for later use. These systems contain several *DC* battery cells that are connected to an *AC*/*DC* converter that usually uses field bus protocols for communication, *e. g.*, *Modbus* over *TCP*. On top, manufacturers either define proprietary registers or apply application-layer data models such as *SunSpec*, which describes vendor-independent Modbus register mappings for converters and *ESSs*. With the *Battery Storage Device Model (802)*, SunSpec includes a *BESS*-specific data model that can communicate nameplate ratings, *SoC* level, and settings as well as heartbeat messages with its basic implementation. Specific models exist for different storage technologies, *e. g.*, redox flow batteries, which offer more advanced battery pack-specific monitoring and control, *e. g.*, cell balancing. The converter can be controlled by the *Basic Storage Control Model (124)*, which offers functionality for requesting real power (charge or discharge) from the battery.

[3] Should not be mixed with *EMS* from the context of transmission system operation, where *EMS* refers to computer-supported tools for monitoring, control, and optimization functions.

Stationary *ESSs* in private households have typically a smaller storage capacity than *EVs*. Similarly, their charging and discharging limits are lower, since fast charging is not required and these batteries only need to supply the usual load patterns of households, which are mostly only a few kilo-watt. The reaction duration and ramping capacity of *ESS* are not limiting factors for energy optimization, which is typically applied in a time resolution of several seconds. *BESSs* even qualify for *ASs*, due to their fast reaction time [63, 117]. Since *ESSs* are mostly used to buffer energy from local *PV* generation and supply stored energy during the night to minimize electricity purchase, the *ESS* is usually fully charged during a sunny day. This limits the actual flexibility of an *ESS* without loss of service quality to only shift the charging of the battery in time during *PV* surplus supply.

Apart from *EVs* and *ESSs*, heat pumps and white goods, *e. g.*, dishwashers, washing machines, or air dryers, may offer flexibility for energy optimization. However, their actual power consumption cannot always be controlled directly. *E. g.*, heat pumps with the *Smart Grid ready* label – a certificate valid in Germany, Austria, and Switzerland – support up to four operation modes: *Power off* (duration of 10 min to 120 min; maximum three times a day), *normal operation* (following heat controller logic), *amplified operation* (suggestion to consume more electrical energy; minimal duration of 10 min), *forced operation* (including increased heat set point; minimal duration of 10 min). Similarly, white goods, such as washing machines, can only be controlled by their start and finish time, *e. g.*, using *EEBus*, but not by their actual power consumption.

From these flexible appliances, only *EVs* and *ESSs* are considered in this thesis for *EMS* energy optimization, since heat pumps and white goods provide smaller flexibility and lower control capabilities. Nevertheless, flexibility *provision-as-a-service* can theoretically utilize any kind of local flexible appliances in Chapter 4, as long as their flexibility potential can be implemented within the *MILP* formulation of the *EMS*.

2.4.2 Literature Overview

Many papers in literature [17, 53, 72, 104, 112, 123, 135, 136, 155] model the *EMS* using a *Linear Programming* (*LP*) or *MILP* problem that optimizes the local flexible assets to various objectives, *e. g.*, cost, sustainability, or users' comfort in case the heating domain is included. Some papers target multiple user satisfaction metrics at once by applying meta heuristics, *e. g.*, NSGA-III [43], and others consider uncertainty using robust optimization [136]. However, most of these papers do not consider the aggregation and scheduling of external flexibility requests, as discussed in Chapter 4.

The aggregation and disaggregation of distributed flexibility resources are intensively discussed in the literature as well. Some papers disaggregate power flexibility requests in a single time slot, which is often used for online flexibility disaggregation, while others schedule the flexibility for multiple time slots at once, which is usually applied for day-ahead scheduling. An overview of applied methodologies is given in Figure 2.7. Most of these papers do not consider power grid constraints, focus only on schedule feasibility, lack detailed modeling of home *EMSs*, or aim for a single objective, *e. g.*, cost minimization for the aggregator or the individual participants.

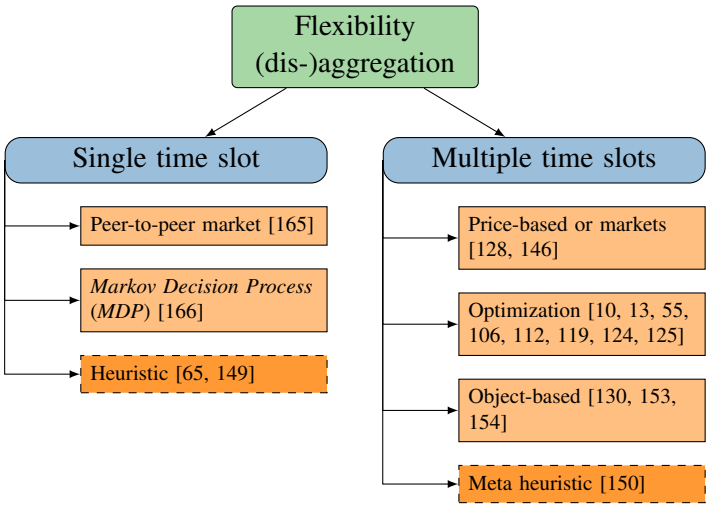


Figure 2.7: Categorization of literature in the domain of flexibility aggregation of distributed *EMS* that considers *QoS* and fairness aspects. The proposed methods in Chapter 4 fit to the two highlighted categories.

Authors of [165] propose a *peer-to-peer* market as a multi-agent system, where *ESSs* are used to optimize the energy prices and customers' utility of the individual agents. Thereby, the *DSO* can employ additional congestion prices to avoid line congestion. The optimization problem defined in [166] aims for economic dispatch at an energy community with flexible assets, among others *EV* charging, that minimizes the aggregated energy cost while avoiding disutility due to delayed charging processes. To tackle uncertainty in the model, a multi-agent *Markov Decision Process* (*MDP*) is defined that infers an action for the next time interval based on the previous state that minimizes the objective function. Since the number of possible states grows exponentially, a *NN* is trained on a set of state-action tuples using the dual problem due to its lower number of decision variables.

The proposal in [65] aims to schedule available resources to flexible customers (mainly *EVs*), while considering grid limitations, respecting the convenience of the users, and reducing the effect of flexibility rebound. The supplying grid feeder, the distributed transformers, and the individual households are represented by agents, which share information and coordinate the usage of home appliances. Thereby, the transformer agents apply a *GA* to assign demand limits to the households such that grid congestion is avoided and the convenience impact factor is minimized. The results of their demand response approach are compared to a *water-filling* algorithm that was proposed for a very similar setup [64]. In [149], *EVs* that support only unidirectional charging are aggregated into a pool of *EVs*, which provides bidirectional flexibility by shifting the charging operation of pending *EVs*. Besides quantification of flexibility, the authors propose a disaggregation scheme that aims to retain positive and negative flexibility using the time-dependent laxity of the *EVs*, which is based on the remaining time and energy to be charged.

The approach in [146] considers power grid constraints and optimizes the demand-side flexibility of customers on multiple time slots in a two-stage approach. In the first stage,

an aggregator sends day-ahead prices to the customers, who locally optimize for minimal cost. In the case of grid threshold violations, the second stage introduces an additional grid usage cost as a set of different incentives, to which each customer optimizes its profile in a rolling window. Finally, the aggregator chooses the best fitting incentives to utilize the required flexibility from the households. In a different approach, the *LP* problem proposed in [128] optimizes all local flexible assets of a household, such as *EVs* and *ESSs* at home *EMS*. Available flexibility is further utilized at the intraday markets following a cost-optimal strategy to minimize the local cost of the *EMSs*.

The multi-objective *MILP* formulated in [55] optimizes the flexibility of multiple apartments in a building at once. Thereby, the energy cost for each apartment is minimized by a separate objective function. Fairness on energy cost is achieved by translating this multi-dimensional objective function into a *goal programming* problem, where the deviation of the minimum cost per apartment to a predefined goal is minimized. The central optimization problem in [10] minimizes the daily operational energy cost of a smart home network while prioritizing *EV* charging. Using their optimization problem formulation, the authors define flexibility indices for *EVs* and *PV* systems to quantify how flexible these assets are. The *EV* flexibility index measures the overlapping of flexibility from consumer assets with *EV* charging intervals and the *PV* flexibility index measures the suitability of consumer appliances to directly consume energy from local *PV* systems. Both indices help to analyze clusters of households with respect to their possible flexibility.

The main goal of the *LP* in [119] is to minimize the total cost for an energy community with *EVs* that support *V2G* functionality. This ultimately leads to an increase in renewable self-consumption, which is in-line with the Portuguese legislation on self-consumption. Authors of [112] formulate a *MILP* for home *EMS* of an energy community that supports reactive power provision and unbalanced three-phase operation. Their model implements voltage constraints on the power grid and utilizes *OLTCs* together with flexibility from *EMSs* to control the voltage in the power distribution grid. The objective function minimizes the total cost of all households and includes a penalty value for discomfort as some means of controlling fairness.

In [124, 125], the focus is on energy flexibility planning, where potential flexibility is characterized by flexibility envelopes in the planning horizon. Flexibility disaggregation is performed by a linear economic dispatch that minimizes the operator's cost, not considering grid constraints. The flexibility of individual assets can also be modeled by convex zonotopes [121], which are inner-approximations of flexibility polytopes for efficient aggregation using the *Minkowski sum*. During disaggregation, a convex cost function is minimized.

Authors of [130, 153, 154] model operational flexibility with so-called *FlexObjects*, which encode the variability of an energy profile (controllable amount of energy for multiple time slots) and the potential to shift this profile in time (time flexibility defined by earliest and latest start time). FlexObjects of single flexible assets are aggregated by aligning their possible start times and profile duration. The disaggregation is performed by assigning the flexibility request proportionally to the energy profiles of the single assets, ignoring potential power grid constraints and the objectives of the single flexibility providers.

Table 2.4: Overview of *QoS* parameters for *EV* charging processes in the literature.

Reference	Autarky or self-consumption	Individual cost	Aggregated cost	Delivery Probability	Discomfort
[150]	✓	✓			
[119]	✓		✓		
[55, 128, 146]		✓			
[165]		✓			
[10, 121, 124, 125]			✓		
[106, 166]			✓	✓	✓
[112]			✓		✓
[13, 65, 149]					✓

Authors of [70] propose a multiple choice combinatorial problem formulation for distributed energy scheduling in a multi-agent system, where the goal is to select a set of solutions that optimize towards a target schedule. This method is extended in [150] to support multi-criteria optimization at the agents, which aim for minimal cost, maximal self-consumption, and peak shaving. Their solution uses a meta heuristic to find suitable schedules for the local optimization problems at the agents.

Table 2.4 summarizes the *QoS* parameters that are targeted in the literature. Additionally, only very few papers discuss fairness aspects of their approaches, including deviation from the original profile [55] and discomfort penalty [112].

The literature review shows that plenty of work is done with respect to user satisfaction of a single *EMS*. However, disaggregation of flexibility is rarely addressed, mainly focusing on cost. This thesis discusses *flexibility-as-a-service* from the perspective of the user, where scheduling flexibility to distributed *EMS* respects the goals of the distributed *ESS* and the needs of the users, and additionally targets fairness aspects.

2.5 General Assumptions

The thesis makes several assumptions about the power grid infrastructure, the *EVs* and their charging equipment, and the distributed *EMSs*. Common assumptions on *EVs* are summarized and discussed in the following. Additional service-specific assumptions are stated in Sections 3.1 and 4.1.

A1 The discussed solutions in this thesis focus on residential home charging. Nevertheless, proposed methods with some adaptions can be applied to workplace charging as well. Public fast charging stations are of less interest because charging service

expectations and possible control flexibility differ. Additionally, controlling public charging stations might not be desired by the charging point operator, whose goal is to maximize charging point utilization, *e. g.*, by reducing additional parking duration with extra parking fees.

A2 *EVs* support only *G2V* and no *V2X* functionality. *V2G* and *V2H* are typically realized with *DC* charging stations, where the converter that needs to support bi-directional power flow is part of the charging station and not part of the *EV*. For *EV* home charging, typically *AC* charging stations are used, where the converter is part of the *EV* and does usually not support bi-directional power flows. *V2L* only supports a single load socket at the *EV* and is therefore not relevant.

A3 *EVs* start charging with a given minimum charging current and can dynamically change their charging current in discrete steps above the minimum charging current during the charging process. Thereby, the charging efficiency remains constant. According to Section 2.3, even the simple *PWM* signaling of the Type-2 charging specification supports such a charging control. The controllability of the variable charging current is also demonstrated in Section 3.6.2.

A4 Whenever a controller changes the charging current, the *EV* changes its charging current instantaneously. Because the main idea of the *PWM* signaling is to limit the current carrying capacity, the *EV* must directly apply the charging current reduction. An increase is also applied within a few seconds as shown in Section 3.6.2.

A5 *EV* batteries are modeled only by constant current charging without a constant voltage phase. This battery saturation phase usually starts earlier with higher charging currents. In the case of *EV* home charging, the maximum charging current is below 96 A (32 A per phase), which is supported up to a high *SoC* by many real *EV* models [147].

A6 *EVs* can configure reactive power behavior during the charging operation to mitigate local voltage violations. This for example is part of the German connection grid code VDE 4100, where *EV* charging equipment must support a reactive power control strategy. The charging control standard ISO 15118 even supports communication of real and reactive power profiles to the *EV*. *EVs*, which cannot control reactive power, can only participate in the mitigation of under-voltage events by reducing the real charging power, which will affect the actual charging service much more.

A7 The future driving and charging requirements of *EV* users are known. This includes the arrival and departure time at home, as well as the required energy to be charged. In the case of home charging, this information can either be inferred from commuting patterns or via manual input to the home *EMS* by the user.

2.6 Chapter Summary

This chapter summarizes the relevant background information on the electrical power system and its operation. This includes the historical development of interconnected syn-

chronous power systems and typically used power grid topologies in the power distribution grid. Furthermore, the term *flexibility* in power systems is defined, characteristics of flexible resources are discussed, and the relation between flexibility and *ASs*, which are required for stable grid operation, is given. Moreover, the notions of *Quality of Service* and *Quality of Experience* from the networking domain are discussed, where a conceptual connection is drawn to the use case of *EV* charging. This is followed by different definitions of fairness from the networking and compute task scheduling.

In addition, the two smart grid applications under investigation are inspected regarding their flexibility and means of control. From the *EV* charging domain, only *EVs* that charge directly from the power grid, called *Plugin Electric Vehicles*, are relevant for this thesis. Thereby, control methods for *AC* charging, which is commonly used for the most flexible *EV* home charging, are explained by the use of IEC 61851-1 and ISO 15118-2. In the domain of power flexibility provision from *EMS*, additionally battery *ESS* are relevant due to their large flexibility potential and easy controllability, *e. g.*, via *Sunspec*. Relevant work that discusses *QoS* and fairness aspects in both smart grid applications is categorized and discussed. Finally, general assumptions on *EV* charging, which is part of both smart grid applications under investigation, are stated and shortly discussed.

CHAPTER 3

Electric Vehicle Home Charging Service

This chapter focuses on *EV charging-as-a-service* that offers *QoS*, *QoE*, and fairness to distributed charging services under the assumptions stated in Section 3.1. Section 3.2 describes the charging process model and Section 3.3 extends the notion of *QoS* and *QoE* from [32, 35]. The queuing approach of [35] is detailed in Section 3.4.1 and distributed probabilistic allocation mechanisms, which are inspired by wireless communication networks and support *QoS* control, are newly introduced in Section 3.4.2. Both online charging service allocation methods are evaluated on realistic input data in Section 3.5, including a sensitivity analysis of the configuration parameters. Finally, the applicability of the proposed solutions – supported by [11, 31] – and a possible extension for improved voltage control [36] is discussed in Section 3.6, before a concluding summary is given.

---

[11] Ammar Alyousef, Dominik Danner, Friederich Kupzog, and Hermann de Meer. "Enhancing power quality in electrical distribution systems using a smart charging architecture". In: *Energy Informatics* 1.1 (Oct. 2018). Article 28, Proceedings of the 7th DACH+ Conference on Energy Informatics, Oldenburg, Germany, Best Paper Award.

[31] Dominik Danner, Ammar Alyousef, Philipp Danner, Wolfgang Duschl, and Hermann de Meer. "Towards grid-friendly electric vehicle charging: Architectural concept and field trials". In: *3rd E-Mobility Power System Integration Symposium*. Dublin, Republic of Ireland: Energynautics, Oct. 2019.

[32] Dominik Danner, Wolfgang Duschl, and Hermann de Meer. "Fair charging service allocation for electric vehicles in the power distribution grid". In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy '19. Poster. Phoenix, AZ, USA: Association for Computing Machinery, June 2019, pages 406–408.

[35] Dominik Danner and Hermann de Meer. "Quality of service and fairness for electric vehicle charging as a service". In: *Energy Informatics* 4.3 (Sept. 2021). Article 16, Proceedings of the 10th DACH+ Conference on Energy Informatics, Virtual.

[36] Dominik Danner and Hermann de Meer. "Max-consensus protocol to determine the regulated node in distributed voltage regulation". In: *Energy Informatics* 5.1 (Sept. 2022). Article 15, Proceedings of the 11th DACH+ Conference on Energy Informatics, Freiburg, Germany.

3.1 Assumptions

In addition to the general Assumptions A1 - A7 on *EV* charging in Section 2.5, this chapter requires further assumptions on the power distribution grid, which are detailed in the following.

A8 The power distribution grid is given in radial topology with a transformer station towards the next higher voltage level as the root node. Furthermore, not all grid parameters are available in the power distribution grid, hence parameters, such as line impedance, line length, or exact loads at the buses, are not known. This is generally valid for historically grown low-voltage power distribution grids, where there is usually no monitoring in place and exact line parameters are not known by the *DSOs*. Despite power distribution grids being built in a ring or meshed topology in real life, they are operated as a radial network for faster fault detection, as discussed in Section 2.1.

A9 For the *EV* charging use case, it is assumed that no reverse power flow happens in the power distribution grid, which would typically be the case with high *PV* penetration. However, the proposed charging service allocation mechanisms are generally applicable to reverse power flow as well. In that case, the current carrying capacity of the lines and the transformer must be considered for both directions. Note that the mechanisms are not designed to solve power grid problems introduced by other generations or loads, but only limit *EV* charging such that this application does not harm the power grid infrastructure.

A10 Critical power grid assets and locations in the power grid are monitored and respective local voltage and utilization measurements are instantaneously available at the local nodes and are communicated to other nodes via *ICT* without delay or data losses. This can be realized by distributed measurement devices and/or by modern smart metering devices at the grid connection points of the customers, *e. g.*, using the *Smart Meter Gateway* (*SMGW*) concept in Germany. In case of communication loss, the proposed mechanisms go to a fail-safe state with reduced charging power to avoid harming the power grid.

A11 It is assumed that applying a coordinated charging service allocation mechanism in time slots of one minute is sufficient for *QoS* control of charging services and fast enough for reasonable load management in the power distribution grid. The evaluation results in Section 3.5.2 underline this assumption and Section 3.6.2 shows that control loops of one minute can be implemented in real-world environments.

3.2 Charging Service Model

A *charging service* defines the requirements for charging an *EV* and is given by (1) time of arrival t_{arr} when the *EV* reaches the charging location, (2) time of departure t_{dep} when the *EV* must leave the charging location at the latest and (3) the required energy to be charged

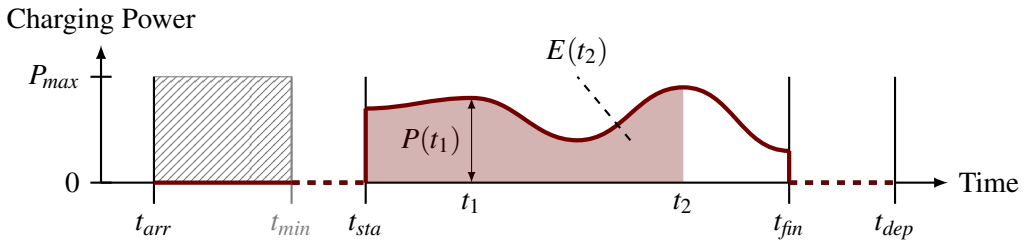


Figure 3.1: Charging service with its arrival time t_{arr} and departure time t_{dep}, the start time t_{sta} and finish time t_{fin} of a charging process, the minimum charging time t_{min} as dashed area, and the charging power profile $P(t)$ colored in solid red. Note that t_{min} must not necessarily be smaller than t_{sta}, and t_{fin} can be equal to t_{dep}.

Table 3.1: Parameters of a charging service.

Parameter	Description
$t_{arr} \in \mathbb{R}$	Arrival time of the *EV* for the charging service
$t_{dep} > t_{arr}$	Departure time of the *EV* from the charging service
$E_{req} > 0$	Required energy of the charging service
$P_{max} \in \mathbb{R}^+$	Maximum charging power
$P(t) \in [0, P_{max}]$	Charging power profile for $t \in [t_{arr}, t_{dep}]$
$E(t) \in \mathbb{R}_0^+$	Charged energy at time t
$t_{sta} \in [t_{arr}, t_{dep}]$	Starting time of the charging process
$t_{fin} \in [t_{arr}, t_{dep}]$	Finish time of the charging process
$t_{min} > t_{arr}$	Minimum possible finish time

$E_{req} > 0$ before departure. *EVs* that do not require any energy at all ($E_{req} = 0$) are not considered for a charging service. Time of arrival and time of departure together form the *available charging time* of the charging service. In Assumption A7, it is assumed that *EV* drivers provide their expected departure time and required energy with high confidence. This assumption is investigated with sensitivity analysis of the input parameters in Section 3.5.3. Figure 3.1 gives an overview of a charging service and its parameters, which are summarized in Table 3.1.

The charging power during a charging service is specified as a function over time t in Equation (3.1), where the maximum charging power P_{max} is limited by the charging equipment.

$$P\colon [t_{arr}, t_{dep}] \to [0, P_{max}]\colon t \to P(t) \tag{3.1}$$

The charged energy during a charging service until time t is denoted in Equation (3.2). As defined in Assumption A3, a constant efficiency factor μ_{EV} is assumed for the charging equipment and the battery of the *EV*, independent of the batteries' *SoC* or the ambient temperature. Note that according to Assumption A2 *V2G* is not considered, because it is counter-intuitive to the user for *charging-as-a-service*. Nevertheless, the following defini-

tions support *V2G* when allowing negative charging powers in Equation (3.1) and adapting the energy calculation in Equation (3.2) with discharging efficiencies.

$$E(t) = \int_{t_{arr}}^{t} P(x) \cdot \mu_{EV} \, \mathrm{d}x \tag{3.2}$$

A *charging process* eventually starts at the first moment when the charging power profile is greater than zero and the *EV* charges energy into its battery. The *start time* is formally defined in Equation (3.3). If no charging process starts within the available charging time of a charging service, the charging power profile is equal to the identical zero function, and t_{sta} is set to t_{dep}. The required starting procedure of a charging process is small compared to the overall charging time and can therefore be neglected [4].

$$t_{sta} = \begin{cases} t \in [t_{arr}, t_{dep}] \colon P(t) > 0 & \\ \qquad\qquad \wedge \nexists t_2 < t \colon P(t_2) > 0 & \text{if } \exists t \in [t_{arr}, t_{dep}] \colon P(t) > 0 \\ t_{dep} & \text{else} \end{cases} \tag{3.3}$$

The *finish time* of a charging process, defined in Equation (3.4), is either the first moment in time when the required energy E_{req} is reached or equal to the departure time.

$$t_{fin} = \begin{cases} t \in [t_{arr}, t_{dep}] \colon E(t) = E_{req} & \\ \qquad\qquad \wedge \nexists t_2 < t \colon E(t) = E(t_2) & \text{if } \exists t \in [t_{arr}, t_{dep}] \colon E(t) = E_{req} \\ t_{dep} & \text{else} \end{cases} \tag{3.4}$$

Furthermore, a charging service *succeeds* if at least the required energy is charged until departure time $E(t_{dep}) \geq E_{req}$. Finally, Equation (3.5) calculates the *minimum finish time* $t_{min} > t_{arr}$, at which the charging process would finish when starting with a constant maximum charging power profile $\forall t \leq t_{min} \colon P(t) = P_{max}$ at arrival t_{arr} following Assumption A5. In case the charging service cannot finish within the available charging time, the minimum charging time is after the departure time $t_{min} > t_{dep}$.

$$t_{min} = t_{arr} + \frac{E_{req}}{P_{max} \cdot \mu_{EV}} \tag{3.5}$$

3.3 Charging Service Assessment

Based on the charging service definition and its properties, this section details *QoS* and *QoE* parameters, as well as the fairness aspects of charging services.

3.3.1 Quality of Service

In the context of this work, the *QoS* of a charging service is defined by the following criteria, considering the charged energy, the start and finish times of the charging process, and the charging power profile. Note, that some *QoS* criteria are calculated at the end of the charging service, while others can also be evaluated during the charging service.

Finished Charging Service The main goal of a charging service is to deliver the required energy E_{req} to the *EV*. A successful charging service receives maximal *QoS*. Otherwise, the *QoS* degrades proportionally with the remaining energy to be charged. This first *QoS* metric, given in Equation (3.6), can be evaluated at each point in time during the charging service, and $QoS_1(t_{dep})$ is the final metric score of the charging service.

$$QoS_1(t) = \frac{E(t)}{E_{req}} \in [0,1] \tag{3.6}$$

For all successful charging services, $QoS_1(t_{dep})$ yields an equal maximum *QoS* score, but the *QoS* can further differ with respect to the time, at which the charging service reached the required energy. Charging services that successfully finish far ahead of their departure time, receive higher *QoS* compared to others, which finish shortly before departure. Therefore, a secondary *QoS* metric QoS_2 is defined in Equation (3.7), which is orthogonal to QoS_1.

$$QoS_2 = \begin{cases} \frac{t_{dep}-t_{fin}}{t_{dep}-t_{min}} \in [0,1] & \text{if } t_{min} \neq t_{dep}, \\ 0 & \text{else} \end{cases} \tag{3.7}$$

If a charging service did not successfully finish, QoS_1 measures the *QoS* within $[0,1]$ and, according to Equation (3.4), the finish time t_{fin} is equal to the departure time t_{dep}; hence QoS_2 is always zero. Only for successful charging services, where $QoS_1 = 1$ and $t_{fin} \geq t_{min}$, QoS_2 varies within $[0,1]$ depending on the finish time of the charging process.

Charging Start Time The second main criterion of a charging service considers the initial waiting time of a charging process. It follows the logic that a charging process that starts earlier has a higher chance to finish in time. Furthermore, a waiting charging process does not receive any service until the charging process actually starts. Therefore, waiting charging services receive a lower *QoS* score, as defined in Equation (3.8).

$$QoS_3 = \begin{cases} 1 - \frac{t_{sta}-t_{arr}}{t_{dep}-t_{arr}} \in [0,1] & \text{if } \exists t \in [t_{arr}, t_{dep}] \colon P(t) > 0, \\ 0 & \text{else} \end{cases} \tag{3.8}$$

In case a charging process did not start at all, the charging power profile $P(t)$ is the identical zero function and hence no t with $P(t) > 0$ exists, which results in $QoS_3 = 0$. An immediately starting charging process has $t_{sta} = t_{arr}$ and receives the highest score with $QoS_3 = 1$. The longer an *EV* waits before charging, the lower its metric score.

There exists a dependency between QoS_1 and QoS_3 if no energy is charged during the charging service. In this case QoS_1 and QoS_3 result always in zero. As soon as QoS_1 is unequal zero, the charging process starts at any time in $[t_{arr}, t_{dep}]$ and QoS_3 can be any value from the interval $[0,1]$. Furthermore, QoS_1 might not reach the maximum value with a late starting time (*e. g.*, $t_{sta} = t_{dep}$). Apart from that, both metrics are independent of each other. A higher dependency exists between QoS_2 and QoS_3 because both consider the time domain. QoS_2 may affect QoS_3 because the later a charging process starts charging, the more time from the interval $[t_{arr}, t_{min}]$ is spent on waiting, which in turn may reduce QoS_2.

Charging Power A third *QoS* criterion is the variation of the charging power $P(t)$ over time, which is one representative of an intrinsic *QoS* parameter that measures infrastructure parameters. In communication networks, this is referred to as *packet jitter*,[4] which measures the variation of packet delays. The *EV* charging service focuses on charging power variation, where high charging power variation results in bad residual charging time estimation. This in turn reduces the quality of the feedback to the user. The respective *QoS* metric assumes a discrete-time model with a constant time slot size of Δ and is defined in Equation (3.9), where $\sigma(X)$ is the sample standard deviation of a set $X = \{x_1, \ldots, x_n\}$, calculated by $\sigma(X) = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2}$. Note that in case no charging process starts, this metric is equal to 0 because no residual charging time estimation can be done.

$$QoS_4 = \begin{cases} 1 - \frac{2 \cdot \sigma(\{P(t) : t_{sta} \leq t \leq t_{fin}\})}{P_{max}} \in [0,1] & \text{if } t_{sta} < t_{fin} \\ 0 & \text{else} \end{cases} \tag{3.9}$$

The sample standard deviation σ returns a value from the interval $[0, \frac{P_{max}}{2}]$ since the charging power profile is limited between 0 and P_{max}. Therefore, QoS_4 normalizes the output to the interval $[0,1]$. Additionally, values of $P(t)$ are only taken from the interval $t \in [t_{sta}, t_{fin}]$, because only the variation during the actual charging process matters. The waiting time before charging and the time after the required energy is charged are not relevant for the remaining charging time estimation.

QoS_4 is independent of QoS_1 because the amount of charged energy $E(t)$ does not have any impact on the standard deviation of the charging power profile. Only in case no energy is charged at all ($QoS_1 = 0$), the charging power profile is the identically zero function and, hence, QoS_4 is zero as well. Similar relation can be stated for QoS_4 with QoS_2 and QoS_3, respectively. If no energy is charged, both are fixed to zero likewise QoS_4. Otherwise, the variation of the charging power profile is almost independent of the start and finish time.

3.3.2 Quality of Experience

Compared to *QoS*, *QoE* takes a bigger picture of the service and measures the delight and annoyance of the user. For *EV* charging, not only the charging service itself (defined by arrival, departure time, and required energy) but also the circumstance of the charging service is considered. Because this thesis focuses on enhancing the *QoE* fairness of a charging service from a technical point of view, aspects like the simplicity of payment or convenience of charging service are not considered in the following *QoE* criteria, even though they play an important role in the users' perception.

Battery State of Charge The first *QoE* metric refers to the battery *SoC* instead of the charged energy of the charging service. Especially, with different battery sizes but the same energy requirements, a user finally does not see the actual energy charged (in kWh), but only the battery *SoC* is displayed in the car (in %). Following this user's recognition, a high *SoC* (near to SoC_{target}) corresponds to a high *QoE* and *vice versa*. Similar to QoS_1,

[4] Interested readers are pointed to RFC 3393.

the *QoE* metric in Equation (3.10) can be evaluated at each point in time during a charging service, and $QoE_1(t_{dep})$ is the final metric score.

$$QoE_1(t) = \frac{SoC(t)}{SoC_{target}} \in [0,1] \tag{3.10}$$

Next Trip A second criterion for *QoE* is whether the *EV* driver will reach the next destination without a charging stop on the way, which is expressed as a binary metric in Equation (3.11). The next trip is defined as *feasible* if the battery holds enough energy to reach the next destination with a *SoC* greater than 10 % at arrival. With the remaining *SoC*, it should be possible to reach the next charging facility near the destination. This again relates to the users' recognition and range anxiety, which let the driver recharge the battery before running out of energy.

$$QoE_2 = \begin{cases} 1 & \text{if the next trip is feasible,} \\ 0 & \text{else} \end{cases} \tag{3.11}$$

It can easily be seen that QoE_1 and QoE_2 are independent to a certain degree because the next trip is independent of the actual battery size and its *SoC*. Nevertheless, with low *SoC* (low QoE_1) it is more unlikely to reach the next destination.

Utilization The third *QoE* criterion targets the utilization of the charging equipment. Users with faster charging stations, *e. g.*, with 22 kW peak charging power, expect that the peak power of the charging equipment is utilized. If the mean charging power during the charging process drops compared to the peak charging power of the charging equipment, the *QoE* lowers as well. Similar to QoS_4, this metric assumes a discrete-time model with a constant time slot size of Δ, and the mean value is calculated by $\mu(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$.

$$QoE_3 = \frac{\mu(\{P(t) : t_{sta} \leq t \leq t_{fin}\})}{P_{max}} \in [0,1] \tag{3.12}$$

This metric is independent of QoE_1 and QoE_2 because the mean charging power does not directly influence the *SoC* at the end of the charging service nor relates to the next trip.

3.3.3 Fairness

Besides a high quality of service and experience scores, charging service operators are interested in how fair the achieved *QoS* and *QoE* metric scores are distributed among the charging services. In contrast to the literature, which often defines fairness based on the instantaneous charging power received by the charging service, this thesis evaluates fairness on the received *QoS* and *QoE* metric scores, which better depict the received service from a user's perspective. A very unfair allocation means that the *QoS* and *QoE* metrics differ significantly among the different charging services, whereas very similar metric scores can be considered fair. To analyze fairness separately from the *QoS* and *QoE* values, the fairness index must be independent of the achieved metric values. Therefore, the fairness index by

Hoßfeld et al. [74] in Equation (2.3) is used. The index calculation uses the metric scores $S = \{s_1, \ldots, s_m\}$ of m different charging services. For every single metric defined above, a fairness index can be computed on a set of charging services, *e. g.*, all charging services of one week. Since all *QoS* and *QoE* metrics from Equations (3.6) - (3.12) are defined within $[L,H] = [0,1]$, the fairness index simplifies to Equation (3.13).

$$F(S) = 1 - \frac{2\sigma(S)}{H-L} = 1 - 2\sigma(S) \tag{3.13}$$

3.4 Methodology

This section presents a distributed hierarchical charging service allocation mechanism based on queuing networks, which is published in [35]. Furthermore, a novel distributed probabilistic charging service allocation protocol is introduced, which is inspired by *MAC* protocols from the networking domain. Both charging service allocation mechanisms include means of controlling *QoS*.

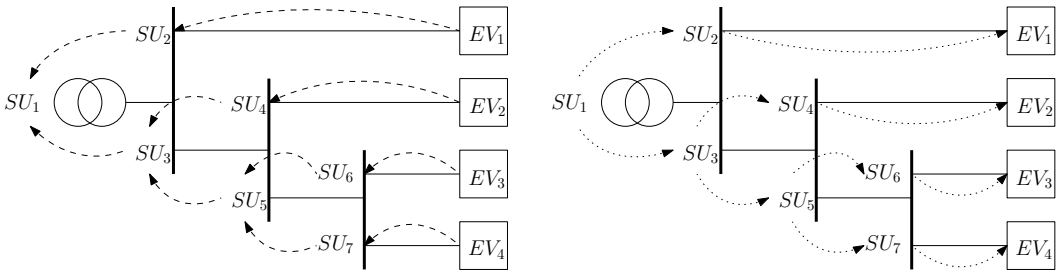
3.4.1 Hierarchical Queuing Networks

In communication networks, multiple information flows pass through the same shared physical link simultaneously, *e. g.*, using *time-division* or *frequency-division* multiplexing. For packet switching, often queuing models are used to send packets over a network of nodes. Each node holds a queue with packets awaiting transmission to another node. Whenever the communication link is free, a scheduler selects the next packet in the queue, normally based on a *first-in-first-out* policy. To establish a certain *QoS*, other policies can be applied such as *EDF*, *LLF*, *Weighted Fair Queuing* (*WFQ*), or *packet prioritization*.

In this work, each *EV* is represented by a flow, and the *EV* can request charging currents by scheduling packets to the power grid in discrete time slots (*e. g.*, one minute), where the packet size p_{size} equals the *EVs*' minimal adjustable charging current. This guarantees that *EV* battery constraints and control capabilities are considered as stated in Assumptions A4 and A5. For example, if the battery of an *EV* limits the charging power to 6.9 kW (10 A on three phases) and the *EV* can change its charging current in discrete 3 A steps (1 A per phase) the *EV* charging service needs to queue ten packets for each phase. Without loss of generality, this section models the packet allocation for a balanced three-phase power system, hence considers only one queue for a single phase.

3.4.1.1 Charging Packet Request and Assignment

The shared network is the underlying power distribution grid whose bandwidth is limited by the available capacities. To not overload power grid assets, a *Scheduling Unit* (*SU*), which contains the queuing logic, is placed at each limiting cable or transformer. Because the power distribution grid is typically operated as a radial network as stated in Assumption A8, the single *SUs* span a tree on top of the power grid infrastructure. Each *EV* requests



(a) Packet requests. **(b)** Packet assignments.

Figure 3.2: Hierarchical composition of *EVs* and *SUs* for the queuing approach [35].

charging current packets to the nearest connected *SU*, typically at the supplying cable or transformer. The requested packets pass the network tree toward the root node as depicted in Figure 3.2(a). Thereby, each SU only forwards packets to the next *SU* as long as the local capacity limit is not reached. Finally, the root node, *e. g.*, responsible for the transformer of a low-voltage network, assigns its available capacity by returning the packets top-down to the *EVs*, as depicted in Figure 3.2(b). Note that voltage violations are treated by a $Q(U)$-controller described in Section 3.5.1.

To determine the available capacity of a shared link, Assumption A10 allows measuring parts of the power distribution grid in each time slot, and from that inferring the available charging capacity of the next time slot using short-term forecasts, *e. g.*, with a persistence model, and data-driven state estimation [34]. The bandwidth calculation can also include an approximation of the power distribution grid losses, which in turn reduces the actual available charging capacity at the end of the radial network, *e. g.*, depending on the cable type and power grid loading. As discussed in Section 3.6.1, the root node may participate in *DR* programs or act as a market agent, which artificially limits the *EV* charging power based on market signals in addition to the pure network-related capacity limitations. This request and response control loop repeats every minute as stated in Assumption A11.

The assigned charging capacity $C_i(t)$ of an *EV* i in time slot $t \in \mathbb{N}$ is calculated in Equation (3.14), where $A_i(t)$ is the set of packets that are assigned to the *EV* by the *SU* and p_{size} is the size of the packet p. Because *EVs* that still have packets in the queue might leave before the next iteration starts, all queues are flushed afterward so that no power grid capacity is lost. Finally, the allocation algorithm is executed again for the next time slot, starting with the request of packets from the *EVs*.

$$C_i(t) = \sum_{p \in A_i(t)} p_{size} \tag{3.14}$$

A reliable and fault-tolerant *ICT* architecture between the *SUs*, measurement devices in the power grid, and the *EVs* is required for the aforementioned procedure. In case of communication issues, like packet delay or drop, affected *EVs* must move to a *fail-safe* mode and pause their ongoing charging processes immediately to avoid damage to the power grid. In this fail-safe mode, charging capacity is still requested to the *SU* until the *EV* receives new packets and can resume the charging operation in normal operational mode. Note that this

work does not target privacy nor security issues like packet injection, that may arise with the operation of *ICT* infrastructure. The proposed approach includes central schedulers at limiting cables or at the transformer station, which can be seen as a single point of failure for all attached *EVs*. However, *SUs* are logical units that do not store state information beyond a single time slot. Hence, *SU* instances can be executed in a cloud environment, which allows fast fault recovery.

The communication effort between entities scales linearly with the number of involved *EVs* and *SUs*. In each time slot, *EVs* (leaf nodes in the tree topology) need to send the requested packets along the path of *SUs* (inner nodes) to the root node, which finally returns a subset of these packets to their origins. Sending packets from one node to another produces a constant cost of $\mathcal{O}(1)$. Because each *EV* and each *SU* sends all packet requests in a single message only once back and forth in each time slot, the total communication cost can be approximated with $\mathcal{O}(n+m)$, where n is the number of inner nodes and m is the number of leaf nodes.

3.4.1.2 Queuing Policies

In the following, different commonly known scheduling policies are presented and discussed with respect to *EV* charging, including the proposed *dynamically weighted fair queuing* approach to reach high *QoS* and *QoE* scores and fairness. All policies can be implemented using the same queuing logic from the previous section. In general, the scheduling policies make use of the two main properties of a charging service: The remaining available charging time and the remaining required energy (or *SoC* level).

***First-Come-First-Served* (*FCFS*)** The typical implementation of a queue is the *first-in-first-out* strategy, where the first element that reaches the queue is the first element that will be processed. In the case of *EV* charging, the policy translates to a *first-come-first-served* policy. The *EV* that arrives earlier will be served first with the maximum possible charging current. If there is available charging capacity left, the *EV* that arrived next will be served, and so forth. Even though considering only the arrival time of *EVs* for the charging service allocation is simple to realize and secure against malicious user inputs, the flexibility (required energy and residual charging time) of the *EV* is not considered at all with this scheduling policy.

***Earliest-Departure-First* (*EDF*)** Contrary to the *FCFS* scheduling, the *earliest-deadline-first* policy executes tasks in the order of the nearest deadline, *e. g.*, used in *Central Processing Unit* (*CPU*) task scheduling [103]. The idea behind this method is to process the more critical tasks first under the assumption that on average each task takes a similar execution time. For *EV* scheduling, *earliest-deadline-first* turns into *earliest-departure-first*, only considering the departure time of the *EV* during scheduling. Similar to *FCFS*, this policy does not consider the required energy, and all packets of the same *EV* are scheduled with the same priority, which results in maximum charging rates for only a few *EVs*.

***Least-Laxity-First* (*LLF*)** The priority of a task is inverse to its slack time, which is equal to the remaining extra time after job execution until its deadline. Note that the slack time can even be negative in case the job cannot finish on time, which however does not change the execution order of *least-laxity-first* scheduling. The slack time $s(t)$ at any time t is calculated in Equation (3.15), where d is the deadline, $r(t)$ is the release time since the start, and $c(t)$ is the residual computation time at time t.

$$s(t) = d - r(t) - c(t) \tag{3.15}$$

In *EV* charging, the departure time is equal to the deadline ($d = t_{dep} - t_{arr}$), the time spent within the available charging time equals the release time ($r(t) = t - t_{arr}$) and the required charging time with an assumed maximum charging power equals the residual computation time ($c(t) = \frac{E_{req} - E(t)}{P_{max} \cdot \mu_{EV}}$). Note that for calculating the residual charging time constant current charging with maximum charging power is considered, following Assumption A5.

***Proportional* (*PROP*)** The *proportional fair* scheduling policy guarantees that every participant receives a fair share of a limited resource proportional to its anticipated resource consumption. Proportional fairness is discussed in literature many times concerning the expected charging power [14, 15, 94, 144, 151], hence the proportional allocation policy is also based on the charging power requests of the *EVs*. In the hierarchical queuing network, proportional fairness is a local property between *EVs* connected to the same *SU*. Capacity limitations along the power distribution grid prevent achieving global proportional fairness.

***Weighted Fair Queuing* (*WFQ*)** Packet *GPS* [42] can be approximated with *weighted fair queuing* and is used to share a resource's capacity fairly between flows, while the weight determines the fraction of capacity that each flow receives. In network scheduling, each of N packet flows, that passes a shared link, is managed by a separate queue i with a weight $w_i \geq 0$, which is determined by the priority of that flow. Every time a new packet p is received, its *virtual finish time* is computed in Equation (3.16), where $virtStart_i$ is the last *virtFinish* time of the same queue i (or the current time if the queue is empty) and b_i is the assigned bandwidth of that queue. The virtual finish time refers to the time when the packet would arrive if all packets of the queue are sequentially sent with the given bandwidth.

$$p_{virtFinish} = virtStart_i + \frac{p_{size}}{b_i} \tag{3.16}$$

The bandwidth calculation of flow i is given in Equation (3.17) using the weights of all single queues w_j and the maximum bandwidth R of the shared link.

$$b_i = \frac{w_i}{\sum_{j=0}^{N} w_j} \cdot R \tag{3.17}$$

Whenever the scheduler can send a packet over the shared link, it selects the queue that contains the packet with the smallest *virtFinish* time and sends the first packet from that queue. Thereby, resources are allocated proportionally fair to the weight of each queue (flow) independent of the packet sizes.

Algorithm 3.1: Packet request aggregation at *SUs*.

Input: Set of packets P requested from *SU* or *EV*
Output: Set of packets R for the request to the next higher level
forall $p \in P$ **do**
    $Q_i \leftarrow getQueue(p)$;
    $p_{virtFinish} \leftarrow$ (3.16), (3.17), (3.18), and (3.19);
    $Q_i \leftarrow Q_i + p$;
end
$R \leftarrow \emptyset$;
while *local capacity limit not reached* **do**
    $Q_i \leftarrow getNextQueue()$;
    $R \leftarrow R \cup \{nextPacket(Q_i)\}$;
end
return R

The pseudo-code for requesting packets at a *SU* is given in Algorithm 3.1 and the packet assignment is shown in Algorithm 3.2. In both cases, Q_i denotes the packet queue of *EV/SU* $i = 1,\ldots,N$, $getQueue(p)$ determines the queue of packet p, $getNextQueue()$ returns the queue with the smallest *virtFinish* time, and $nextPacket(Q)$ returns the packet with the smallest *virtFinsh* time of queue Q. In networking, the packet size p_{size} denotes the number of bits of the packet and the bandwidth defines how many bits can be transmitted per second (bit/s). For *EV* charging, the packet size is given by the minimum adjustable charging current. Because the requests are only valid for a discrete-time Δ, the actual packet size can be seen as the electrical charge (Ah) that needs to be transmitted by the power grid. Analogously, the bandwidth is the current carrying capacity of the power grid (in A).

In statically *WFQ*, the weight of one charging service is determined at the beginning of the charging service when the *EV* arrives at the charging station. Similar to [179], the weight is based on the comparison of the required *SoC* with the current *SoC* of the *EV*. The weight of each *EV* i is calculated in Equation (3.18). *EVs* that require a full charge receive a weight of 11 and *EVs* that arrive at home with the required *SoC* in the battery obtain a weight of 1, hence will only charge with minimum priority.

$$w_i = \frac{\max(SoC_{target,i} - SoC_i(t_{arr}), 0)}{100\,\%} \cdot 10 + 1 \tag{3.18}$$

Dynamically Weighted Fair Queuing (DWFQ) By dynamically changing the weights of the flows, *WFQ* can be utilized to control the *QoS* for each flow. In contrast to statically weighted fair queuing by [179], the *DWFQ* policy considers both flexibility aspects of the charging service, namely available charging time and required energy. Because for *WFQ* the weights $w_i(t)$ must be greater than zero, the slack time cannot be used to dynamically estimate the weight of a charging service like in *LLF* policy. Instead, the remaining charging time $c(t)$ is divided by the remaining time until departure as in Equation (3.19). Charging services that cannot finish on time are given a weight greater than 1, while others that may finish until departure are given a weight less or equal to 1. Once the weight crosses

Algorithm 3.2: Packet assignment at *SUs*.

Input: Set of assigned packets P for this *SU*
Output: Assignments A_i for *EV/SU* $i = 1, \ldots, N$
$\forall i = 1, \ldots, N$: $A_i \leftarrow \emptyset$;
$C \leftarrow \sum_{p \in P} p_{size}$;
while min(C, *local capacity limit*) *not reached* **do**
  $Q_i \leftarrow getNextQueue()$;
  $p \leftarrow nextPacket(Q_i)$;
  $A_i \leftarrow A_i \cup \{nextPacket(Q_i)\}$;
end
return $(A_i)_{i=1,\ldots,N}$

the uniform value of 1, it stays greater than 1 until departure, since under Assumption A5 the remaining available charging time is not sufficient for the charging service to succeed.

$$w_i(t) = \frac{\left(\frac{E_{req} - E(t)}{P_{max} \cdot \mu_{EV}}\right)}{t_{dep} - t} \tag{3.19}$$

3.4.2 Distributed Probabilistic Protocols

In communication systems, there exist many distributed mechanisms for *MAC* on a shared medium. For instance, the *Carrier Sense Multiple Access / Collision Detection* (*CSMA/CD*) method, implemented in *Ethernet* networks as defined in IEEE 802.3, ensures that a sender does not start transmitting a frame if the shared medium is *busy* (*carrier sense*) and immediately stops transmission if another sender transmits simultaneously (*collision detection* by listening to the medium in parallel).[5] Whenever a *collision* is detected, all senders stop transmitting immediately and sample a random backoff time (waiting time), until which the senders retry sending the frame. This backoff time is typically sampled from a *Binary Exponential Backoff* (*BEB*) window that increases exponentially with the number of consecutive *collisions*. Wired Ethernet networks allow *collision* detection, whereas in wireless networks simultaneously receiving during transmission is challenging and, due to the *hidden node problem* [108], carrier sensing might even fail. Therefore, wireless network protocols apply *Carrier Sense Multiple Access / Collision Avoidance* (*CSMA/CA*) implemented by the *Distributed Coordination Function* (*DCF*), defined in IEEE 802.11, where instead of detecting (and tolerating) *collisions*, the goal is to avoid *collision* upfront. To do so, the sampled random waiting time from the *BEB*, in the context of IEEE 802.11 called *Contention Window* (*CW*), decreases only when the medium is not sensed *busy*. This drastically reduces the likelihood that multiple senders start transmitting at the same time.

[5] Modern Ethernet networks are built with switches and isolated full-duplex cables, which makes the *CSMA/CD* obsolete, but the specification is still valid for backward compatibility.

3.4.2.1 Distributed Coordination Function

The power distribution grid tolerates parallel charging processes up to a certain limit. A *collision* occurs whenever the current flowing through a power grid asset crosses a defined threshold, *e. g.*, transformer or cable overloading. This however cannot be detected in a fully decentralized manner,[6] but a detected *collision* can be broadcast to all affected charging locations, *e. g.*, all charging stations of the low-voltage power grid if the transformer is overloaded. The power distribution grid can be considered as *busy* (additional charging processes are not possible) if the loading of the power grid assets is near the *collision* threshold, *e. g.*, 10 % below the limit. If the power grid is neither sensed *busy* nor a *collision* is detected, it is considered to be *free* for charging.

Similar to the queuing network approach in Section 3.4.1, the time domain is split into discrete time slots, *e. g.*, 1 min. To ensure the timely starting of a charging process, the maximum mean delay time should be below 10 min, hence the smallest CW is defined by $[0, CW_{min} - 1]$ ($CW_{min} = 2^4 = 16$ min). The maximum CW is limited to $[0, CW_{max} - 1]$ ($CW_{max} = 2^8 = 256$ min), which is approximately one-quarter of the mean parking duration of the data set in Section 3.5.1. After arrival, a charging process is allowed to immediately start charging if the power grid is not sensed *busy*. Otherwise, the charging service must sample a uniformly random value from the CW interval. As long as the power distribution grid is not sensed *busy*, a charging service decreases its backoff timer every time slot until the charging process finally starts when the backoff timer reaches zero. The charging is kept active for at least a predefined *charging frame duration* (*e. g.*, 15 min) unless a *collision* occurs. If the power grid is sensed *busy* after that duration, the charging process stops and proceeds with a newly sampled random backoff from the CW. Otherwise, the charging process can remain to charge as long as the power grid is not *busy*. In both cases, pending charging services are allowed to reduce their random backoff timers, which offers them potential access to the power distribution grid. Retaining charging after the charging frame duration when the power grid is not *busy* improves the power grid utilization.

Compared to communication networks, it is not desirable that after a *collision* all charging processes stop immediately because this will most likely cause a voltage dip (short rise of the voltage), which may harm electrical equipment. Instead, each active charging process samples a discrete random variable $X \in \{0, 1\}$ with $P(X = 0) = P(X = 1) = 0.5$, which expresses an additional waiting time. The charging process stops immediately only if the random variable equals 0. Otherwise, it keeps charging for one additional time slot and re-evaluates whether the *collision* persists. If the *collision* is already resolved, the charging process proceeds to charge without interruption. Otherwise, the random variable X is sampled again. Overall, this ensures that all active charging processes together perform an exponential decrease of their charging current in case of *collision* events. Similar to networking, a detected *collision* during active charging exponentially increases the CW interval. Whenever a charging frame duration is charged without *collision*, the CW is halved. This implements an *exponential-increase-exponential-decrease* method for the CW.

[6] One may think of detecting voltage problems locally, however, the voltage does not necessarily correlate with asset overloading and, as a local phenomenon, it is not guaranteed that all charging locations form a consensus, which may result in discrimination of some charging locations.

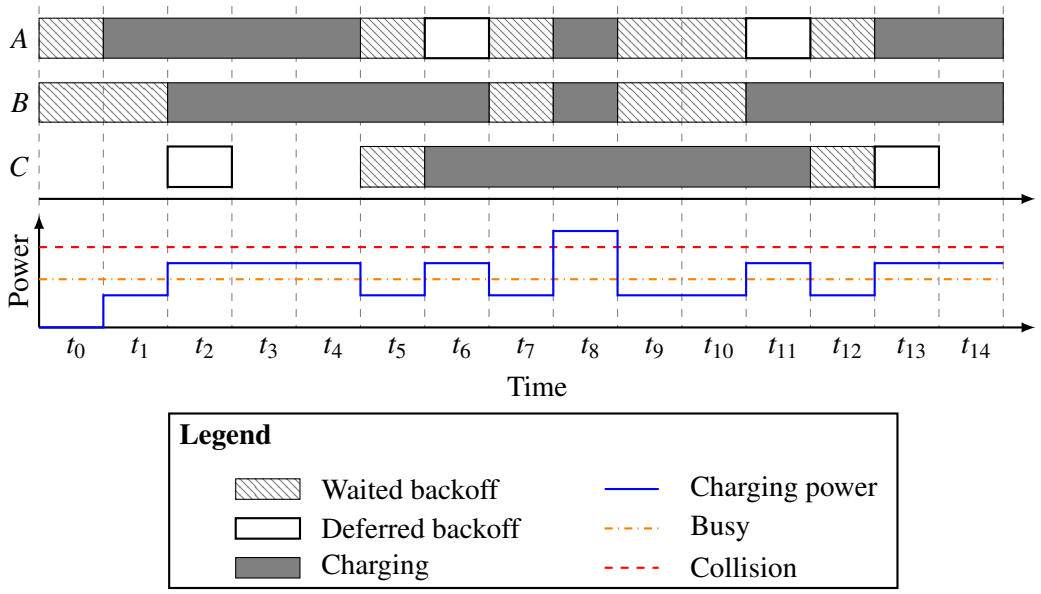


Figure 3.3: Example of the *DCF* for three *EV* charging services using *BEB*. Between t_2 and t_4 charging service *C* pauses to count down its random backoff because the power grid is sensed *busy*. At time t_8 a *collision* occurs, which forces *A* and *B* to stop charging.

In the following, the *DCF*-based probabilistic charging service allocation protocol is explained by an example, that has three pending charging services *A*, *B*, and *C*, where only two charging processes can be active simultaneously. The timeline of the events is depicted in Figure 3.3. It is assumed that the charging services *A* and *B* arrive before t_0 and still need to wait the remaining time from their random backoff from the initial assumed *CW* interval of $[0,3]$. At t_1, *A* starts charging and because the power grid is not sensed *busy*, *B* continues to count down its backoff time by one. The third charging service *C* arrives at t_2 and samples a random backoff time of 1 which is instantaneously deferred because *B* starts charging at that time and causes the power grid to be sensed *busy*. At time t_5, charging process *A* completes its charging frame duration of 4 time slots and stops charging immediately, because the power grid is sensed *busy*. This allows other charging services, like *C*, to resume decrementing their backoff timers. Additionally, charging process *A* samples a new random backoff of 2. Charging process *B* does not directly stop after the charging frame duration of 4 time slots, but continuous charging until starting of *C* causes the power grid to become *busy*. *B* samples a new random backoff of 1. Because the charging process of *C* is still charging at t_8 and the power grid is not sensed *busy*, *A* and *B* start charging simultaneously, which causes a *collision* (power threshold is crossed). As a result, all three charging services sample a random value from $\{0,1\}$, where *A* and *B* draw 0 (stop charging) and *C* draws 1 (keeps charging). Because the *collision* is solved by the reaction of *A* and *B*, and the power grid is not sensed *busy* anymore, *C* charges until the power grid is detected *busy* at t_{11}. After the *collision* at time t_8, all charging services exponentially increase their *CW* to $[0,7]$. Charging process *A* samples a new backoff time of 3 and process *B* samples 2, hence *B* starts charging at t_{11} before *A* starts at t_{13}. Because *C* can successfully finish its charging frame duration at t_{11}, its *CW* is reduced back to $[0,3]$.

Because *EVs* can control their charging power in discrete steps, *e. g.*, in 1 A steps per phase, any charging service may run several *DCF* procedures in parallel, where each procedure decides on whether to charge with a portion of the controllable charging current. The sum over all *DCF* procedures gives the overall charging current at the next time slot. If it is below the minimum charging current, the charging process needs to stop or will not start yet. As a result, the protocol supports variable charging currents, which fits the means of controlling an *EV* charging process as in Assumption A4.

3.4.2.2 Enhancement with Quality of Service Control

The *MAC* protocol extension in IEEE 802.11e adds means of controlling *QoS* to the wireless communication standard via *Enhanced Distributed Channel Access* (*EDCA*). Thereby, traffic is assigned to access categories with different priorities, among them *voice*, *video*, *best-effort*, and *background* traffic. Frames from prioritized access categories have a higher chance of being transmitted whenever the shared medium is *free*. This is guaranteed via shorter waiting times and smaller *CW* configurations. For example, voice frames are sent with the highest priority, because voice applications typically require low latency and voice interrupts highly impact the service quality. For video applications, the latency is less important because received frames are buffered for a short time before the video is finally decoded. Best-effort and background traffic, *e. g.*, sending E-Mails or downloading data, have the lowest priority because they are typically not time critical. This concept maps to the *CoS* introduced in Section 2.2.

In wireless communication, frames from different access categories need to wait a different amount of time, called *Arbitration Inter-Frame Spacing* (*AIFS*), after the shared medium is identified to be *free*. Access categories with higher priority wait for shorter *AIFS* and lower prioritized categories wait longer. The backoff time is counted down after the specific *AIFS* time of the access category has passed. To further improve the service quality, the minimum and maximum size of the *CW* is reduced, which consequently prioritizes frames from different access categories after *collisions*.

The priority of the charging service depends on its laxity, which may dynamically change over time depending on whether the charging process is active or not. Therefore, to distinguish the urgency of the different charging services, they are sorted into six *QoS* classes, similar to the four access categories in wireless communication. However, a charging service is not statically assigned to a *QoS* class, but the *QoS* class may dynamically change when the urgency of the charging service increases or decreases. The assignment to a *QoS* class is given in Equation (3.20) and is defined on how often the residual charging time (with assumed maximum possible charging rate and constant current charging) fits into the remaining available charging time. Additionally, the $\ln(.)$ function is used to better spread charging services with medium to low laxity to the lower *QoS* classes.

$$QoS_{class}(t) = \max\left(0, \min\left(\left\lfloor \ln\left(\frac{t_{dep}-t}{\left(\frac{E_{req}-E(t)}{P_{max}\cdot\mu_{EV}}\right)}\right)+1\right\rfloor, 5\right)\right) \tag{3.20}$$

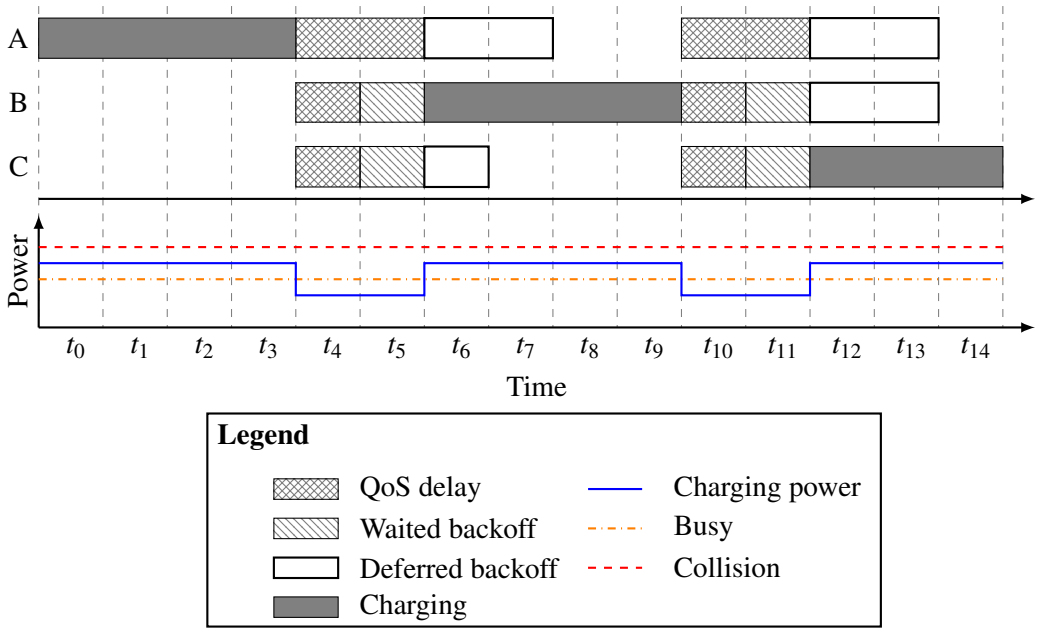


Figure 3.4: Application of *QoS* classes to the *EV* charging service allocation protocol.

Note that the *QoS* class is inverse to the weight function of the *DWFQ* policy in Equation (3.19). Hence, charging services with negative laxity will be in *QoS* class 0, charging services with small laxity smaller will be in *QoS* class 1, and so on. Charging services that are not actively charging may switch to a lower *QoS* class with higher priority, whereas active charging services reduce their residual charging time and may therefore switch to a higher, less prioritized *QoS* class. Once a charging service is in *QoS* class 0, it cannot change its *QoS* class anymore and will stay with the highest priority until departure.

Instead of waiting a specific *AIFS* time depending on the access category, a charging service simply waits an additional time equal to its *QoS* class, which gives a maximum additional waiting of 5 min. Furthermore, the enhanced *QoS* version of the *DCF* charging assignment protocol selects the minimum (aCW_{min}) and maximum (aCW_{max}) *CW* size depending on the current *QoS* class at time t, given in Table 3.2. Note that only the lower *QoS* classes obtain reduced *CW* window sizes, such that the upper waiting time limit after a sequence of *collisions* remains the same.

Figure 3.4 depicts an example with three charging services A, B, and C with different *QoS* classes, where only one charging process can charge at once. In the beginning, A is still charging until t_3, after which it waits its *QoS* class of 2. Because B and C have *QoS* class 1, they only wait one additional time slot at t_4 before counting down their backoffs. C reduces its backoff by one and defers the remaining time, while B waits for its remaining random backoff of 1 and starts charging in the next time slot t_6. After waiting for the *QoS* delay, A needs to defer the full backoff time in t_6. At time t_9, B completes its charging frame duration and all charging services again need to wait their *QoS* class, before they continue counting down their backoffs.

Table 3.2: Calculation of *QoS*-aware *CW* boundaries and *QoS* class delays.

QoS class	aCW_{min}	aCW_{max}	*AIFS* delay
0	$\frac{1}{4}\,CW_{min} = 4\,\text{min}$	$\frac{1}{32}\,CW_{max} = 8\,\text{min}$	0 min
1	$\frac{1}{4}\,CW_{min} = 4\,\text{min}$	$\frac{1}{16}\,CW_{max} = 16\,\text{min}$	1 min
2	$\frac{1}{2}\,CW_{min} = 8\,\text{min}$	$\frac{1}{8}\,CW_{max} = 32\,\text{min}$	2 min
3	$\frac{1}{2}\,CW_{min} = 8\,\text{min}$	$\frac{1}{4}\,CW_{max} = 64\,\text{min}$	3 min
4	$CW_{min} = 16\,\text{min}$	$\frac{1}{2}\,CW_{max} = 128\,\text{min}$	4 min
5	$CW_{min} = 16\,\text{min}$	$CW_{max} = 256\,\text{min}$	5 min

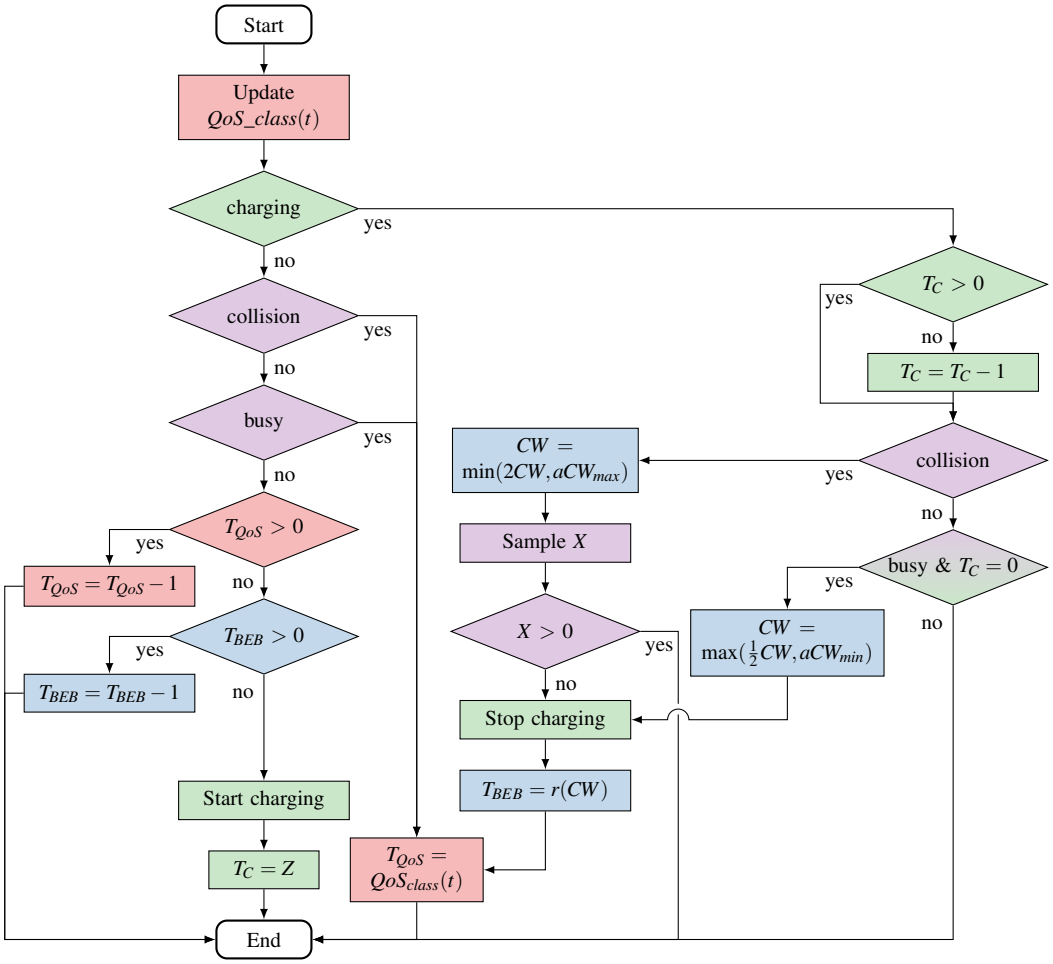


Figure 3.5: Sequence diagram of the probabilistic *EDCA EV* charging protocol. The different colors highlight the parts that are related to the power grid (purple), charging (green), binary exponential backoff (blue), and *QoS* control (red).

For the *QoS*-enhanced version of the *EV* charging control schema, the sequence diagram in Figure 3.5 is invoked at every time step of the charging service. It is assumed that every charging process has already chosen an initial random backoff time if the power grid is sensed *busy* at arrival. Otherwise, the backoff timer is set to $T_{BEB} = 0$. The timer for counting down the *QoS* class-specific *AIFS* delay is denoted by T_{QoS} and the timer for tracking the charging frame duration is denoted by T_C. A new random backoff time is uniformly sampled from the interval $[0, CW - 1]$ with the function $r(CW)$ and the charging frame duration is set to Z.

The communication effort of the distributed probabilistic protocols scales linearly with the number of *SUs* that are placed at critical assets in the power grid. Each *SU* needs to broadcast a signal (*free*, *busy*, or *collision*) to each *EV* that is located underneath the power grid. If this can be realized with broadcast messages, the required communication can be approximated by $\mathcal{O}(n)$, where n is the number of critical assets in the power grid.

3.4.3 Discussion

Some queuing policies share similarities with the probabilistic allocation protocols. For example, *PROP* aims to assign each *EV* charging service a proportional amount of charging capacity comparing the possible charging current of each charging process. A similar result is achieved by the *DCF* approach, where competing charging services receive on average a similar charging capacity. Thereby, both *PROP* and *DCF* utilize variable charging currents for the charging processes. Second, *DWFQ* uses the laxity of the charging service for prioritization, similar to *EDCA*, where charging services are grouped into *QoS* classes according to their laxity. Again, both methods use variable charging currents.

The major difference between the hierarchical queuing approach and the probabilistic protocols is in terms of required communication effort and computational power at the participating entities. Both methods require *SUs* in place to assign currents according to a policy (queuing) or generate *busy*/*collision* signals (probabilistic protocols). Obviously, in both cases, the *SUs* require local information as stated in Assumption A10, *e. g.*, current flow measurements. The main difference is the way of communicating the signals to the *EV* charging processes. The hierarchical queuing network performs a two-step *request-response* schema with bidirectional communication, where a charging request from the charging processes is forwarded to the uppermost *SU* before the response is sent back the request chain. Hence, the number of sequential messaging is linear to the longest chain of *SUs* to the root, which limits the reaction time of the algorithm. On the other hand, the probabilistic protocols only require one-directional communication from the *SUs* to the charging processes, which directly receive a congestion broadcast signal from any *SU*. Hence, the required sequence of communication is constant in $\mathcal{O}(1)$. In both methods, the computational requirement is rather low, *e. g.*, limited to charging packet switching and random number sampling. Such tasks can easily be accomplished by embedded hardware such as network switches or *EV* charging controllers. In addition, the time requirement of the calculation is low, since the charging current allocation is re-evaluated every minute when a new control signal is computed. In contrast to the reduced communication effort, the power grid utilization of the probabilistic methods is lower due to the waiting times.

It is worth mentioning that both methods rely on the *SUs* where congestion occurs. If these *SU* are not working correctly, *e. g.*, produce a wrong output or do not respond at all, this has a similar effect on both methods. In the case of the hierarchical queuing approach, a not responding *SU* will cause its sub-tree to not charge at all since no charging processes grouped underneath will receive a message with the assigned charging current. Similarly, without a *free* signal, charging processes cannot be sure that the shared power grid is ready for charging, and hence need to wait. If the messages/signals from the *SUs* are (maliciously) incorrect, *e. g.*, ignoring overload situations, the charging processes are not informed correctly to mitigate this event. Also, the logic of the charging service must be trusted. Hence, the integrity of all the participating entities (*SUs* and charging processes) and reliant communication must be ensured as stated in Assumption A10.

3.5 Evaluation

This section first describes the co-simulation setup for the evaluation, including the used charging pattern for *EV* charging services, the underlying low-voltage power distribution grid, and the connection of the different simulators. In the second part, the hierarchical queuing and probabilistic charging service allocation mechanisms are evaluated with regard to *QoS*, *QoE*, and fairness, as well as their impact on the power system. A sensitivity analysis of the simulation parameters finally demonstrates the impact of user inputs, power grid limitations, and the voltage controller.

3.5.1 Setup

The evaluation is performed via co-simulation of the different *EV* charging service allocation mechanisms, simulated *EVs*, and a simulation of the IEEE 906 low-voltage power distribution grid. All simulators are connected with the co-simulation framework *mosaik*[7].

3.5.1.1 Electric Vehicle Charging Pattern and Battery Model

In the literature, authors use different *EV* charging patterns to evaluate their control algorithms, which can be categorized as follows.

- One category uses *stochastic modeling* of charging data, where typical arrival and departure times are defined, and the charging services are sampled from a normal distribution around these times [107, 171]. These charging patterns are accurate for large scale simulations because they reflect the expected typical behavior around the chosen arrival and departure times. However, when the simulation scenario contains only a few *EVs*, probabilistic charging patterns are less suitable, because outliers, which are not modeled well, may have a huge impact on the results.

[7] https://mosaik.offis.de (version 2.6.1)

- The second category uses charging patterns based on *real data sets* [101, 102, 149]. These charging patterns are the most realistic, however, the circumstances of the data set must fit the scenario under investigation to avoid biasing the result. For example, public charging differs from *EV* home charging in terms of charging power rating (fast *vs.* slow charging), parking times (during the day *vs.* overnight), and charging flexibility (short charging stays with large energy requirement *vs.* longer charging stays with shiftable charging time). Additionally, early *EVs* have shorter driving ranges and are therefore typically used as secondary cars, which does not reflect the expected charging behavior with very high *EV* penetration. Several public and workplace charging data exists [16, 100, 122], however data for private households is less common, due to the lack of recording capabilities, or only reflects the mobility requirement of the recorded area [156].
- The third category uses *realistic data sets* generated from mobility surveys [35, 37, 148]. These typically nationwide surveys cover representative mobility patterns of households in different regions with different mobility needs. First, the survey data is filtered to fit the required means of transport and the regional scenario under investigation. Afterward, charging patterns are created from the driven distance, the parking time of the car, and the assumed *EV* characteristics including battery size, charging speed, and average consumption. These realistic data sets have both statistical relevance and possible outliers to be realistic even on a smaller number of charging patterns, *e. g.*, for studies on a low-voltage power distribution grid.

In this thesis it is assumed that most people will not (like to) change their driving behavior drastically when switching from combustion engine vehicles to *EVs* in the future. Therefore, data from the *Mobility Panel Germany*[8], which provides one-week survey data on travel behavior in Germany [47], serves as the basis for *EV* charging behavior. This data record contains 1757 surveyed households with more than 2000 individual trips in which the car is the main means of transportation. Each registered trip consists of trip sections with destination, time of departure and arrival, and distance covered.

The mobility survey data is filtered to fit the rural region of the investigated low-voltage power distribution grid, described in Section 3.5.1.2, and the trip sections are aggregated such that each aggregated trip starts and ends at home. The resulting data contains the arrival time at home t_{arr}, the departure time from home t_{dep}, and the distance d of the last trip before arriving at home. It is assumed that the vehicle will not be charged if the stay between two trips is less than 1 h, due to convenience reasons of the *EV* driver, and the corresponding distance is added to the next trip. It is further assumed that all *EV* drivers want to recharge the consumed energy of the last trip during their stay at home, which is reasonable because home charging is typically less expensive than public charging. Assuming a battery storage capacity of 40 kWh[9] and an average energy consumption of

[8] Mobility Panel Germany: MOP 2016/17, *Bundesministerium für Verkehr und digitale Infrastruktur* (English: Federal Ministry of Transport and Digital Infrastructure)

[9] Compared to the battery capacities offered by *EV* manufacturers, 40 kWh seems rather small, however, the author of this thesis believes that for most of the trips, the advertised huge driving range of more than 400 km is not required. This estimation is underpinned by the fact that more than 98 % of the trips in the data set can be driven with the assumed battery capacity.

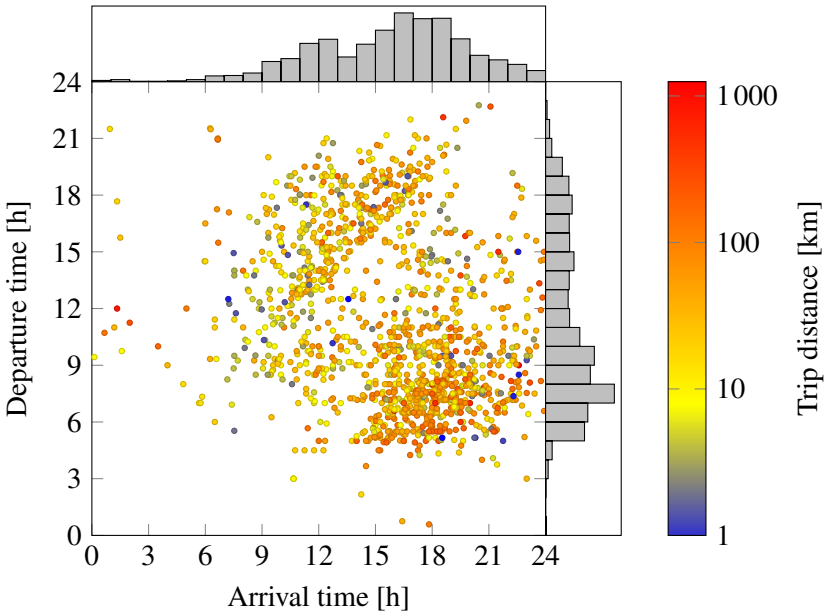


Figure 3.6: Commuting behavior of the charging patterns. The majority of *EVs* arrive around 18:00 and leave at 7:00, which can be seen from the histograms beside the axis. Additionally, some *EVs* are arriving around noon with shorter stays. *EVs* that have an earlier departure time than arrival time in this plot, depart on the following days.

17 kWh per 100 km, the required energy of the charging service, which is upper limited by the battery capacity, can be estimated using Equation (3.21). In case the battery capacity of an *EV* is not large enough to cover the whole trip distance, it is assumed that the driver visits a public charging station during the trip, where only the required additional energy is charged. With this worst-case assumption, the *EV* will arrive home with an empty battery and requires a full charge cycle. The departure times are assumed to be strict deadlines, hence *EV* that missed charging E_m kWh requires more energy in the next charging service.

$$E_{req} = \max(40\,\text{kWh}, d \cdot \frac{17\,\text{kWh}}{100\,\text{km}} + E_m) \tag{3.21}$$

The number of charging services per *EV* ranges between 0 and 15 per week. The mean parking duration is approximately 16.6 h and the mean driving distance is equal to 39.1 km per trip. As can be expected, many commuting *EVs* reach home between 17:00 and 18:30 and need to leave between 6:30 and 8:00 on the next day, which can be seen in Figure 3.6. In addition, the data set also contains 13.2 % shorter charging stops with less than 3 h, which arrive almost normally distributed around noon.

The batteries of the *EVs* are modeled using a constant current battery model according to Assumption A5. However, other battery models are theoretically possible, which would require a more advanced estimate of the remaining charging time. Both the queuing models and the probabilistic protocols support variable charge requests to handle variable charging power requirements from the battery model in their charging packet requests and try-error

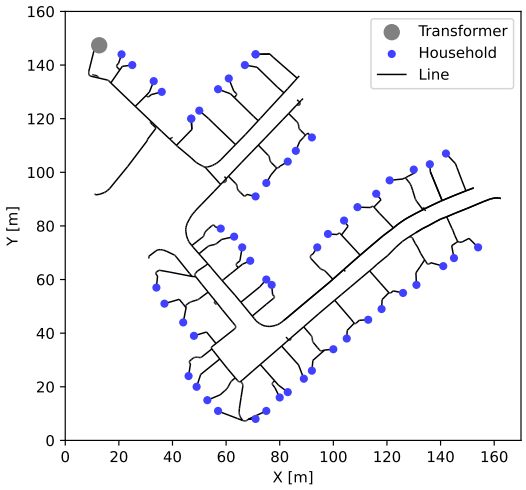


Figure 3.7: IEEE 906 low-voltage test feeder, where the transformer is located in the top left side and 55 households are connected to the power distribution grid.

approaches, respectively. Furthermore, following Assumption A3, the constant charging efficiency is set to 95 %, and the minimum adjustable charging current is given by 3 A. Additionally, the minimum charging current of the *EV* is limited by 18 A (6 A per phase), similar to the control capabilities in IEC 61851-1 as described in Section 2.3. To encode the minimum charging current, the first queued packet of each charging service has a packet size of 18 A. All following packets have the minimal adjustable size of 3 A and are only assigned after the first packet according to the queuing logic. The distributed allocation strategy only starts charging if enough sub-procedures decide to start. The threshold for *busy* is configured to be 10 % below the transformer limitations. Consequently, the absolute difference between *busy* and *collision* is greater than the minimum charging current of an *EV* for all experiments, such that starting a single *EV* cannot jump from *free* to *collision*.

3.5.1.2 Power Grid and Simulation Scenarios

The evaluation is carried out on the simulated IEEE 906 low-voltage test feeder. This typical European low-voltage power distribution grid, shown in Figure 3.7, connects 55 households on a three-phase power system. Since balanced *EV* charging is assumed, the power flow simulation in PyPower connects all households balanced to all three phases. Furthermore, the load profiles of the households have a power factor of 0.95 inductive and each household owns two *EVs*[10] that can charge in parallel at two 22 kW wallboxes.[11] The aforementioned charging patterns are randomly assigned to the *EVs*.

[10] Households in rural areas in Germany own an average of 1.64 vehicles in 2020 [143]. Hence, two *EVs* is a rather conservative assumption that reflects a very high *EV* penetration rate.

[11] Note that according to NAV §19 (*Niederspannungsanschlussverordnung*; English: Low-voltage connection regulation) wallboxes with up to 12 kVA must only be registered to the power grid operator in Germany, while above 12 kVA approval by the power grid operator is required.

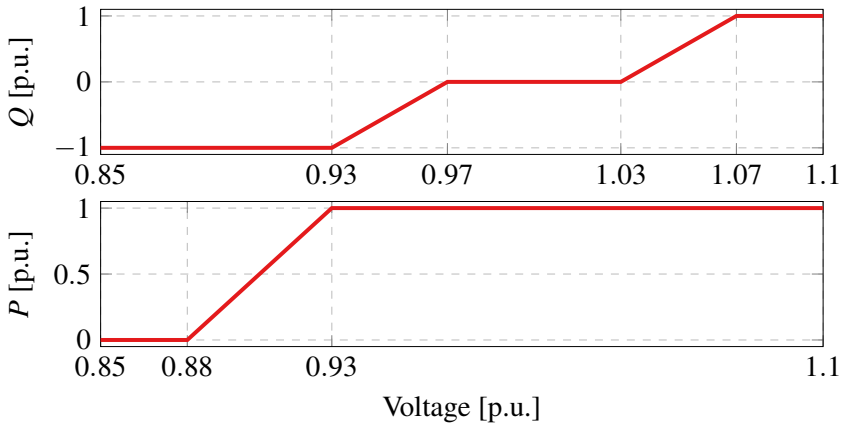


Figure 3.8: Voltage controller according to VDE-AR-N 4100. Real and reactive power limits depend on the locally measured voltage value [35].

In the baseline scenario without *EV* charging, the low-voltage power grid has a peak load of 60.5 kVA at the transformer, which is substantially smaller than the total installed charging capacity of 2.4 MVA. Nevertheless, uncontrolled charging with 22 kW and the aforementioned charging patterns results in a peak load of approximately 312 kVA due to the simultaneity factor. Because this would increase the peak load at the low-voltage transformer by more than five times, which applied to many low-voltage power grids can cause critical peak loads in the superior power grid infrastructure, the maximum loading at the transformers' *SU* is limited to 100 % of its baseline peak.

The proposed queuing and probabilistic allocation mechanisms act as load management. To additionally react to voltage violations, a linear voltage droop controller that uses reactive power control capabilities according to Assumption A6 is implemented as in Figure 3.8. This decentralized voltage controller changes the reactive power behavior of the rectifier and in critical situations even reduces the real power demand of the *EV*. To avoid grid losses and keep the reactive power ratio in the low-voltage power grid in a reasonable range, the voltage controller implementation is configured to have always a power factor better than 0.9. To not exceed the provided current capacity from the allocation mechanism, besides the adaption of the reactive power behavior, the real power demand is slightly reduced between 0.93 p.u. and 0.97 p.u. (1.03 p.u. and above respectively). Furthermore, the reactive power decreases with the real power demand below 0.93 p.u. to stay with the defined minimum power factor. To avoid oscillations, a *first-order lag filter* is applied to the control signal as given in Equations (3.22) and (3.23).

$$P(t) = k \cdot \hat{P}(t) + (1-k) \cdot P(t-1) \tag{3.22}$$

$$Q(t) = k \cdot \hat{Q}(t) + (1-k) \cdot Q(t-1) \tag{3.23}$$

$\hat{P}(t)$ and $\hat{Q}(t)$ are the target signal value from the voltage controller limited by the assigned charging current from the allocation mechanism. The factor k has to be configured to avoid oscillations, but still reach the target signal value within the desired time. The co-simulation

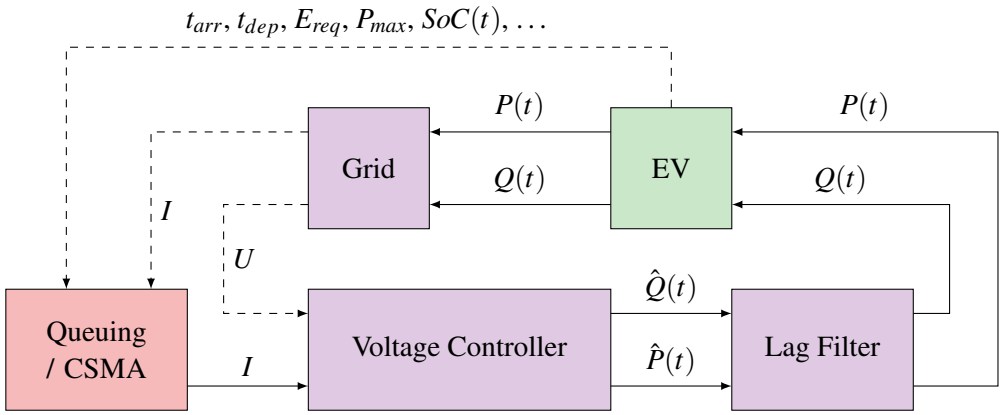


Figure 3.9: Control flow in the co-simulation. The different colors highlight grid-related simulators (purple), charging (green), and *QoS* control (red). The dotted connections are time-shifted and respective information flow is used from the last simulation step.

steps in one-minute steps and the target value should be reached the latest after five steps to provide a fast enough reaction, therefore k is set to $1 - \frac{1}{e} \approx 0.632$.

Figure 3.9 visualizes the control flow of the co-simulation. First, the allocation mechanisms start with allocating charging current I to the single *EVs* based on their charging service parameters $(t_{arr}, t_{dep}, E_{req}, P_{max}, SoC(t), \ldots)$ and the calculated available charging capacity coming from the power grid simulator. Before actually controlling the charging processes, the voltage controller calculates the real and reactive power values of the charging service $\hat{P}(t)$ and $\hat{Q}(t)$ based on the local voltage measurement U. The first-order lag filter stabilizes the control signal before the parameters are finally sent to the *EV* model, which charges the *EV* according to its battery model and passes the values back to the power grid simulation. The updated charging service parameters are finally passed to the allocation mechanisms in the next simulated time step. Similarly, the available current capacity and node voltage from the power grid simulator close the control loop, which is executed every minute.

3.5.2 Analysis

All the following results are obtained from a 7-day week simulation of the low-voltage power grid with the aforementioned scenario setup and assumptions. During this simulated week, the individual *EVs* require between 1 and 13 charging services, on average 5.06. The mean energy demand of a charging service is 6.68 kWh, which is approximately 16.7 % of the assumed battery capacity. The total energy demand of all occurring 557 charging services is 13.2 % bigger than the total baseline demand of the households.

All the following results are obtained by simulations with 22 kW wallboxes, hence an upper charging current limitation of 96 A, and transformer limitation at 100 % of the baseline scenario, hence an upper current limit at the transformer of 263 A. In all the simulations, the limiting *SU* was the transformer, hence it is valid to compare all charging services with each other. Other configurations are investigated in the sensitivity analysis in Section 3.5.3.

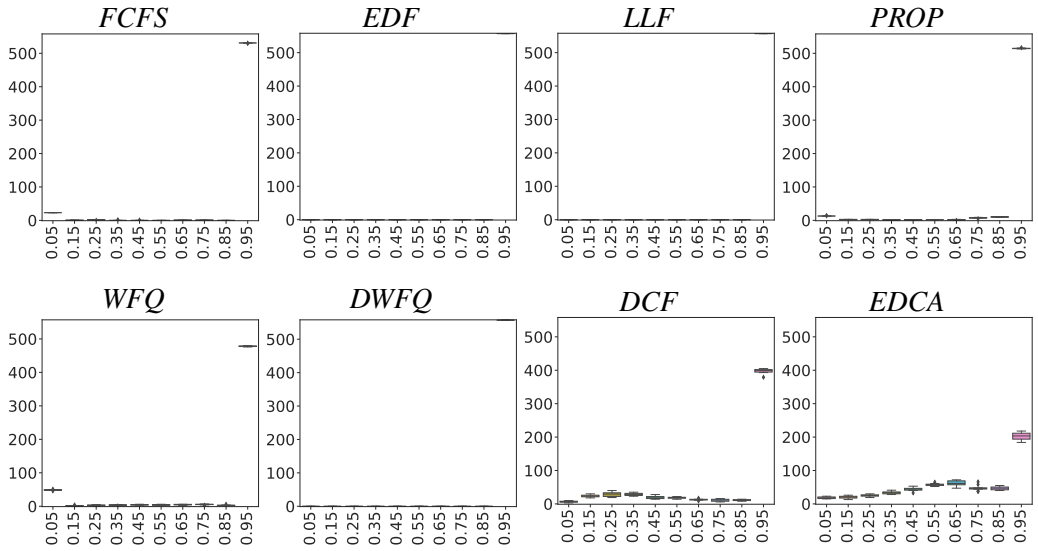
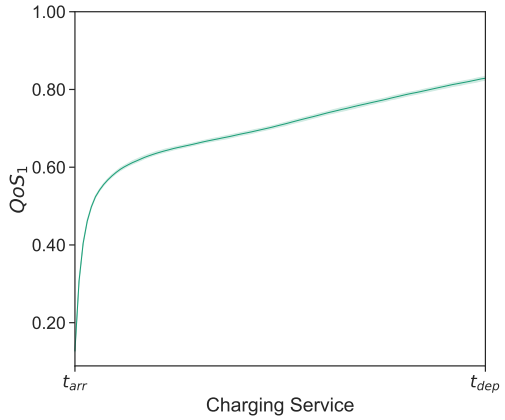


Figure 3.10: Histogram in box plot style that shows the impact of randomness on the QoS_1 metric from $N = 10$ independent simulation runs. Note that for *EDF*, *LLF* and *DWFQ* all charging services reach the maximum QoS_1 value, hence the plots seems to be empty.

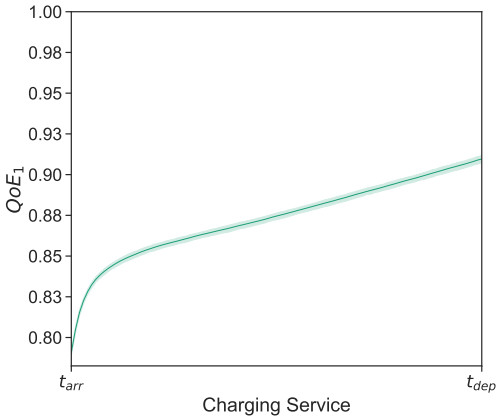
3.5.2.1 Impact of Randomness

The simulation scenario is repeated ten times with random charging profile assignments to the different households. Furthermore, the probabilistic charging allocation mechanism uses different initial random seeds for the random number generation. Figure 3.10 shows the histogram of achieved QoS_1 over all charging services in box plot style to summarize the variance of the metric values over multiple random simulation runs. The random location of the charging services in the power grid does only slightly impact the deterministic queuing allocation mechanisms, *e. g.*, only very small variations are experienced with *PROP* and *WFQ*. The two probabilistic allocation mechanisms, *DCF* and *EDCA*, experience slightly higher impact from the random seeds, however, the values of the different simulation runs are still very dense. Because charging control is done on a timescale of 1 min, which is much smaller than the average available charging time of the charging service, every charging service needs to sample many random backoff values and, hence, the sampled random numbers reflect the uniform probability distribution well. At the end, this leads to stable service quality among all charging services. The remaining *QoS* and *QoE* metrics show similar results.

The two metrics QoS_1 and QoE_1 can be computed for each time step during the charging service and the mean metric values with a shaded 95 % confidence interval are shown in Figure 3.11. It can be seen that even during the charging service there are only minor differences. In conclusion, the random selection of the *EV* mobility demand profiles nor the random number sampling of the exponential backoff algorithms in the probabilistic allocation protocols have a critical impact on the results. Therefore, the following analysis is performed on the results of only a single simulation run.



(a) Mean QoS_1 metric values from $N = 10$ independent simulation runs using *DCF*.



(b) Mean QoE_1 metric values from $N = 10$ independent simulation runs using *DCF*.

Figure 3.11: Impact of randomness on the mean QoS_1 and QoE_1 metric values throughout the charging service runs using *DCF*. The very dense shaded green area is the 95 % confidence interval.

3.5.2.2 Quality of Service/Experience and Fairness

First, the obtained metric values for the quality of service and experience of the different queuing policies and probabilistic allocation mechanisms are analyzed. Figure 3.12 shows box plots over all 557 charging services, where the mean metric values are indicated by black circles. Below the box plots, the fairness index F from Equation (3.13) is depicted for each metric and policy.

All policies of the hierarchical queuing networks achieve high QoS_1 values that are close to 1.0 for most charging services, however, the number and variation of outliers vary significantly among the queuing policies. *EDF*, *LLF*, and *DWFQ* achieve the maximum quality, while *FCFS*, *PROP*, and *WFQ* have slightly lower mean values due to many outliers. There even exist some charging services that do not receive any service at all ($E(t_{dep}) = 0$) and, hence, obtain $QoS_1(t_{dep}) = \frac{E(t_{dep})}{E_{req}} = 0$. That is the case when charging services are blocked by other charging services that have higher flexibility but are prioritized by the applied queuing policy, *e. g.*, *FCFS* blocks charging services that arrive later. The two probabilistic protocols *DCF* and *EDCA* achieve mean metric values between 0.6 and 0.8, but experience much greater variation among the different charging services. This can be explained by the lack of direct coordination, since the probabilistic protocols do not take a coordinated decision in advance, but base their charging decisions on measured *busy* and *collision* signals, which reduces the overall efficiency of these two methods. The *QoS* class delay of *EDCA* causes additional waiting times, where no *EV* is charging even though the power grid is not sensed *busy*. This further reduces the efficiency of the *EDCA* compared to the more simplistic *DCF* approach, which is reflected by a lower mean QoS_1 value. Obviously, *EDF*, *LLF*, and *DWFQ* achieve a maximum fairness index of 1.0, because all charging services receive equally maximum service. Since *WFQ* uses the *SoC* for prioritization, many charg-

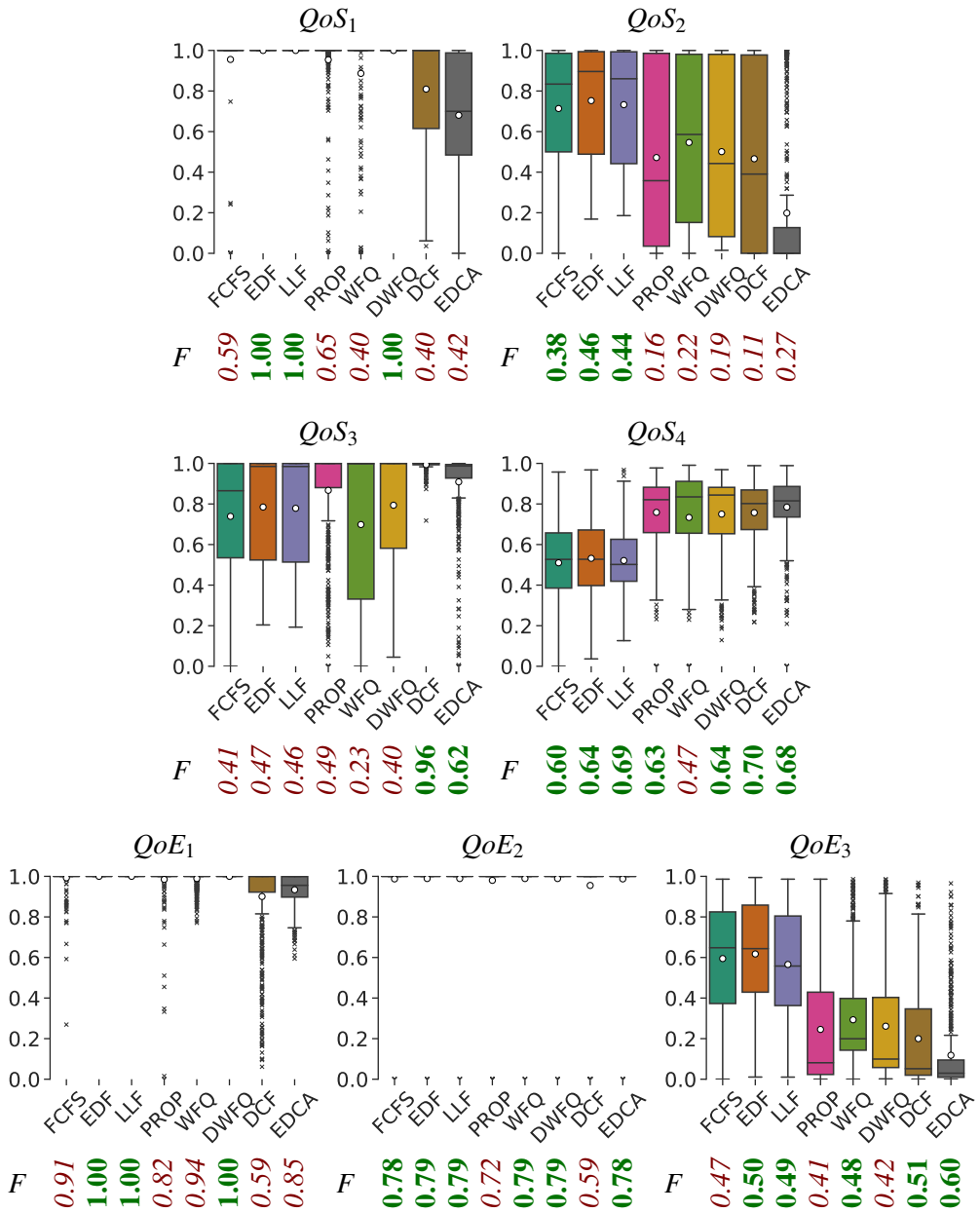


Figure 3.12: *QoS* and *QoE* metrics of the different allocation policies for all 557 charging services. The box plots show the resulting distribution among the charging services and the circle denotes the mean value. Below the box plots, the achieved fairness index F is given. Note that for QoS_1, QoE_1 and QoE_2 most of the charging services have very high service quality, hence the boxes are very near to 1.0.

ing services with already high SOC at arrival receive no energy at all and, hence, obtain a QoS_1 value equal to zero. This results in a low fairness index of QoS_1.

Because QoS_2 depicts the finish time of a charging service, it is most relevant for the queuing policies *EDF*, *LLF*, and *DWFQ*, which finish all charging services in time. From these three policies, *EDF* and *LLF* have higher mean metric values (earlier finish times) with smaller variation (better fairness) than *DWFQ*, because of their prioritization nature. With *DWFQ* policy, charging services that can finish in time receive a very low weight (below 1). Hence, *DWFQ* does not focus on early finish times but provides each charging service a share of the available charging current throughout the charging service. Note that the QoS_2 metric highly depends on the charging service. Finishing 10 min before departure yields a better QoS_2 value if the total available charging time is smaller. Not finished charging service receive $QoS_2 = 0$, therefore the QoS_2 metric of the other queuing policies and the two probabilistic approaches range from 0 (not finished) to 1 (finished). Because *EDCA* only finishes 126 out of 557 charging services, its QoS_2 values are very low.

The two metrics QoS_3 (starting time) and QoS_4 (power variation) target QoS during the charging time. All queuing policies except *PROP* have comparable QoS_3 values with a slightly larger quartile box for *WFQ*, which results in a lower fairness index. The weight definition of *WFQ* is based on the *SoC* at arrival time instead of the energy requirement of the charging service, which delays starting of the charging services with an already high *SoC*. *PROP* targets equal charging current for all parallel charging services and, therefore, offers early starting times to most of the charging services. The two probabilistic methods are not aware of other parallel charging services and operate by probing whether charging is possible. This contributes to early charging start times. For QoS_4, the mean metric value of the first three queuing policies (*FCFS*, *EDF*, and *LLF*) is noticeably lower than with the other methods. This is because the latter also enables variable charging currents, whereas the first three queuing policies operate as purely time-division multiplexing by only switching the charging process on and off, which increases the variation in charging power. Nevertheless, the fairness index is comparable between all approaches, except for *WFQ*, which achieves a very low fairness index because many charging services with high *SoC* at arrival are not charged at all, which results in QoS_4 being equal to 0.

Similar to QoS_1, *EDF*, *LLF*, and *DWFQ* achieve maximum quality in QoE_1, which measures the charging service concerning the battery *SoC*. It can be seen that *EDCA* improves over *DCF* by sorting the charging services into QoS classes, which results in a drastically reduced number of outliers. That is because *EDCA* prioritizes charging services with higher energy requirements, which results in overall better *SoC* at departure time. This can also be seen in the improved fairness index compared to *DCF*. Since a fully charged battery is not sufficient to reach the next destination for some trips, the QoE_2 metric cannot reach the maximum value for any allocation mechanism. Therefore, all allocation mechanisms work almost equally well, except *DCF* and *PROP*, since they do not apply any preference from the charging service requirements.

Opposite to QoS_4, the QoE_3 metric is higher with the queuing policies *FCFS*, *EDF*, and *LLF*, because they either charge with a maximum charging power or do not charge at all. *EDCA* has a very low metric value because the introduced QoS classes create additional

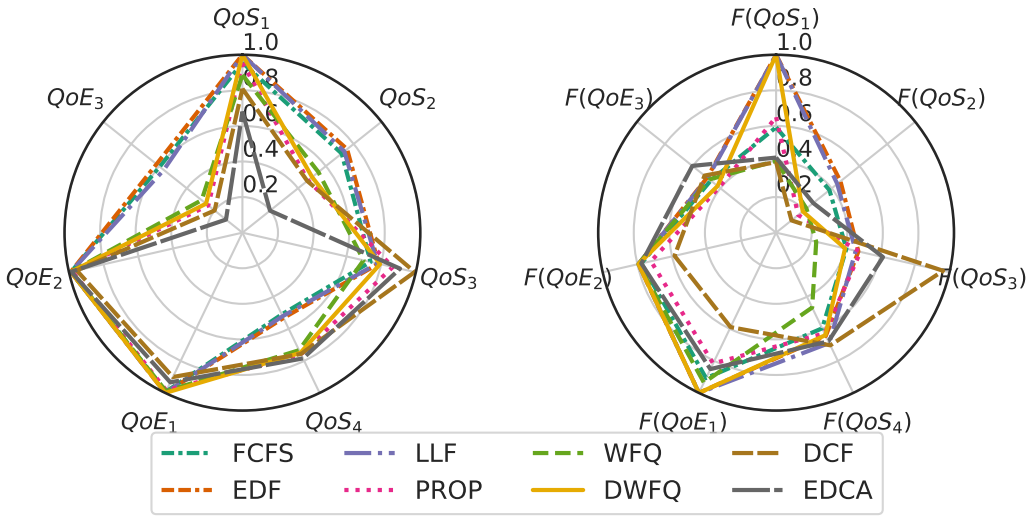


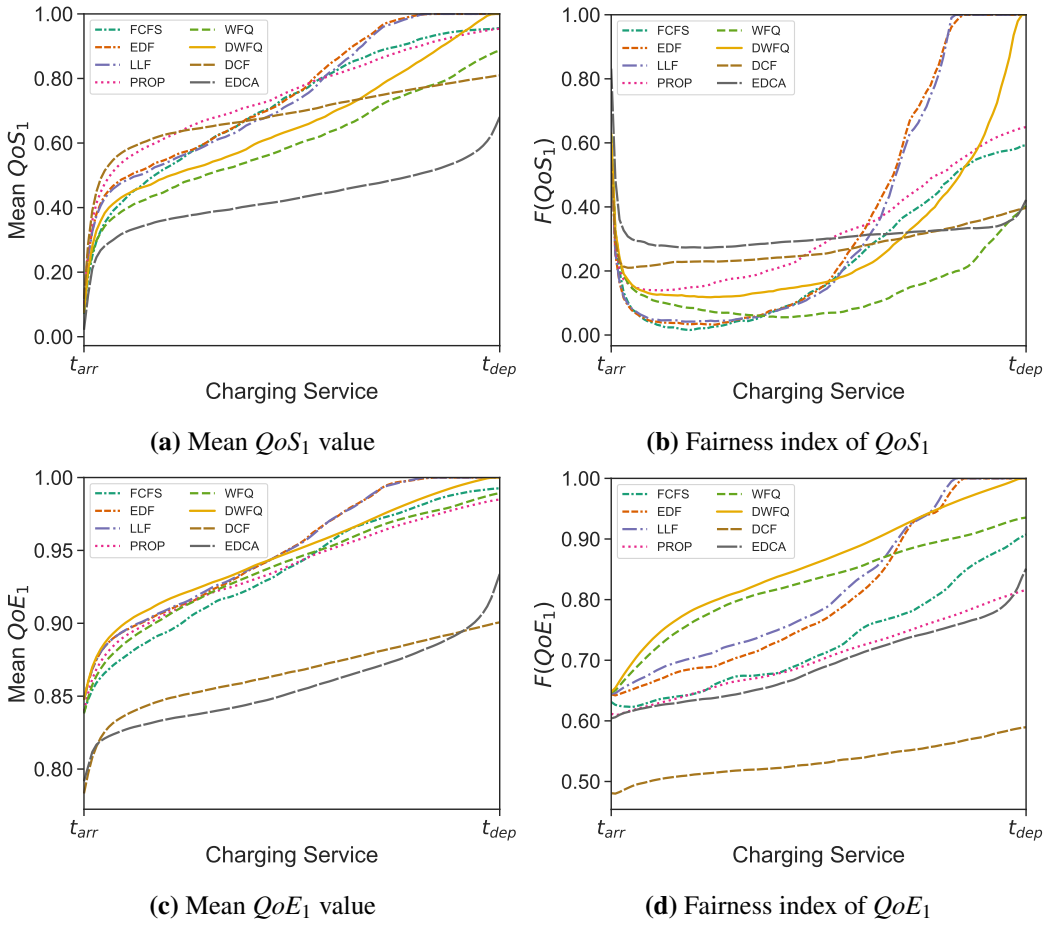
Figure 3.13: Mean *QoS* and *QoE* metric values and fairness index of different allocation mechanisms.

delays, which in turn reduces the mean charging power during the charging process. Note that the fairness index is independent of the mean metric value. This can be seen in the case of *EDCA*, which obtains the best fairness index, but has the lowest mean metric values.

Figure 3.13 compares the mean *QoS* and *QoE* values and the fairness indices obtained by the different allocation mechanisms. The spider plots clearly show that none of the methods outperforms the others in all metrics. However, there exist some allocation mechanisms that are more suitable than others as they achieve better metric values in a subset of the metrics. The three queuing policies without variable charging currents, namely *FCFS*, *EDF*, and *LLF*, perform well in QoS_2 and QoE_3, however, achieve low values in QoS_4. From the probabilistic methods, *EDCA* has slightly lower mean metric values compared to *DCF*, except for a larger gap in QoS_2, where a higher number of charging services do not finish. However, the fairness indices of *EDCA* are much better, except for QoS_3, which is due to a higher number of charging processes that wait longer to start because of their *QoS* class delay. *PROP* and *DCF*, which do not consider charging service parameters, have lower fairness indices for most *QoS* and *QoE* metrics, even though they aim for equally distributed fair resource sharing among connected charging processes. Hence, the in-cooperation of departure times plays an important role in charging service allocation due to the better performance of *EDF*, *LLF*, and *DWFQ*.

3.5.2.3 Quality Metrics during the Charging Services

The two metrics QoS_1 and QoE_1 can also be measured at every point in time during a charging service. Figure 3.14 depicts the evolution of the QoS_1 and QoE_1 mean values, as well as their fairness indices during the charging services. The x-axis is normalized to range between arrival time t_{arr} and departure time t_{dep} of all charging services, which makes



(a) Mean QoS_1 value

(b) Fairness index of QoS_1

(c) Mean QoE_1 value

(d) Fairness index of QoE_1

Figure 3.14: Mean QoS_1 and QoE_1 metric values and fairness index during the charging services using different allocation mechanisms. All charging services are normalized to the range $[t_{arr}, t_{dep}]$.

them comparable on the same timescale, even though the charging services have different duration and do not take place at the same time. This re-sampling represents the view of the *EV* owner on the charging services.

Figure 3.14(a) shows the evolution of the mean QoS_1 value during the charging services. *PROP* and *DCF* have higher mean values in the first half of the charging service because both aim for equal power-sharing among charging processes, where those with small energy requirement E_{req} benefit with early high QoS_1 metric values. Likewise, the fairness index is reasonably high during this time, as can be seen in Figure 3.14(b). It can also be identified that *EDF* and *LLF*, which focus on early finishing of the charging process, reach the maximum QoS_1 value approximately after 80 % of the time of the charging services. However, they experience quite low fairness indices during the first half of the charging services. *DWFQ* achieves a nearly linear increase of the mean QoS_1 value during the charging service, while the fairness index is best among the queuing policies that reach maximum QoS_1 at least during the first half of the charging services. For *EDCA* one can observe a

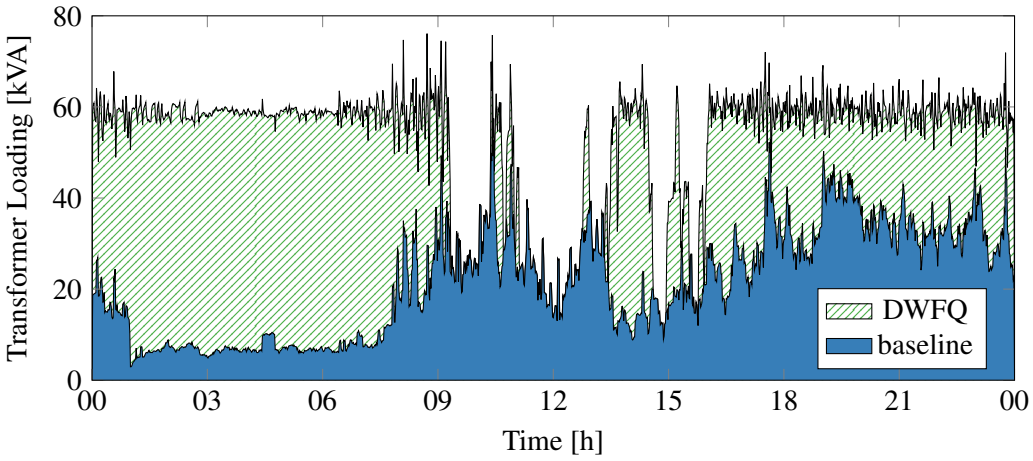
slight increase in mean QoS_1 value and fairness towards the end of the charging service, which results from the charging service obtaining a smaller *QoS* class and, therefore, having a higher probability to charge.

From Figure 3.14(c), it can be seen that the mean QoE_1 value of all queuing policies evolves quite similarly during charging services, except for *EDF* and *LLF*, which finish their charging processes earlier. Compared to the other policies, *FCFS* has a slightly lower value during the first half of the charging services, because newly arrived *EVs* are blocked until all previous charging services are fully served. Only *EDF*, *LLF*, and *DWFQ* finally reach the maximum at departure time. Because *EDF* and *LLF* schedule only maximum charging power to the most critical charging services with regard to time and remaining available charging time, both reach the maximum metric value earlier than *DWFQ*. In contrast, *DWFQ* focuses on a fair allocation throughout the whole charging service, which results in a higher mean quality metric during the first half of the charging service. Although *DWFQ* has a slightly lower mean QoE_1 value at the last third of the charging service, most of the time this policy dominates the fairness index shown in Figure 3.14(d) except at the very end. With a higher fairness index during the first half of the charging service, *EVs* are served more fairly in case they need to leave earlier than the planned departure time. Furthermore, it can be expected that *DWFQ* (and also *WFQ*) are more robust against malicious user inputs (*e. g.*, incorrect departure times), because opposite to *EDF* all charging services always obtain a portion of the available charging capacity according to their weight. This aspect is further investigated in Section 3.5.3.4. The mean metric values of the probabilistic protocols are nearly constantly 0.10 points smaller due to lower effectiveness. Despite *EDCA* having a slightly lower mean QoE_1 value during most of the charging service, its fairness index is much better than with *DCF* allocation.

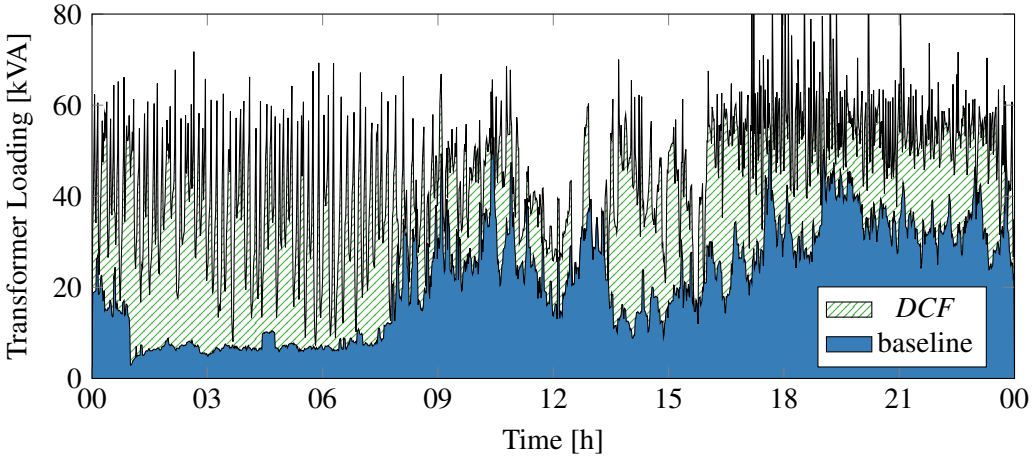
3.5.2.4 Impact on the Low-Voltage Power Grid

DSOs are concerned with the impact of the different allocation mechanisms on the low-voltage power grid in terms of power quality and grid losses. Regardless of which policy is used for the hierarchical queuing model, the configured threshold in all simulations smoothly limits the transformer loading. Figure 3.15(a) shows the transformer loading of a single day; once the baseline scenario without charging and once with the *DWFQ* policy. Except for some short spikes, the transformer is smoothly limited at around 60 kVA, which is similar to the peak load of the baseline at 10:25. As can be seen, the spikes occur when the baseline load has abrupt changes, *e. g.*, during the power ramp-up between 8:00 and 9:00 in the morning. Since the *SUs* decide on the available charging current only based on the past measurements, these baseline spikes transfer to the aggregated load profile and balance out in the next algorithm iteration within 1 min. This effect could be dampened by using a good short-term prediction model instead of a persistence model like used in the simulation. All the other queuing policies have very similar transformer load profiles.

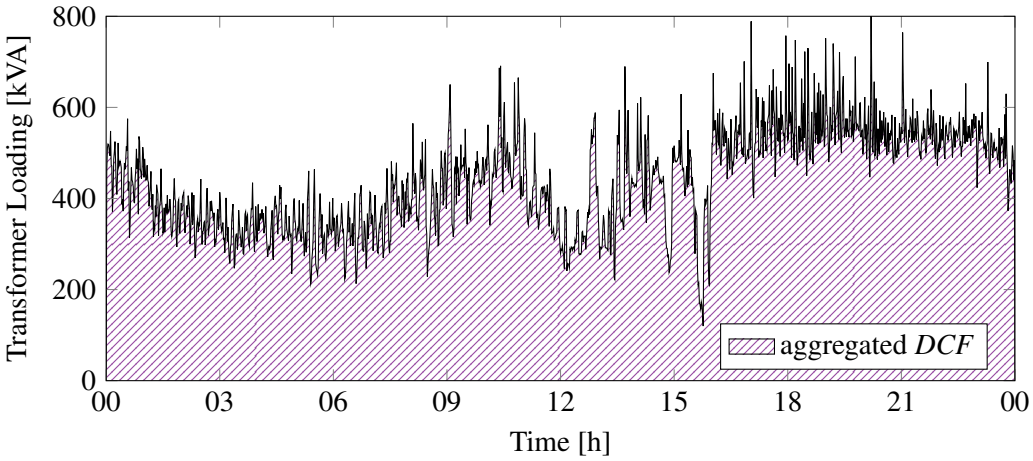
Figure 3.15(b) shows the transformer loading using the probabilistic *DCF* protocol. The fluctuations at the transformer are drastically higher than with the queuing approach. This is due to the randomized *MAC* protocol that does not have a global coordinator, and spikes occur due to randomized probing of the network. Note that the upper transformer limit



(a) Baseline and central queuing model using *DWFQ* policy.



(b) Baseline and probabilistic methods using *DCF*.



(c) Aggregated transformer loading using *DCF* from ten independent simulation runs.

Figure 3.15: Transformer loading during a simulated day.

Table 3.3: Charging statistics and impact on the low-voltage power grid. U_{10min}^{avg} is the 10-minute average voltage value according to EN 50160.

Allocation mechanism	Mean SoC (QoE_1) [%]	Min SoC (QoE_1) [%]	Min U_{10min}^{avg} [V]	Grid Losses [%]
FCFS	*99.24*	*25.31*	*216.27*	*6.87*
EDF	**100.00**	**100.00**	*216.99*	*6.88*
LLF	**100.00**	**100.00**	*217.68*	*6.75*
PROP	*98.46*	*0.00*	**219.34**	**5.90**
WFQ	*98.89*	*77.09*	**220.76**	**5.99**
DWFQ	**100.00**	**100.00**	**220.05**	**5.99**
DCF	*90.08*	*4.10*	**222.03**	**5.42**
EDCA	*93.64*	*56.80*	**221.73**	**5.55**
Baseline	-	-	231.55	0.00
No control	100.00	100.00	176.05	13.06
U-control	100.00	100.00	203.40	12.09

is respected, except for some short violations in the evening when many *EVs* arrive home. These spikes cause the *collision* counters to increase the *CW* size of the charging processes, which in turn reduced the number of spikes throughout the night. The probabilistic protocols have lower efficiency in providing energy for the charging service, which is reflected by the recurrent under-loading of the transformer, especially in the first third of the day.

Numerous spikes and their deep valleys seem to be undesirable from a power grid operator's point of view. However, this effect only occurs on a single low-voltage transformer. Applying the probabilistic protocol to a set of transformers in the same medium voltage power grid, one can expect a more uniform load profile, similar to Figure 3.15(c), which shows the sum of the transformer load profiles of the ten independent simulation runs with different seeds. The graphs for *EDCA* look similar and are provided in Figure A.1.

Table 3.3 summarizes the achieved mean and minimum *SoC* at departure time (expressed by QoE_1) of the different allocation mechanisms and also provides power grid statistics extracted from the power flow simulation. The minimum of the 10-minute average voltage values (as defined in EN 50160) from all buses is taken as an indicator for the voltage impacts of the allocation mechanisms. The grid losses are calculated by comparing the charged energy of all charging services with the additional energy that passes the transformer. Values from the baseline scenario without charging (*Baseline*), uncontrolled charging (*uncontrolled*), and only using the aforementioned local voltage controller (U*-control*) are given in Table 3.3 for comparison.

All allocation mechanisms reach an acceptable voltage level, but the three queuing policies with variable charging rates (*PROP*, *WFQ*, and *DWFQ*) improve the voltage level by more than 2 V compared to the other queuing policies. This is because the charging capacity is shared among more charging services, with each receiving a smaller share, thereby reducing the voltage drop at the charging locations. The minimum 10-minute average voltage

values of the two probabilistic protocols are higher by more than another 1 V, however, they come along with high voltage fluctuations due to the probabilistic probing.

In addition, variable charging rates (*PROP*, *WFQ*, and *DWFQ*) reduce grid losses that are caused due to *EV* charging by approximately 1 %, which over the simulated week is equal to the yearly energy consumption of one household (assumed 2000 kWh). Similar applies to the probabilistic protocols. Note that this study does not consider reduced charging efficiency of *EV* charging equipment at lower charging rates as stated in Assumption A3. Because *DWFQ* achieves maximum *SoC*, higher minimum 10-minute average voltage value, and low additional grid losses due to *EV* charging, it outperforms the other queuing policies in terms of service and power quality.

3.5.3 Sensitivity of Configuration Parameters

This section analyzes the impact of different configuration parameters on the proposed system. Among others, experiments investigate the effect of different scenario setups in terms of wallbox power rating, the number of *EVs*, or the transformer limitation. Furthermore, the impact of the local voltage control on the *QoS* and fairness of the charging service is inspected in a separate simulation setup. Finally, the effect of earlier departure times on service quality and fairness is analyzed.

3.5.3.1 Wallbox Power Rating

The results from Figure 3.12 in metric QoE_3 suggest that most of the charging services do not fully utilize the charging capability of the 22 kW wallboxes, especially with the methods that allow variable charging currents. To assess the impact of the wallbox peak power, the same simulation runs are repeated with the maximum charging power at the wallboxes limited to 11 kW instead of 22 kW. As can be expected, the main difference is a general improvement of the QoE_3 value (wallbox peak power utilization), especially for the allocation mechanisms with variable charging currents. On the other side, the difference between the variable and non-variable charging current strategies is reduced in metric QoS_4, which measures the variation of the charging currents. The other metrics only show minor changes and are summarized in Figure A.2.

None of the allocation mechanisms could fully serve all charging services with wall box rating of 11 kW, where the best-performing policies *EDF*, *LLF*, and *DWFQ* have one single charging service with a *SoC* of around 93 % at departure. Obviously, the maximum transformer loading is not influenced by the charging capacity of the wallboxes. However, according to Table 3.4, the minimum 10-minute average voltage value shows an improvement of more than 2 V, except for the *WFQ* policy. This is due to the limited charging power that can be drawn at each node in the power grid, which decreases punctual loads and increases the number of parallel charging processes. Using an appropriate charging capacity allocation mechanism, wallboxes with 11 kW peak power are sufficient to serve almost all charging services, even with a high *EV* penetration and a transformer peak limited to 100 % of the baseline load.

Table 3.4: Charging statistics and impact on the low-voltage power grid of the different allocation mechanisms and queuing policies with 11 kW wallboxes (difference to 22 kW is shown in brackets). U^{avg}_{10min} is the 10-minute average voltage value.

Allocation mechanism	Mean SoC (QoE_1) [%]	Min SoC (QoE_1) [%]	Min U^{avg}_{10min} [V]	Grid Losses [%]
FCFS	*99.09* (↓0.15)	*19.63* (↓5.68)	*220.40* (↑4.13)	*6.00* (↓0.87)
EDF	**99.99** (↓0.01)	**93.16** (↓6.84)	*219.92* (↑2.93)	*6.07* (↓0.81)
LLF	**99.99** (↓0.01)	**93.13** (↓6.87)	*220.33* (↑2.65)	*6.02* (↓0.73)
PROP	*98.28* (↓0.18)	*0.00* (↓0.00)	**222.10** (↑2.76)	**5.60** (↓0.30)
WFQ	*98.83* (↓0.06)	*77.09* (↓0.00)	**221.23** (↑0.97)	**5.74** (↓0.25)
DWFQ	**99.99** (↓0.01)	**93.02** (↓6.98)	**222.42** (↑2.37)	**5.75** (↓0.24)
DCF	*86.47* (↓3.61)	*1.26* (↓2.84)	**224.39** (↑2.36)	**5.01** (↓0.41)
EDCA	*93.02* (↓0.62)	*51.88* (↓4.92)	**224.66** (↑2.93)	**5.27** (↓0.28)

3.5.3.2 Transformer Limit and Vehicle Penetration

For all the simulations above, the number of *EVs* is fixed to a rather high *EV* penetration of two *EVs* per household. Furthermore, the *SU* at the transformer is limited to 100 % of the baseline load. To investigate the effect of both parameters on the performance of the allocation mechanisms, additional simulations with different numbers of *EVs* (0.5, 1, 2 *EVs* per household) and transformer limits (80 %, 90 %, 100 %, 110 %, 120 %, 140 %, and 160 %) are performed.

As can be seen in Figure 3.16, a smaller number of *EVs* and higher transformer limits do not impose any problems to finish all charging services, which is depicted by the QoS_1 metric. It can even be expected that with unlimited transformer (and cable) limitations, all policies work exactly similarly since all charging requests can directly be served and the power grid will never be sensed *busy*. However, increasing *EV* penetration leads to all charging policies except for *EDF*, *LLF*, and *DWFQ* not fully serving the charging services at high transformer limits. This is due to a sub-optimal charging capacity allocation to competing charging services during parallel charging processes. On the other side, a low number of *EVs* with small transformer limits mainly cause problems for *FCFS* and *WFQ*. One reason is the blocking behavior of *FCFS*, which results in short charging service to be not served at all. The *WFQ* policy suffers from static weight assignment at arrival time that is only based on the *SoC*. Charging services with high *SoC* receive low service rates independent of their available charging time. With both a high number of *EVs* and a low transformer limit, all allocation mechanisms result in reduced service quality due to the mismatch of energy demand and supply. However, service degradation is smaller for *EDF*, *LLF*, and *DWFQ*.

LLF and *DWFQ* have lower mean QoS_1 values than *EDF* because these two policies use laxity for prioritization. Compared to *EDF*, a lower priority is assigned to charging services with short duration and small energy requirements. As a result, more than twice the number of unfinished charging services with already high *SoC* can be counted at departure with *LLF*

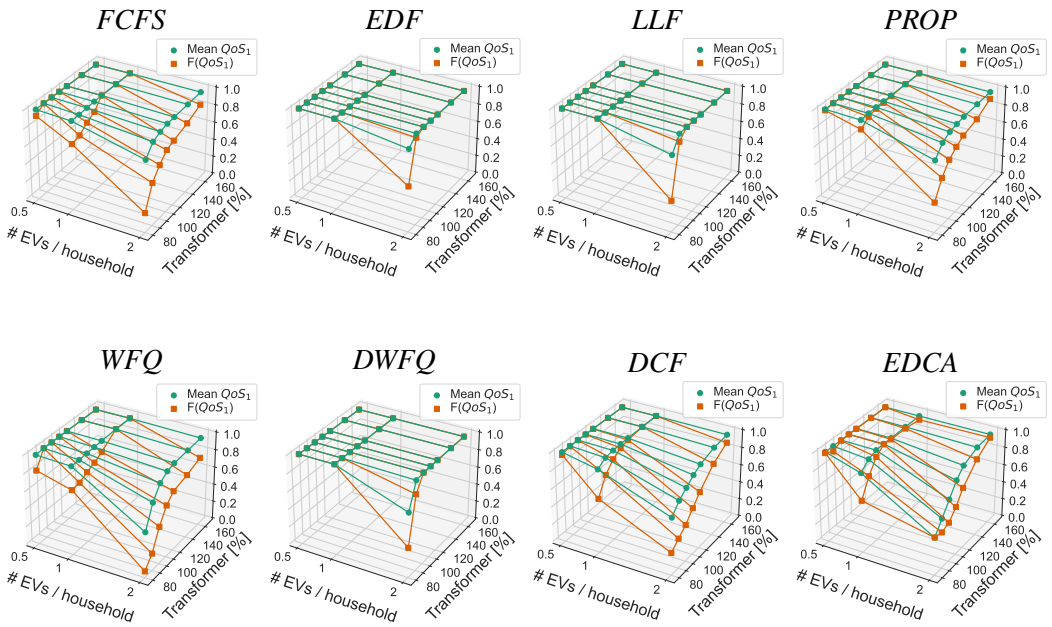


Figure 3.16: Impact of number of *EVs* and transformer limit on the mean QoS_1 metric value and its fairness.

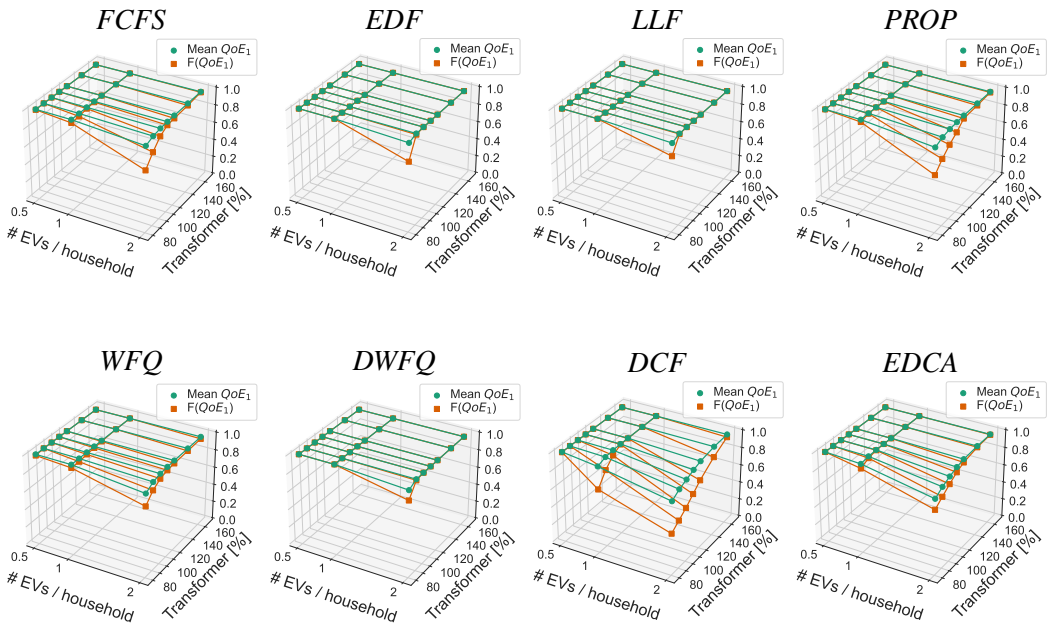


Figure 3.17: Impact of number of *EVs* and transformer limit on the mean QoE_1 metric value and its fairness.

and *DWFQ* policy. However, looking at the mean QoE_1 metrics, these charging services achieve a higher metric score compared to *EDF*, which is reflected by the fairness index in Figure 3.17. It is worth mentioning, that *DWFQ* outperforms the other policies at all *EV* and transformer limit configurations while reaching a high mean metric score.

It can be seen that the laxity (*LLF* and *DWFQ*) plays an important role in an optimal allocation with restricted resource availability in terms of QoS_1 and QoE_1. Additionally, the purely deadline-driven *EDF* policy performs well, however, results in a few charging services with very low *SoC* at departure, which impacts its fairness. Hence, the best queuing policy should be chosen to ensure high quality of charging service and fairness even with highly limited scenarios.

3.5.3.3 Local Voltage Control

The charging service allocation mechanisms in Sections 3.4.1 and 3.4.2 focus on *QoS* and fairness of the charging current assignment to single charging processes and only implement a load management solution. However, the voltage values in the power grid are kept within allowed boundaries using a local voltage controller at the charging stations. Undervoltage mitigation is achieved by supplying reactive power to the power grid, which slightly reduces active power usage to respect the assigned charging current limitations. Ultimately, the charging real power reduces to zero in case of voltage values below 0.88 p.u.. This however can cause the charging service to receive lower *QoS* and, therefore, biases fairness among the charging services.

To demonstrate the effect of the different charging strategies on *QoS* and fairness in presence of local under-voltage, an artificial charging scenario is defined based on a simple four-node power grid, depicted in Figure 3.18. This power grid contains only two charging locations (EV_1 and EV_2) and no further load or generation. A standard (over-sized) 800 kVA transformer and two 70 mm^2 copper cables with lengths 100 m and 400 m connect the four buses. The values are chosen to experience a voltage drop at charging station EV_2, such that only EV_2 needs to supply reactive power during charging. The scenario further schedules one charging service for each *EV*, with similar charging requirements, but slightly different arrival and departure time. EV_1 arrives at minute 5 after the simulation start and departs at minute 205. EV_2 arrives at 15 and leaves at 195, hence has a 20 min shorter stay. Both *EVs* have a battery of size 40 kWh with a charging efficiency of 100 %, can charge with up to 22 kW, and arrive with 10 % of *SoC*, hence having very similar laxity.

In case there is no charging current limitation at the transformer, all charging service allocation mechanisms behave similarly by allowing both *EVs* to charge with their maximum charging current. However, depending on the voltage value, the local voltage controller may impact the actual charging rate of the charging process. Consequently, local voltage threshold violations may reduce the *QoS* and, therefore, decrease fairness among the charging services. This effect cannot be avoided by the allocation mechanisms without knowing and optimizing the local voltage values, which is not possible under Assumption A8, therefore only the current limitations are considered by design in the *QoS*-aware load management.

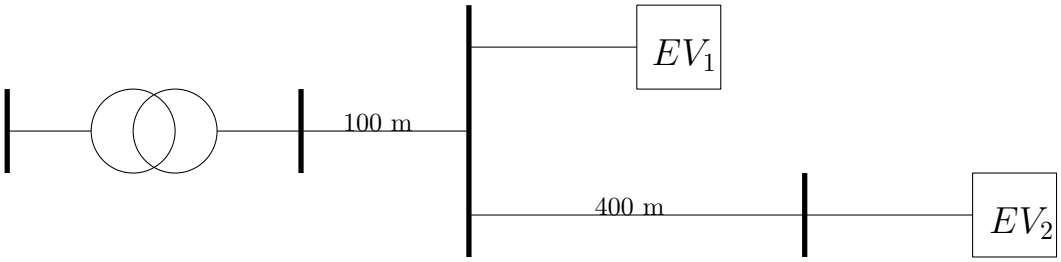


Figure 3.18: Simple low-voltage power grid for investigation of voltage effects on the charging allocation mechanisms.

In the following experiment, the transformer is limited to only support one *EV* charging at a time, hence the available charging current is limited to 22 kVA (32 A per phase). The resulting *SoC* curves of the two *EVs* are depicted in Figure 3.19. It can be seen that with fully sequential charging (*EDF* or *FCFS*) the slope of the *SoC* curve of EV_2 is slightly smaller than for EV_1 due to the required 8.32 kVA reactive power injection, which limits the active charging power to 17.17 kW (26 % reduced maximum charging power compared to EV_1). As a result, only one *EV* will be fully charged. The small difference of the final *SoC* in Table 3.5 is because with *FCFS* policy no *EV* charges after 3:15. *LLF* prioritizes the incoming EV_2 due to the slightly smaller laxity until both *EVs* reach the same laxity. From then on, both *EVs* are charged sequentially in chunks, such that they keep almost the same laxity. This iterative switching between the *EVs* compensates for the lower charging efficiency of EV_2 due to reactive power supply, such that both *EVs* reach very similar *SoC* at departure. All other policies charge both *EVs* in parallel with reduced charging power. This results in a smaller voltage drop at EV_2 and, therefore, overall less required reactive power. *PROP* and *WFQ* policies perform nearly equally because due to equal energy requirements, both receive the same weight, which is equal to the proportional allocation of the charging power. Because the maximum reactive power requirement from EV_2 is only –2.59 kVA, the charging real power only differs by 7 %. As the stochastic variant of *PROP*, *DCF* performs very similarly, but with in total lower efficiency because the transformer is not utilized with 22 kVA all the time. *DWFQ* assigns EV_2 an 18 % higher charging current compared to EV_1, because of its slightly smaller laxity. Furthermore, the charging power increases throughout the charging processes to compensate for the lower efficiency of EV_2 due to the reactive power supply. A similar effect can be observed with *EDCA*, where the charging rate of EV_2 increases towards the end of the charging service, where the lower charging efficiency due to reactive power causes EV_2 to change to a lower *QoS* class.

Table 3.5 summarizes the findings of the impact of voltage on the QoE_1 metric. It can be seen that only *LLF*, *DWFQ*, and *EDCA* reach a high fairness index, while among them *DWFQ* achieves the best single metric values for both *EVs*. Furthermore, the voltage values are better for the allocation mechanisms with variable charging currents, which leads to reduced reactive power requirements and therefore higher efficiency.

In case *EVs* do not support reactive power injection contrary to Assumption A6, only the active power reduction according to Figure 3.8 can be applied to mitigate under-voltage situations. This however amplifies the difference in the maximum possible charging power

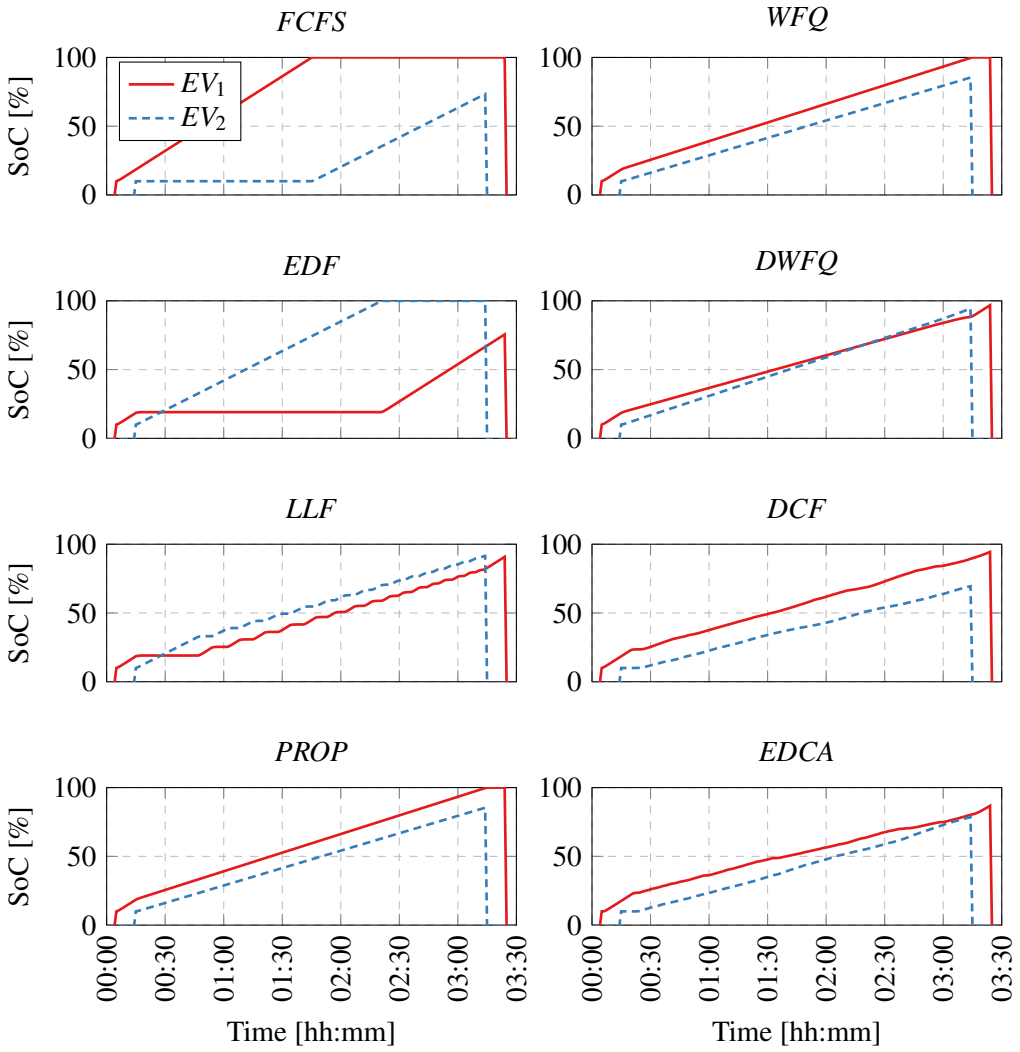


Figure 3.19: Behavior of different charging strategies in under-voltage events.

Table 3.5: Effect of voltage on the charging service allocation mechanisms.

Allocation mechanism	EV_1 (QoE_1) [%]	EV_2 (QoE_1) [%]	$F(QoE_1)$ [%]	Min U^{avg}_{10min} [V]	Q at EV_2 [kVA]
FCFS	**100.00**	*73.35*	*62.31*	*213.13*	*−8.32*
EDF	*75.56*	**100.00**	*65.43*	*213.13*	*−8.32*
LLF	**90.85**	**91.45**	**99.15**	215.14	*−8.32*
PROP	**100.00**	85.38	79.33	**217.87**	**−2.59**
WFQ	**100.00**	85.37	79.30	**217.87**	**−2.59**
DWFQ	**96.78**	**94.66**	**97.01**	**217.32**	**−3.35**
DCF	**94.41**	*69.42*	*64.66*	**217.01**	**−1.10**
EDCA	86.76	*78.53*	**88.37**	**217.22**	**−1.63**

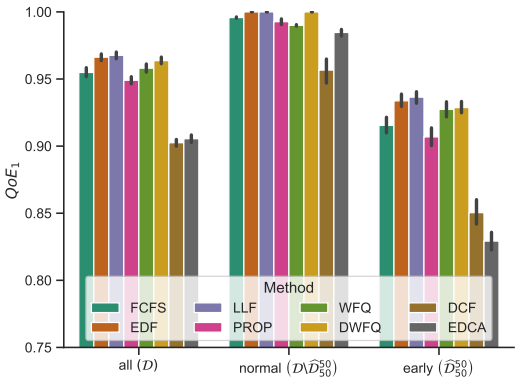
between the two charging processes. Note that in any case, the assigned charging current capacity is only the allowed upper limit for the *EV* charging process. First, because local voltage may force to reduce the charging power and, second, because the communication protocols between the charging station and *EV* usually only define upper limits, which may not be reached, *e. g.*, due to battery temperatures.

3.5.3.4 Departure Time

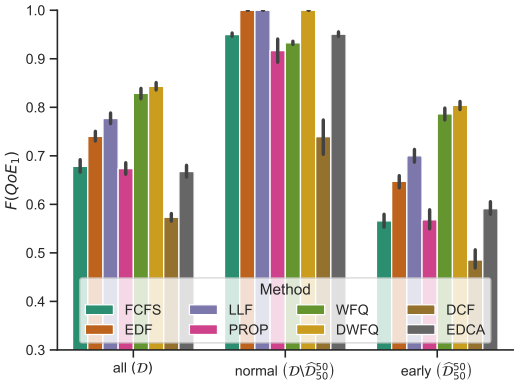
In Assumption A7 and Section 3.5.1.1, it is assumed that the departure time of a charging service is known at arrival, and it is a strict deadline. It is further assumed that *EV* drivers will most likely visit a fast charging station on their next trip if the *SoC* of the *EV* is not sufficiently high. This is more time efficient than waiting for a full charge at home, which is usually only slow charging with mode 1 or 2. As already seen in Section 3.5.2.3, the time-dependent service metrics highly differ among the policies during the charging service. The following experiments investigate the service quality of charging services when *EV* drivers leave earlier than their planned departure time. Therefore, a subset $\widehat{\mathcal{D}} \subset \mathcal{D}$ of 50 % (20 %) of all *EV* drivers $\mathcal{D}$ is selected randomly to depart at 50 % (80 %) of their available charging time. Let $\widehat{\mathcal{D}}_c^d$ denote the subset of drivers with cardinality $|\widehat{\mathcal{D}}_c^d| = \frac{c}{100} \cdot |\mathcal{D}|$ that depart at d% of the planned charging stay. Again, ten independent simulation runs are executed with different seeds to investigate the impact of random subset selections.

First, it can be seen that the random selection of the early departing *EVs* from the ten independent simulation runs has little effect on the results, as illustrated by the dense 95 % confidence interval in Figure 3.20. Second, the mean QoE_1 metric in Figure 3.20(a) is degraded for all early departing *EVs* in subset $\widehat{\mathcal{D}}_{50}^{50}$, regardless of which charging service allocation mechanisms are used. Comparing the three policies with the overall best mean metric values (*EDF*, *LLF*, and *DWFQ*), the mean service quality degradation of early leaving *EVs* is very similar. However, the fairness index in Figure 3.20(b) shows the advantage of the packet weighting approaches, where both *WFQ* and *DWFQ* perform well. This is because all charging processes receive a fair share of the available capacity at every point in time during the charging service according to their weight and no charging service is starving. This effect is even amplified with a lower number of early departing *EVs*, as can be seen in Figure 3.20(c). On the contrary in Figure 3.20(e), if 50 % of the *EVs* depart at 80 % of the available charging time, *EDF* and *LLF* reach a higher mean *SoC* and fairness index than *WFQ* and *DWFQ*. This is already implied in Figure 3.14, where *EDF* and *LLF* finish most charging services at around 80 % of the available charging time. Similar results are obtained with a lower number of *EVs* that leave earlier, which is given in Figure A.3.

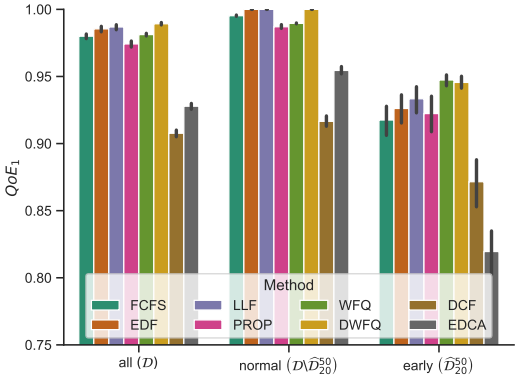
One possible way to improve charging service quality for *EV* drivers with untrustworthy user input is to include reputation in the charging power allocation algorithm, like in [6]. Their optimization-based approach integrates the reputation of how accurately the user predicts its departure time in the allocation prioritization. However, in reality, it is not that simple to detect whether a user entered the wrong departure times or simply was parking longer than planned. In addition, *EV* drivers that typically predict well (usually overnight, where there is enough energy available) can still game the system for a single charging service, which will not sustainably reduce their reputation.



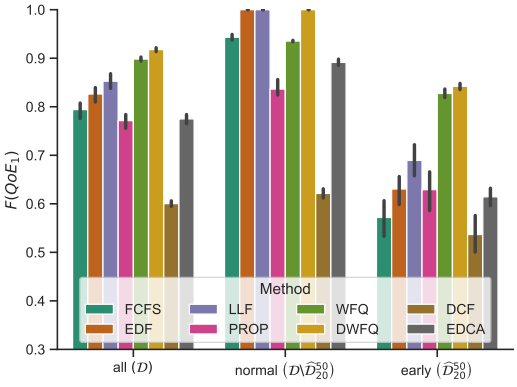
(a) Mean QoE_1 metric value for $\widehat{\mathcal{D}}_{50}^{50}$.



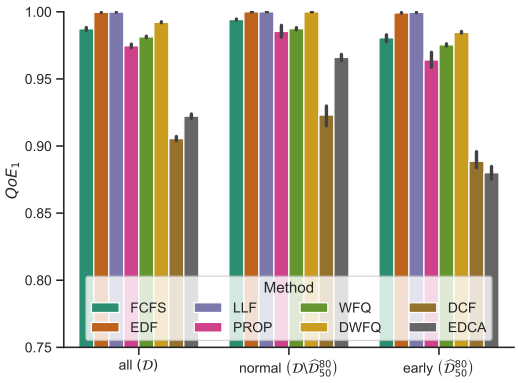
(b) Fairness index of QoE_1 for $\widehat{\mathcal{D}}_{50}^{50}$.



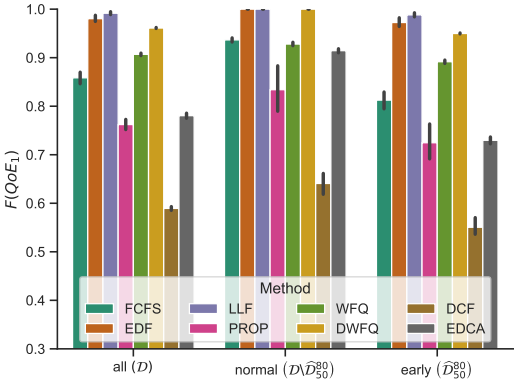
(c) Mean QoE_1 metric value for $\widehat{\mathcal{D}}_{20}^{50}$.



(d) Fairness index of QoE_1 for $\widehat{\mathcal{D}}_{20}^{50}$.



(e) Mean QoE_1 metric value for $\widehat{\mathcal{D}}_{50}^{80}$.



(f) Fairness index of QoE_1 for $\widehat{\mathcal{D}}_{50}^{80}$.

Figure 3.20: Impact of earlier departure on the charging service quality.

3.6 Applicability

This section deals with the practical implementation of the queuing network from Section 3.4.1 and the probabilistic protocols from Section 3.4.2, respectively. Thereby, possible application scopes are introduced, to which the proposed approaches may be deployed in real-life. Furthermore, potential technical and legal challenges are discussed, as well as ideas for further extension.

3.6.1 Scope of Application

The charging service allocation mechanisms allow limiting charging capacity resources (infrastructure as well as energy availability) while ensuring the quality of charging service to the *EV* driver. This offers potential for different applications in the power grid, where three are discussed in the following.

- The charging service allocation mechanisms can be applied to **avoid or postpone power grid infrastructure expansion** in the low-voltage power distribution grids, while ensuring high quality of *EV* charging service and, therefore, acceptance from the users. Grid expansions are usually required in older low-voltage power grids that are built in the 1980s and before, without large *PV* or *EV* penetration considered in the planning horizon. The proposed charging service allocation mechanisms can increase infrastructure utilization, while ensuring that *EV* charging does not harm the power grid stability in terms of overloading and voltage problems. Power grid expansion can thus be avoided or at least postponed if rapid increase in *EVs* cannot be accommodated by sufficiently rapid power grid expansion.

  The proposed approaches require measurement devices in place at the critical assets of the power grid and the software components, *e. g.*, the *SUs*, can be deployed in a cloud computing environment, where only communication connections to the measurement device and the wallbox controllers are required. The cost for measurement and cloud computing infrastructure is usually less than the expected power grid expansions, as shown by the European project *ELECTRIFIC* for a very similar setup [131]. As a positive side effect, continuous power grid measurements provide insights into the actual power grid situations to the *DSOs*, which usually operate their low-voltage power grids without monitoring. This may help to better understand the operation of their low-voltage power grids and may lead to an improved future power grid planning process. However, an *ICT*-based solution as an alternative to power grid expansion carries the risk of cyberattacks that can not only disturb service operations but also damage the power grid infrastructure.

- The allocation mechanisms can also be used to **align *EV* charging demands with renewable energy generation** by reducing the available charging current at times with low renewable energy generation. This can be achieved in the queuing network by modifying the algorithm at the root *SU* to reduce the available charging current at times when low renewable generation is available in the power grid. In the case of

the probabilistic protocols, respective *busy* and *collision* signals can be sent at lower threshold levels. For a smooth transition, first, the *busy* threshold is reduced for at least a charging frame duration, before the *collision* threshold is adapted. Increasing the total available charging current is done in the opposite order. Such a control strategy has the potential to increase the use of renewable energies for *EV* charging while considering the quality of charging service during charging current assignment.

Note that the use of local renewable energy is indirectly strengthened even without variable charging current control at the transformer. Local generation reduces the transformer loading, which in turn increases the available charging current in the low-voltage power grid. As a result, more *EVs* can charge during the availability of local renewable energies. In energy communities with high local renewable generation, the proposed charging service allocation mechanisms can be used to achieve high autarky by configuring the transformer threshold to zero. This will only charge the *EVs* using surplus energy from the energy community while the charging service allocation ensures high quality of service and fairness among the community members.

- Finally, the aggregated *EV* charging demand at the root *SU* can be **offered to the flexibility market**, while ensuring service quality during disaggregation to the single *EVs*. This however is only possible with the queuing network approach, because the probabilistic protocols do not offer a bi-directional communication channel to the *SUs*. Hence, the flexibility potential is not known at *SUs*. With the queuing networks, it is possible to offer positive and negative flexibility since the maximum charging requirement can be extracted from the charging requests, and in addition to the assigned charging currents, it is possible to estimate the positive and negative charging flexibility at the current moment. To trade flexibility on the market, a prediction of the *EV* charging demand is required for a given time horizon.

3.6.2 Technical Challenges

Some technical challenges exist that need to be addressed before applying the proposed methodology to the field. First, targeting Assumption A7, an easy way for the *EV* driver to provide relevant data is required. This includes static information on the wallbox and the *EV* specifications, *e. g.*, maximum charging rate, battery model, battery capacity, as well as charging service-specific parameters, *e. g.*, current *SoC* at arrival time and during the charging processes, and the planned departure time t_{dep}. Most of these parameters can be queried automatically from the *EV* via ISO 15118 or the cloud service of car manufacturers. Otherwise, the user can enter this information into a mobile application that serves as an interface to the *charging-as-a-service* platform. Missing data could be derived from historical charging behavior. However, the performance of the proposed methods on estimated charging parameters is not validated in this thesis and requires further investigation.

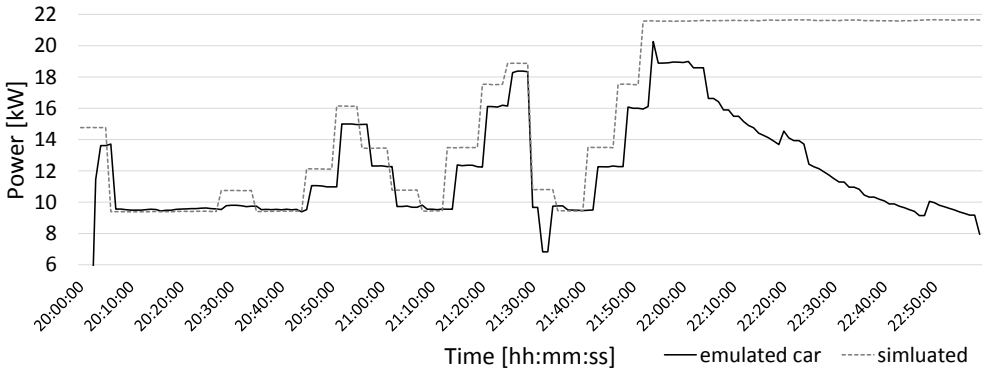
A second technical challenge is the behavior of real *EVs*, which may differ from Assumptions A2 - A6. The assumptions on *EV* charging controllability fit very well with existing standards, however, the battery management system of the *EV* could cause different

charging behavior, *e. g.*, due to battery overheating or during the battery saturation phase. Similarly, the reaction time of *EVs* on control signals may influence the actual outcome of the algorithms. Finally, *EVs* that lack reactive power control capabilities can only participate with real power reduction in under-voltage mitigation, which drastically influences its received quality of charging service.

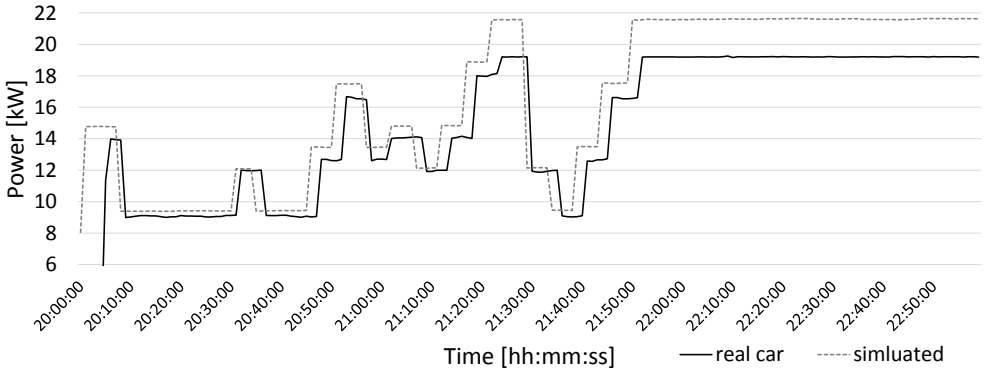
The reaction of one real and one emulated *EV* on varying charging signals is analyzed by the author of this thesis in a joint publication with Alyousef et al. in [11]. While the charging algorithm from this paper is different from the proposed solutions, the experimental results from the *power-hardware-in-the-loop* co-simulation at the laboratory of the *Austrian Institute of Technology* show pretty well how *EVs* react to charging signal changes. In this setup, a real and an emulated *EV* is connected to a charging station with a Type 2 charging outlet. The varying charging signals from the algorithm are sent to the *EV* via the IEC 61851-1 *PWM* specification of the charging outlet, as detailed in Section 2.3. Figure 3.21 compares the charging signal (dotted line) with the emulated/real behavior of an *EV*. It can be seen that in both cases the *EV* follows the charging signals with only minor deviation: *(i)* The emulated *EV* reaches its battery saturation phase in which the charging signal of 22 kW can no longer be followed and the *EV* gradually reduces its charging power despite a constant charging signal. *(ii)* Sometimes the expected charging power from the charging signal does not match exactly with the actual charging demand. However, the charging power of the *EV* does at no point in time overshoot the charging signal. Consequently, charging control via *PWM* signaling in Type 2 charging outlets works sufficiently well for being used by *QoS*-aware charging-as-a-service algorithms.

The same control algorithm from [11] is tested in the context of the European project *ELECTRIFIC* in a low-voltage power grid in Bavaria, Germany [4, 31]. Figure 3.22 shows the reactive power behavior of control signals sent to an *EV* that can only control the upper current that limits the apparent power, but not the power factor. As can be seen, when the charging power is 10 kW or more – which corresponds to 45 % of the maximum charging rate of 22 kW – the lagging power factor is greater than 0.9. Below this utilization, the power factor decreases nearly linearly. The *EV* charging stopped with a minimum charging power of 3 kW, where the lagging power factor is 0.4, which translates to reactive power of approximately 6.9 kVA. This shows that with legacy *EVs* one needs to consider that their converters are designed to optimally operate in the region of the rated charging power. This contradicts Assumption A4, where a constant efficiency is assumed with reduced charging currents. Since controllable charging processes will become more important, it can be expected that newer *EVs* are built with better-performing converters, even at lower charging currents. In addition, the communication standard ISO 15118 foresees the option to control active and reactive power separately.

Another outcome of the field trials of the research project is that communication from a measurement device via a cloud-based charging control algorithm to the final reaction of the *EV* takes less than 1 min [4]. Thereby, the reaction of the *EV* after setting the control variable at the charging station is almost instantaneous. This confirms the feasibility of a one-minute control loop for a smart charging cloud architecture as proposed in Section 3.4.1 based on the Assumption A11.



(a) Simulated charging signal with the reaction of an emulated *EV*. At the end of the charging process, the emulated *EV* reaches its saturation phase and reduces its charging power gradually.



(b) Simulated charging signal with the reaction of a real *EV*. The tested real *EV* follows nearly the simulated power, except for delayed reactions.

Figure 3.21: Reaction of a real and an emulated *EV* on charging signals [11].

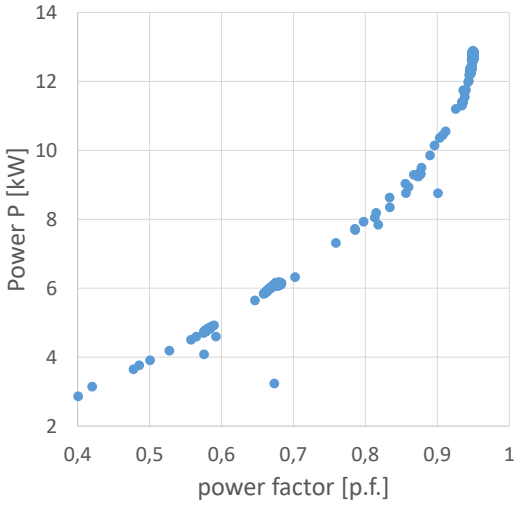


Figure 3.22: Reactive power behavior of an *EV* with reduced charging power [31].

3.6.3 Legal Framework

In Germany, there is by law (EnWG §14a) the option for controllable consumer devices in low-voltage networks. This regulation enforces the power grid operator to charge reduced grid usage fees if in turn the controllable device is agreed to be used for power grid supporting purposes and the device is metered with a separate grid metering point. These controllable consumer devices include *EV* charging, heat pumps, air conditioning, night storage heaters, and electrical energy storage systems. The amendment of EnWG §14a in mid of 2022 highlights the usage of controllable consumer devices for peak shaving purposes, which allows delayed power grid expansion, and leaves the concrete regulation and implementation to the Federal Network Agency[12]. Legislation foresees the use of the German *SMGW* [25] solution to communicate control signals to consumer devices. However, this does not impact the proposed approach. On the contrary, the *SMGW* can establish the connection between the wallboxes and the first supplying *SU* and can provide access to local power grid measurement data such as the local voltage values.

The proposed charging service allocation mechanisms are superior to the discussed peak shaving applications in the context of EnWG §14a, where *EVs* and heat pumps are mainly considered switchable loads that are either on or off. As the experiment results in Section 3.5.2 show, reduced charging power at all charging processes at the same time can help to allocate available charging currents in a more service-oriented and fair manner, while on the other hand also stabilizing the voltage levels.

3.6.4 Possible Extension with Coordinated Voltage Control

In the simulation setup, a local droop-based feedback controller stabilizes the local voltage level. Despite its simplicity and utilization in power grid connection standards, it has been shown that such controllers allocate reactive power in a sub-optimal manner [97], which leads to unnecessary grid losses. The distributed algorithm in [97] solves overvoltage events with nearly minimum grid losses by utilizing reactive power capabilities of the *DRESs* at the node that is located nearest to the node with the highest measured voltage value, also called the *regulated node* [96]. The same concept can be applied to *EV* charging, where the *EV* injects reactive power to mitigate under-voltage events. Because low-voltage power grids have a high R/X ratio, the reaction of local *DRESs*/*EVs* at the regulated node might not be sufficient to fully mitigate the voltage violation, but other nearby nodes can support. In order to minimize the grid losses, the approach of [97] selects the next best node according to the sensitivity matrix, which measures the impact of reactive power variations on the voltage at the regulated node. The sensitivity matrix of the power grid is calculated using the time-invariant line impedance (following the approach in [23]) instead of the inverse Jacobian matrix. In this way, the sensitivity matrix, and thus the corresponding delay times for the algorithm, can be precomputed based on the power grid topology only. For a given voltage variation at a certain node, a node with higher sensitivity will need to absorb/inject less reactive power than a node with lower sensitivity. Reactive power absorption/injection in low-voltage power grids correlates with power losses due to higher current flows. Conse-

[12] in German: *Bundesnetzagentur*

quently, selecting a node with high sensitivity and therefore low reactive power contribution will result in near-minimum power losses [97].

This nearly optimal reactive power allocation can easily be integrated into the hierarchical queuing network, where the local voltage measurement is sent along the hierarchical power grid infrastructure, similar to a charging request, and the maximum (minimum) voltage node is determined centrally at the root *SU*. The result is sent back to the charging processes together with the charging current assignment packets. In the case of the distributed probabilistic protocols, only a uni-directional communication channel from the *SUs* to the charging processes exists. However, the determination of the maximum (minimum) voltage node can be done using a modified max-consensus protocol, as the author of this thesis has investigated for the *DRES* use-case in [36]. Thereby, every agent of the voltage control network iteratively shares its local information state only with its direct neighbors. Upon reception of a message, every agent updates its local information state to point to the new regulated node, before sending the updated information state to its neighbors in the next communication round. Eventually, all connected agents converge to the regulated node within a limited number of iterations. In [36] different communication overlay networks are investigated with regard to convergence time and required traffic. The simplest overlay network is a local broadcast network, where all agents (reactive power agents and purely measurement agents) communicate only with the nearest direct neighbors, *e. g.*, via meshed communication network like *Zigbee*. Alternatively, a static minimum ring topology could be embedded in the underlying communication network. Finally, a *sensitivity ring* is proposed, which includes the sensitivity matrix to dynamically construct a ring network, where the next neighbor is determined based on its sensitivity on the currently known regulated node. This ensures that nodes with high impact on the regulated node are informed quite early in the ring and hence react faster on voltage events. The simulation results show that the optimal overlay network depends on the underlying communication infrastructure and is a trade-off between convergence time and total required traffic. Most important, the results show that communication delay should not be neglected when designing a distributed voltage regulation mechanism [36].

3.7 Chapter Summary

This chapter introduces a model for *EV charging-as-a-service* that represents a charging service by three main parameters: *arrival time*, *departure time*, and *required energy* to be charged during the available charging time of the charging service. This definition is followed by a set of *QoS* and *QoE* metrics that quantitatively measure the performance of *EV* charging allocation mechanisms with regard to the quality of the charging service and its perception by the *EV* driver.

Two charging allocation mechanisms to control distributed *EV* charging services in the power distribution grid are introduced. The first one is based on hierarchical queuing networks that are built on top of the radial power distribution grid, where each *EV* requests charging current packets that are then assigned by different queuing policies. Among others, the proposed *DWFQ* policy uses the remaining energy to be charged and the residual

available charging time to prioritize charging services. The second mechanism implements *MAC* protocols, known from the networking domain, on the charging service allocation problem. Thereby, each *EV* reacts to *busy* and *collision* signals from power grid assets by sampling probabilistic waiting time intervals according to a *BEB* strategy. The simple *CSMA/CA*-based method is extended with means of controlling *QoS* from wireless networks, where different *QoS* classes prioritize charging services. Additional waiting time is introduced for low-priority charging services and the exponential waiting time interval after a *collision* is limited for high-priority charging services.

The two allocation mechanisms are evaluated on a co-simulation of the IEEE 906 low-voltage test feeder using charging patterns from a mobility survey in Germany to capture realistic driving behavior. Furthermore, a sensitivity analysis investigates the impact of the simulation configuration parameters. Finally, this chapter discusses different application scopes of the proposed methods as well as potential technical and legal challenges. The main findings of the chapter are as follows.

- None of the discussed allocation mechanisms and policies outperforms all others in all seven considered *QoS* and *QoE* metrics. However, results show that laxity of the charging service – as it is used by the proposed *DWFQ* policy – plays an important role to achieve high mean metric scores for the most important metrics: Charged energy and the resulting *SoC* level of the *EV* measured by QoS_1 and QoE_1.
- When it comes to fairness, it is important to consider the metric evolution over time during the charging service. This is important especially when *EVs* need to leave earlier than their planned departure time.
- With intelligent smart charging, all charging services in the simulated low-voltage power grid can be served sufficiently with a charging power limitation of the original baseline peak load, even at high EV penetration. This is possible due to the high flexibility of *EV* home charging. In addition, allocation methods that reduce charging currents during the charging process are superior to sequential charging policies when it comes to voltage levels and grid losses.

CHAPTER 4

Power Flexibility Service from Battery Storage of Home Energy Management Systems

This chapter discusses the smart grid application of *flexibility provision-as-a-service* to distributed *Energy Management Systems* (*EMSs*), where required additional assumptions are introduced in Section 4.1. The following Section 4.2 describes the *Mixed Integer Linear Programming* (*MILP*) model of the *EMS* published in [37], which optimally schedules local appliances such as *EV* charging and a stationary *Energy Storage System* (*ESS*). This *EMS* optimization model can also integrate external flexibility requests by adding additional constraints to the planning problem. For trading flexibility on the energy market, multiple *EMSs* are combined into a *flexibility pool*, which acts as a resource aggregating *Virtual Power Plant* (*VPP*). Besides flexibility resource aggregation, the *resource aggregator* must disaggregate a flexibility request to the single *EMSs*. Section 4.3 discusses related *QoS* aspects and fairness of flexibility provision. In Section 4.4, two heuristic methods for flexibility request disaggregation are presented. The *linear heuristic* published in [37] is described in Section 4.4.1 and a *meta heuristic* based on a *Genetic Algorithm* (*GA*) published in [33] is detailed in Section 4.4.2. Both methods are evaluated in terms of *QoS* on household load and *PV* generation data from real installed *EMSs* in Section 4.5. Thereby, the meta heuristic also includes the power flow simulation of the same low-voltage power grid that is introduced during evaluation in Section 3.5.1.2. Finally, the applicability of the proposed solutions and possible extension for *aFRR* as *Ancillary Service* (*AS*) in terms of *reserve power* requests are discussed in Section 4.6, before a concluding summary is given.

---

[33] Dominik Danner, Robin Huwa, and Hermann de Meer. “Multi-objective flexibility disaggregation to distributed energy management systems”. In: *ACM SIGEnergy Energy Informatics Review* 2.2 (June 2022), pages 1–12.

[37] Dominik Danner, Jan Seidemann, Michael Lechl, and Hermann de Meer. “Flexibility disaggregation under forecast conditions”. In: *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. e-Energy ’21. Virtual Event, Italy: Association for Computing Machinery, June 2021, pages 27–38.

4.1 Assumptions

In addition to the general Assumptions A1 - A7 on *EV* charging in Section 2.5, which also apply in this chapter, further assumptions on the power distribution grid and the *EMSs* are made. These assumptions are discussed in the following. Note that Assumptions A8 - A11 are explicitly not required in this chapter.

Power Grid

A12 All required parameters of the power distribution grid are provided such that power flow equations can be solved. This includes the power grid topology, the impedance of the cables, as well as any other power system equipment in place. This information is usually available by the *DSO*. Alternatively, a black box service is required for estimating the grid losses and voltage level violations when applying certain power grid profiles at the grid connection points of the *EMSs*.

A13 Flexibility requests that are assigned to a single *EMS* are applied to their locally planned grid profile. When aggregating flexibility to a flexibility pool, grid losses that occur due to the spatial distribution of the *EMSs* in the power distribution grid are not considered during disaggregation. This fits the general idea that parties connected to the grid are measured and accounted for at their grid connection point, and grid losses are dealt with by the power system operator. This creates a natural objective for the *DSO* to minimize grid losses, which is tackled by the fitness function f_2 in Section 4.4.2.1.

Energy Management System

A14 The *ESS* can change its charging and discharging rate nearly instantaneously between the minimum and maximum limitations and these changes apply nearly instantaneously without ramp-rate limitations. Even though the offline *EMS* optimization usually runs with a 15-minute time resolution, a nearly instantaneous reaction within a few seconds is required to achieve the planned energy quantities within the time slots. Communication and control delay of local *EMS* and battery chemistry are fast enough as discussed in Section 2.4.

A15 Similar to Assumption A5, the stationary *ESS* is modeled with constant-current charging/discharging. High charging and discharging powers, which may be affected by the battery saturation phase, are minimized by objective $O_{4.1}$ in Section 4.2, and reaching the saturation phase itself is reduced by objective $O_{4.2}$, which aims to avoid high and low battery *SoC*. Furthermore, the efficiency factor – as the ratio between charged energy to discharged energy – is assumed to be constant and not influenced by external factors, *e. g.*, battery temperature, which is anyhow more stable compared to *EVs* since stationary *ESS* are typically installed inside a building.

A16 Forecast models for *PV* generation and household load exist that provide expected generation and load profiles in the desired time resolution. These forecast models further specify their error probability function $P(X = x)$, which defines the probability for having a forecast error of x. Ideally, the error distribution is symmetric around zero to avoid error aggregation over time. Possible forecast models are discussed by the co-authors in [37].

A17 The *ESS* of an *EMS* can offer automatic frequency control as *AS* to the power grid. This typically requires specific hardware controllers in place that measure the power grid frequency and activate charging or discharging of the battery accordingly, *e. g.*, using a droop-based frequency-watt controller as defined by the *SunSpec Model 134*. This assumption is relevant for the possible extension toward reserve power flexibility scheduling in Section 4.6.4.

4.2 Power Flexibility Model

In contrast to the online control methods for *EV* charging in Chapter 3, this chapter considers offline planning of different behind-the-meter appliances that are managed by an *EMS*, including *EV* charging and stationary *ESS*. This section discusses the *MILP* modeling and external flexibility request scheduling to the *EMS*, as well as flexibility aggregation into flexibility pools.

4.2.1 Home Energy Management System

The *Energy Management System* (*EMS*) of a single household consists of several non-controllable and controllable assets, which can be utilized to maximize the use of energy from local *PV* installations while respecting user-specific constraints such as the planned departure time of the *EVs* or storage limitations of a stationary *ESS*. This scheduling problem can be solved using the following *MILP*, which determines optimal operation profiles for all appliances that are managed by the *EMS*. The *MILP* problem is defined over a discrete time horizon with T time slots of length $\Delta_t, t = 1, \ldots, T$ and optimizes one *ESS* and up to N *EV* charging processes.

4.2.1.1 Decision Variables and Constraints

The indices, parameters, and decision variables of the *MILP* optimization problem are summarized in Table 4.1 and are detailed for each appliance in the following paragraphs.

Electric Vehicle Charging *EV* charging processes can be specified by their available charging time $\alpha_i(t)$, the minimum required energy $E^{req}_{EV,i}$, and hardware-specific charging limitations. In European households mainly *AC* charging with Type-2 connectors is installed, which do not support *V2G* operation as stated in Assumption A2. According to the control capabilities of IEC 61851-1, which are explained in Section 2.3, the maximum

Table 4.1: Indices, parameters, and decision variables of the *MILP* grouped by appliances.

Type		Symbol	Unit	Description
Indices		$t \in \{1,\dots,T\}$		Index for time slot t
		$i \in \{1,\dots,N\}$		Index for *EV* charging process i
Parameters	*EMS*	$\Delta_t > 0$	h	Size of the time slot
		$c^{buy}(t) > 0$	ct/kWh	Power grid energy cost
		$P_{PV}(t) \geq 0$	kW	*PV* generation forecast
		$P_{HH}(t) \geq 0$	kW	Household load forecast
		$P_G^{max}(t) \geq 0$	kW	Upper power grid limit
		$P_G^{min}(t) \leq 0$	kW	Lower power grid limit
	EV	$\alpha_i(t) \in \{0,1\}$		*EV* is available (1) or not (0)
		$\mu_{EV,i} \in (0,1]$		Charging efficiency
		$P_{EV,i}^{min} \geq 0$	kW	Minimum charging power
		$P_{EV,i}^{max} \geq P_{EV,i}^{min}$	kW	Maximum charging power
		$E_{EV,i}^{max} > 0$	kWh	Energy storage capacity
		$E_{EV,i}^{init} \in [0, E_{EV,i}^{max}]$	kWh	Initially stored energy
		$E_{EV,i}^{req} \in [0, E_{EV,i}^{max} - E_{EV,i}^{init}]$	kWh	Energy requirement
	ESS	$\mu_{ESS} \in (0,1]$		Efficiency
		$P_{ESS}^{max} \geq 0$	kW	Maximum (dis-)charging power
		$E_{ESS}^{max} \geq 0$	kWh	Energy storage capacity
		$E_{ESS}^{init} \in [0, E_{ESS}^{max}]$	kWh	Initially stored energy
Variables	*EMS*	$P_{Grid}(t) \in [P_G^{min}(t), P_G^{max}(t)]$	kW	Grid profile (dependent variable)
	EV	$M_{EV,i}(t) \in \{0,1\}$		Charge (0) or do not charge (1)
		$P_{EV,i}(t) \in [0, P_{EV,i}^{max}]$	kW	Charging power
	ESS	$M_{ESS}(t) \in \{0,1\}$		Charge (0) or discharge (1)
		$P_{ESS}^{+}(t) \in [0, P_{ESS}^{max}]$	kW	Charging power
		$P_{ESS}^{-}(t) \in [0, P_{ESS}^{max}]$	kW	Discharging power
		$E_{ESS}(t) \in [0, E_{ESS}^{max}]$	kWh	Stored energy (dependent variable)

current of these charging stations is only controllable in the range between 6 A to 32 A per phase. Hence, the problem formulation must ensure that if an *EV* i is charging, it charges with a minimum charging power of $P_{EV,i}^{min}$. Therefore, two decision variables are defined for each time slot t, and each *EV* charging process i.

$$M_{EV,i}(t) \in \{0,1\} \tag{4.1}$$
$$P_{EV,i}(t) \in \left[0, P_{EV,i}^{max}\right] \tag{4.2}$$

The integer variable in Equation (4.1) encodes whether to charge ($M_{EV,i}(t) = 0$) or not charge ($M_{EV,i}(t) = 1$), and the variable in Equation (4.2) defines the used charging power. To ensure a minimum charging power, the constraint in Equation (4.3) must hold, which further limits $P_{EV,i}(t)$ to $\{0\} \cup [P_{EV,i}^{min}, P_{EV,i}^{max}]$. Note that the relatively small charging power steps from Assumption A3 are neglected during the offline planning since the planning is based on (inaccurate) *PV* and *Household* (*HH*) forecasts and, hence, when applying the schedule the nearest power step can be chosen.

$$P_{EV,i}^{min} \leq P_{EV,i}^{max} \cdot M_{EV,i}(t) + P_{EV,i}(t) \leq P_{EV,i}^{max} \tag{4.3}$$

If the *EV* decides to not charge ($M_{EV,i}(t) = 1$), the only option for the charging power is 0 kW, whereas charging is only possible with $P_{EV,i}^{min} \leq P_{EV,i}(t) \leq P_{EV,i}^{max}$, which guarantees that each *EV* receives its minimum charging power. Additionally, the constraint in Equation (4.4) ensures that an *EV* can only be charged when the *EV* is available ($\alpha_i(t) = 1$). If it is available, $M_{EV,i}(t)$ can be chosen arbitrarily, otherwise $M_{EV,i}(t)$ must be equal to one and the *EV* cannot charge.

$$M_{EV,i}(t) + \alpha_i(t) \geq 1 \tag{4.4}$$

The following constraint in Equation (4.5) ensures that the minimum required energy is charged by each *EV* charging process. Thereby, the charging process is modeled as *constant current* charging with a constant efficiency factor $\mu_{EV,i}$ as stated in Assumptions A3 and A5.

$$E_{EV,i}^{req} \leq \sum_{t=1}^{T} P_{EV,i}(t) \cdot \Delta_t \cdot \mu_{EV,i} \leq E_{EV,i}^{max} - E_{EV,i}^{init} \tag{4.5}$$

Note that the interdependence of the battery *SoC* between two charging processes of the same *EV* is not considered, because this would require a forecast of the actual energy consumption of the trip, including potential public charging processes, which is out of scope of this thesis.

Energy Storage System The second and most flexible appliance is the *ESS*. In contrast to *EVs*, the *ESS* is always available and supports bi-directional charging. To model charging and discharging, three decision variables are defined for each time slot t.

$$M_{ESS}(t) \in \{0,1\} \tag{4.6}$$
$$P_{ESS}^{+}(t) \in [0, P_{ESS}^{max}] \tag{4.7}$$
$$P_{ESS}^{-}(t) \in [0, P_{ESS}^{max}] \tag{4.8}$$

Equation (4.6) indicates whether to charge ($M_{ESS}(t) = 0$) or discharge ($M_{ESS}(t) = 1$) the *ESS*, and the variables in Equation (4.7) and (4.8) define the respective charging and discharging power, which can be changed nearly instantaneously according to Assumption A14. To avoid simultaneous charging and discharging, two additional constraints in Equations (4.9) and (4.10) are required.

$$0 \leq M_{ESS}(t) \cdot P_{ESS}^{max} + P_{ESS}^{+}(t) \leq P_{ESS}^{max} \tag{4.9}$$

$$0 \leq (1 - M_{ESS}(t)) \cdot P_{ESS}^{max} + P_{ESS}^{-}(t) \leq P_{ESS}^{max} \tag{4.10}$$

If the solver decides to discharge the *ESS* ($M_{ESS}(t) = 1$), the constraint in Equation (4.9) fixes the charging power $P_{ESS}^{+}(t)$ to be equal to zero and Equation (4.10) allows any discharging powers in the interval $[0, P_{ESS}^{max}]$. In the case of *ESS* charging, Equation (4.10) fixes the discharging power to zero, and Equation (4.9) allows for charging with any power values. Note that two solutions exist for neither charging nor discharging the *ESS* ($M_{ESS}(t) = 0; P_{ESS}^{+}(t) = 0$ and $M_{ESS}(t) = 1; P_{ESS}^{-}(t) = 0$), nevertheless the overall profile of the *ESS*, which is calculated by $P_{ESS}(t) = P_{ESS}^{+}(t) - P_{ESS}^{-}(t)$, stays the same. The time-varying energy stored in the *ESS* is modeled by a dependent decision variable $E_{ESS}(t)$ in Equation (4.11), where $E_{ESS}(0) = E_{ESS}^{init}$. As stated in Assumption A15, the charging and discharging of the *ESS* uses a *constant current* model with constant efficiency factors μ_{ESS} and $\frac{1}{\mu_{ESS}}$ which results in a round-trip storage efficiency of $(\mu_{ESS})^2$.

$$\begin{aligned} E_{ESS}(t) = {} & E_{ESS}(t-1) \\ & + \left(P_{ESS}^{+}(t) \cdot \Delta_t \cdot \mu_{ESS}\right) - \left(P_{ESS}^{-}(t) \cdot \Delta_t \cdot \frac{1}{\mu_{ESS}}\right) \end{aligned} \tag{4.11}$$

Energy Management System Finally, the *EMS* must operate its appliances within operational constraints. First, (time-dependent) power grid limitation must be considered, *e. g.*, limitations of the grid connection fuses or maximum *PV* feed-in during peak hours. Therefore, the power grid profile, which measures the power flow at the grid connection point of the *EMS*, is calculated using a dependent decision variable $P_G(t)$ and the linear equality constraint in Equation (4.12). Time-dependent power grid limitations are applied using the constraint in Equation (4.13).

$$P_G(t) = P_{HH}(t) + \left(\sum_{i=1}^{N} P_{EV,i}(t)\right) + P_{ESS}(t) - P_{PV}(t) \tag{4.12}$$

$$P_G^{min}(t) < P_G(t) < P_G^{max}(t) \tag{4.13}$$

Note that the *MILP* formulation assumes a perfect forecast of $P_{HH}(t)$ and $P_{PV}(t)$, even though Assumption A16 states that the forecast models may experience forecast errors. One way to deal with forecast uncertainty in offline optimization is using *stochastic optimization* [157]. However, the integration of stochastic optimization into the *MILP* would drastically increase the complexity of the model. Therefore, during optimal *EMS* behind-the-meter schedule optimization a perfect forecast is assumed and stochastic behavior is only considered during the power flexibility provision in Section 4.2.2.

Second, to avoid critical reverse power flows, it is preferable to not charge an *ESS* from the power grid or to discharge the *ESS* to the power grid. Both restrictions can be modeled with the following two constraints in Equations (4.14) and (4.15), where $P_{diff}(t) = P_{PV}(t) - P_{HH}(t) - \sum_{i=1}^{N} P_{EV,i}(t)$ describes the *PV* surplus or deficit load of the *EMS*.

$$P^{+}_{ESS}(t) \leq \quad P_{diff}(t) + M_{ESS}(t) \qquad \cdot (P^{max}_{G}(t) + P^{max}_{ESS}) \tag{4.14}$$

$$P^{-}_{ESS}(t) \leq -P_{diff}(t) + (1 - M_{ESS}(t)) \cdot (P^{min}_{G}(t) + P^{max}_{ESS}) \tag{4.15}$$

In the case of power surplus ($P_{diff}(t) > 0$), only charging the *ESS* ($M_{ESS}(t) = 0$) is permitted by Equation (4.15), which would become infeasible otherwise. Furthermore, the charging power $P^{+}_{ESS}(t)$ is limited by the surplus $P_{diff}(t)$ in Equation (4.14). During a power deficit ($P_{diff}(t) < 0$), only discharging the *ESS* ($M_{ESS}(t) = 1$) is permitted by Equation (4.14) with the discharging power $P^{-}_{ESS}(t)$ limited by $P_{diff}(t)$ in Equation (4.15).

4.2.1.2 Multi-Objective Function

To optimally utilize the flexibility of all local appliances, the *EMS* solves the *MILP* problem with the following optimization objectives with hierarchical priorities.

Cost Optimization The main goal of the *EMS* is to minimize the total operational cost by minimizing the cost of consuming energy from the power grid, which in turn maximizes the degree of energy autarky of the *EMS*. Furthermore, storing local surplus *PV* generation into the *EVs* or the *ESS* optimizes self-consumption. Both goals are expressed by objective O_1 in Equation (4.16).

$$\begin{aligned} O_1 = & \sum_{t=1}^{T} P^{buy}_{G}(t) \cdot \Delta_t \cdot c^{buy}(t) \\ & - c^{mean} \cdot \left(E_{ESS}(T) - E^{init}_{ESS}\right) \cdot (\mu_{ESS})^2 \\ & - c^{mean} \cdot \sum_{i=1}^{N} \sum_{t=1}^{T} P_{EV,i}(t) \cdot \Delta_t \cdot \mu_{EV,i} \end{aligned} \tag{4.16}$$

Since objective O_1 is minimized and $c^{buy}(t) > 0$, the helper variable $P^{buy}_{G}(t) = \max(0, P_G(t))$ can be linearized by the two constraints in Equations (4.17) and (4.18).

$$P^{buy}_{G}(t) \geq 0 \tag{4.17}$$

$$P^{buy}_{G}(t) \geq P_G(t) \tag{4.18}$$

In addition, the cost objective O_1 prices the stored energy in the *ESS* and the *EVs* with the average buying cost $c^{mean} = \frac{1}{T} \sum_{t=1}^{T} c^{buy}(t)$. This favors storing PV generation for later use, while the load is served from the grid in case of below-average energy costs. Note that feed-in tariffs are not supported in the objective function and the final objective value does not express the actual payment stream.

Electric Vehicle Profile Shaping Besides cost optimization, also the shape of *EV* charging profiles is important because *EVs* should be charged in blocks to avoid charging interrupts during the charging process. This is essential if multiple *EVs* are managed by the same *EMS* to avoid continuous switching between the charging processes. Block charging can be achieved by minimizing the absolute difference between consecutive charging powers with $O_2 = \sum_{i=1}^{N} \sum_{t=1}^{T-1} |P_{EV,i}(t) - P_{EV,i}(t+1)|$. As a positive side effect, the *EV* charging profiles will be constant as well as possible, which has a positive impact on the battery *SoH*. Minimization of an absolute value can be done with helper variables $B_i^+(t) \geq 0$ and $B_t^-(t) \geq 0$ and the constraint in Equation (4.19).

$$P_{EV,i}(t) - P_{EV,i}(t+1) = B_i^+(t) - B_t^-(t) \tag{4.19}$$

The objective function O_2 is redefined in Equation (4.20) using these helper variables.

$$O_2 = \sum_{i=1}^{N} \sum_{t=1}^{T-1} \left(B_i^+(t) + B_t^-(t) \right) \tag{4.20}$$

Peak Shaving In favor of the *DSO*, the *EMS* should yield smooth power grid profiles to avoid peaks at aggregation points in the power grid. Additionally, smooth power grid profiles allow the *EMS* to compensate for *PV* generation and household load forecast errors, because of larger safety margins towards the grid limits. Therefore, the difference between the maximum grid consumption and the minimum grid feed-in is minimized by objective O_3 in Equation (4.21). The max(.) and min(.) functions can be reformulated to linear equations similarly as explained for objective O_1.

$$O_3 = \max_{t=1,\ldots,T} (P_G(t)) - \min_{t=1,\ldots,T} (P_G(t)) \tag{4.21}$$

EMS Flexibility The last objective of the *EMS* is to utilize the *ESS* in a way such that it allows providing maximum flexibility during online operation to compensate for forecast errors. Therefore, the difference between the maximum charging and discharging power of the *ESS* is minimized in Equation (4.22), which provides a safety margin of the *ESS* power flexibility. Second, the mean *SoC* of the *ESS* is kept near 50 % to ensure energy flexibility, which is encoded in Equation (4.23). In this way, power and energy flexibility can likely be provided in both directions.

$$O_{4.1} = \max_{t=1,\ldots,T} (P_{ESS}^+(t)) + \max_{t=1,\ldots,T} (P_{ESS}^-(t)) \tag{4.22}$$

$$O_{4.2} = \left| \frac{1}{2} E_{ESS}^{max} - \frac{1}{T} \sum_{t=1}^{T} E_{ESS}(t) \right| \tag{4.23}$$

Both flexibility objectives $O_{4.1}$ and $O_{4.2}$ can be reformulated to avoid absolute values $|.|$ and max(.) functions by applying the aforementioned techniques. Because both objectives focus on the same decision variables, they are combined using linearization with normalized

weights $\beta_1 = \frac{1}{2P_{ESS}^{max}}$ and $\beta_2 = \frac{1}{\frac{1}{2}E_{ESS}^{max}}$ to form objective O_4 in Equation (4.24).

$$O_4 = \beta_1 O_{4.1} + \beta_2 O_{4.2} \tag{4.24}$$

The overall *MILP* problem of the *EMS* is a multi-objective optimization problem that is solved hierarchically with decreasing prioritization in the dimensions. The *MILP* formulation is given by

$$\begin{aligned} &\min_{M_{EV,i}(t),P_{EV,i}(t),M_{ESS}(t),P_{ESS}^{+}(t),P_{ESS}^{-}(t)} && (O_1, O_2, O_3, O_4) \\ &\text{subject to} && (4.3) - (4.5), (4.9) - (4.15), (4.17) - (4.19). \end{aligned}$$

First, objective O_1 is minimized and an additional constraint is added to the problem formulation using the resulting minimum value $\widetilde{O}_1$ as in Equation (4.25). Then, the second objective is to minimize the remaining solution space, and so on.

$$O_k = \widetilde{O}_k \tag{4.25}$$

Thereby, the required additional amount of energy and the best time for demanding from the power grid is determined first. Afterward, the *EV* charging processes are scheduled in a blocked manner and the power grid profile is shaped. Finally, the usage of the *ESS* is optimized to provide future flexibility in case of forecast errors. Note that the cost of the optimal solution can be determined by only solving the first objective O_1 of the *MILP* problem due to the hierarchical optimization approach. However, to obtain the final power grid profile, all objectives must be solved hierarchically, since the last objective O_4 optimizes the usage of the *ESS*, which can potentially impact the power grid profile.

4.2.2 Power Flexibility Scheduling

The *MILP* formulation of the energy optimization problem of a household can be used to schedule external flexibility requests to the *EMS*, which are finally delivered by the flexible appliances. This flexibility provided by the *EMS* does not need to consider ramp-rate limitations, because *EVs* can change their charging power nearly instantaneously according to Assumption A3, and flexibility provided by the *ESS* is instantaneously available according to Assumption A14.

Power Flexibility Request The power flexibility provided by an *EMS* at time t is defined as the deviation from the optimal power grid profile at the grid connection point from a previous optimization run and is given by the variable $P_{flex}(t)$. A power flexibility request $P_{flex}(t)$ to the *EMS* can be scheduled by fixing the grid profile with an additional constraint as in Equation (4.26), where $\widetilde{P}_G(t)$ is the optimal grid profile of a previous optimization run.

$$P_G(t) = \widetilde{P}_G(t) - P_{flex}(t) \tag{4.26}$$

The upper and lower flexibility bounds can be determined by maximizing and minimizing the flexibility variable accordingly as in Equations (4.27) and (4.28).

$$P_{flex}^{max}(t) = \max_{M_{EV,i}(t),P_{EV,i}(t),M_{ESS}(t),P_{ESS}^{+}(t),P_{ESS}^{-}(t)} P_{flex}(t) \tag{4.27}$$

$$P_{flex}^{min}(t) = \min_{M_{EV,i}(t),P_{EV,i}(t),M_{ESS}(t),P_{ESS}^{+}(t),P_{ESS}^{-}(t)} P_{flex}(t) \tag{4.28}$$

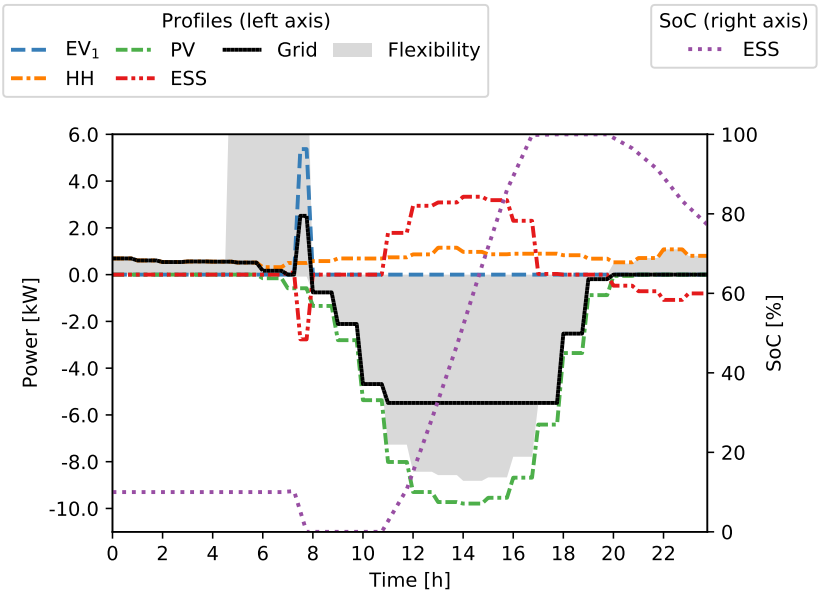
Note that in some cases a value of $P_{flex}(t)$ cannot be scheduled within the boundaries $[P_{flex}^{min}(t), P_{flex}^{max}(t)]$, because the only available flexible appliance is an *EV*, which requires a minimum power to charge. However, the *ESS* can compensate for these small power gaps in most cases.

Figure 4.1(a) shows the optimal operational profiles of an exemplary *EMS* with one *PV* system, an *ESS* ($E_{ESS}^{max} = 12\,\text{kWh}$, $P_{ESS}^{max} = 9\,\text{kW}$), and one *EV* charging process ($E_{EV,1}^{req} = 2.55\,\text{kWh}$, $P_{EV,1}^{min} = 4.3\,\text{kW}$, $P_{EV,1}^{max} = 11\,\text{kW}$, available between 4:45 and 7:45). In this setup, Δ_t is equal to 15 min, $c^{buy}(t)$ is constant for all time slots, and the *ESS* is not allowed to charge from the grid or discharge to the grid. Possible power flexibility of the power profile at the grid connection point (black solid line) is shown by the gray-shaded area, where negative flexibility (area above the grid profile) is mainly limited by *EV* availability and positive flexibility is limited by *PV* generation that is stored into the *ESS*. Figure 4.1(b) shows the same *EMS*, where power flexibility of $-1.5\,\text{kW}$ is scheduled between 12:00 and 13:00. The *MILP* decides on when to optimally compensate required energy for the flexibility request, which results in a grid profile with almost constant additional feed-in during *PV* generation. Because the *MILP* problem is used for offline day-ahead planning, compensation of the energy flexibility can also take place in advance, before the time of the flexibility request, like it is done between 11:00 and 12:00. If no compensation effect is desired at any time t', the flexibility request profile $P_{flex}(t')$ can be fixed to zero and the *MILP* solution will stick to the initial profile from the previous optimization run.

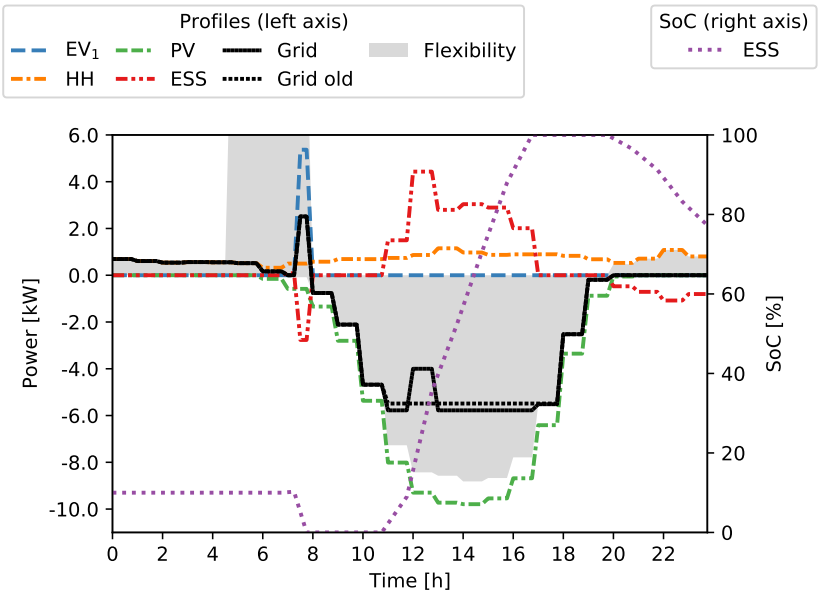
Power Flexibility Cost The cost $\gamma(x)$ of a flexibility request x is the difference between the cost of the optimal *EMS* solution before (O_1^b) and after (O_1^a) scheduling of the power flexibility request. Only the first objective (cost objective) is used because due to the hierarchical optimization, the latter objective optimizations cannot impact the minimum cost value. Furthermore, because O_1^b is minimal, scheduling a flexibility request cannot yield a smaller objective value ($O_1^b \leq O_1^a$) and, therefore, the cost of flexibility in Equation (4.29) is always greater or equal to zero.

$$\gamma(x) = O_1^a - O_1^b \tag{4.29}$$

The exemplary flexibility request in Figure 4.1 results in zero cost, because the energy cost is constant during the day and the power flexibility is provided by shifting the *ESS* charging operation. However, scheduling a power flexibility request during the *EV* availability from 4:45 to 7:45 will yield a flexibility cost in case the *EV* must be charged at non-optimal times before *PV* generation to fulfill the flexibility request.



(a) Optimal *EMS* profile before scheduling flexibility.



(b) Optimal *EMS* profile after scheduling power flexibility of −1.5 kW between 12:00 and 13:00.

Figure 4.1: The exemplary *EMS* with $\Delta_t = 15\,\text{min}$ consists of one *PV* system with an hourly forecast, one *EV* charging process between 4:45 and 7:45, and an *ESS* with $E_{ESS}^{max} = 12\,\text{kWh}$, $P_{ESS}^{max} = 9\,\text{kW}$.

Power Flexibility Uncertainty The forecast error of *PV* generation and household load can impact the probability to deliver a flexibility request. The forecast errors in Equations (4.30) and (4.31) are calculated by subtracting the actual values from the predicted values. A positive error refers to a lower actual value than expected.

$$\xi_{PV}(t) = P_{PV}(t) - \widehat{P}_{PV}(t) \tag{4.30}$$

$$\xi_{HH}(t) = P_{HH}(t) - \widehat{P}_{HH}(t) \tag{4.31}$$

PV generation and household load inversely impact the total grid demand profile $P_G(t)$. Therefore, the combined forecast error is calculated according to Equation (4.32), where a positive error refers to a lower actual grid demand than expected.

$$\xi(t) = \xi_{HH}(t) - \xi_{PV}(t) \tag{4.32}$$

For example, let there be an optimistic *PV* generation forecast of 8 kW at an arbitrary time slot, although the actual generation is only 5 kW. For the same time slot, let there be a pessimistic load forecast of 5 kW, although the actual demand is 2 kW. The positive *PV* forecast error $\xi_{PV}(t) = +3\,\text{kW}$ and the positive load forecast error $\xi_{HH}(t) = +3\,\text{kW}$ however result in a combined forecast error of 0 kW, which will not have any effect on the uncertainty of flexibility delivery.

In contrast to public charging stations, where *EVs* show up randomly, the *EMS* optimization focuses on private charging, for which arrival and departure times are known according to Assumption A7. Short deviations of *EV* availability and energy requirement can be covered by the *ESS* and, hence, these are not considered for flexibility delivery uncertainty.

If not provided along with the forecast models as stated in Assumption A16, the *error distribution function* of the forecast models can be obtained by applying *kernel density estimation* on the empirical distribution function of the forecast error of a large data set of *PV* generation and household load forecasts. The required empirical data can be recorded by the *EMS*. The error distribution function defines the probability of a forecast model to forecast with a certain error. For scheduling flexibility requests to an *EMS*, it is most interesting to determine the probability of a certain maximum error value x, for which the flexibility request cannot be delivered anymore. The probability $P_t(X \leq x)$ is expressed by the *Cumulative Distribution Function* (*CDF*), which can be created for each time slot t of a day, as depicted for hourly data in Figure 4.2. As can be seen, the *PV* forecast error mainly impacts the error distribution during daylight, while the error distribution during the night has a narrow dispersion.

The delivery probability $\rho(x)$ of a flexibility request x depends on the certainty of the forecast models. For example, if a positive flexibility request is scheduled with the maximum possible power flexibility of an *EMS*, any negative forecast error directly impacts the feasibility of the flexibility request. With a safety margin between the scheduled power flexibility and the maximum possible power flexibility, additional flexibility can be used to compensate for forecast errors and to deliver the planned grid profile. The delivery probability $\rho(x) \in [0,1]$ of a power flexibility $x = P_{flex}(t)$ is calculated as in Equation (4.33), which considers the positive and negative flexibility limits.

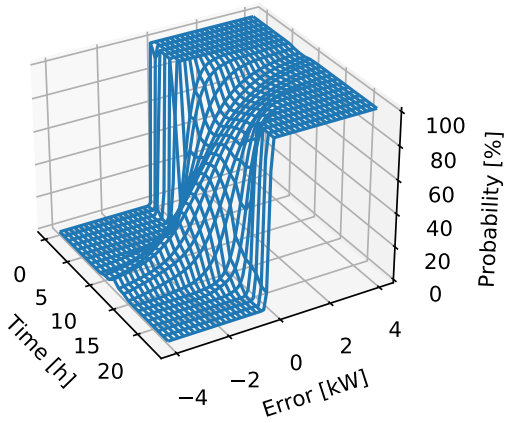


Figure 4.2: Hourly *CDF* curve of combined *PV* and household forecast error [37].

$$\rho(x) = \begin{cases} 1 - P_t(X \leq x - P_{flex}^{max}(t)) & \text{if } x > 0 \\ P_t(X \leq x - P_{flex}^{min}(t)) & \text{else} \end{cases} \tag{4.33}$$

Note that flexibility is planned against the forecast grid profile of the *EMS* and is finally delivered by the controllable assets, which adopt their operational profile. Uncertainty however is a matter of uncontrollable *PV* generation and household load profiles. In the case of forecast errors, the flexibility is still delivered by the controllable assets, while the uncertain *PV* and household load profiles determine the feasibility range at the *EMS*.

4.2.3 Flexibility Aggregation to Virtual Power Plants

A single *EMS* does not provide sufficient power flexibility to be traded on the energy market. Therefore, multiple *EMSs* are aggregated in *flexibility pools*, sometimes integrated into a *VPP*. To avoid multi-indices, but still identify *EMS e* out of *K EMSs* of the *flexibility pool*, a *context* notation $[\![.]\!]_e$ is used, where variables inside the brackets refer to *EMS e*.

Because the flexibility aggregator does not have insight into the *EMS* optimization, it only aggregates the power flexibility boundaries provided by the *EMSs*, between which it assumes that flexibility requests are feasible. The upper and lower boundaries of the flexibility pool are determined as in Equations (4.34) and (4.35).

$$P_{flex}^{max,Pool}(t) = \sum_{e=1}^{K} [\![P_{flex}^{max}(t)]\!]_e \tag{4.34}$$

$$P_{flex}^{min,Pool}(t) = \sum_{e=1}^{K} [\![P_{flex}^{min}(t)]\!]_e \tag{4.35}$$

Note that during flexibility provision the aggregator must ensure to not assign more power flexibility to a single *EMS* than it can deliver according to its local boundaries.

To disaggregate a flexibility request $P_{flex}^{Pool}(t)$ at time t to the K *EMSs*, an algorithm needs to decide to which portion each *EMS* $e = 1, \ldots, K$ contributes to the flexibility request. The disaggregation is defined by a vector $x = \left(\left[\!\left[P_{flex}(t)\right]\!\right]_1, \ldots, \left[\!\left[P_{flex}(t)\right]\!\right]_K\right)^\intercal$ containing flexibility requests to each *EMS* of the pool, where the condition in Equation (4.36) must hold. Note that according to Assumption A13, grid losses are not considered during disaggregation and the flexibility is accounted for at the grid connection point of the *EMSs*.

$$P_{flex}^{Pool}(t) = \sum_{e=1}^{K} x_e \tag{4.36}$$

The disaggregation of a power flexibility request $(P_{flex}^{Pool}(t))_{t\in\mathcal{T}}$ over multiple time slots $t \in \mathcal{T} \subseteq \{1, \ldots, T\}$ is represented by a matrix as in Equation (4.37). Each row vector is the power flexibility delivered by one *EMS* over all time slots $\mathcal{T}$ and each column vector is the disaggregation of the power flexibility request to the *EMSs* at a single time slot.

$$\begin{pmatrix} \left[\!\left[P_{flex}(t_1)\right]\!\right]_1 & \cdots & \left[\!\left[P_{flex}(t_{|\mathcal{T}|})\right]\!\right]_1 \\ \vdots & \ddots & \vdots \\ \left[\!\left[P_{flex}(t_1)\right]\!\right]_K & \cdots & \left[\!\left[P_{flex}(t_{|\mathcal{T}|})\right]\!\right]_K \end{pmatrix} \in \mathbb{R}^{K\times|\mathcal{T}|} \tag{4.37}$$

A flexibility disaggregation is called *feasible* if all *EMSs* can schedule their power flexibility requests. The cost of power flexibility disaggregation is the sum of individual costs for each *EMS* and is calculated in Equation (4.38). Similarly, the delivery probability is the multiplication of delivery probabilities of the single *EMSs* given in Equation (4.39).

$$\gamma^{Pool}(x) = \sum_{t\in\mathcal{T}} \sum_{e=1}^{K} \left[\!\left[\gamma(P_{flex}(t))\right]\!\right]_e \tag{4.38}$$

$$\rho^{Pool}(x) = \prod_{t\in\mathcal{T}} \prod_{e=1}^{K} \left[\!\left[\rho(P_{flex}(t))\right]\!\right]_e \in [0,1] \tag{4.39}$$

4.3 Power Flexibility Provision Assessment

Based on the *MILP* formulation of the *EMSs* and the flexibility service aggregation to *flexibility pools*, this section details *QoS* and *QoE* parameters for the different participating actors, as well as fairness considerations for *flexibility provision as-a-service*. Since this smart grid application is different from the *EV* charging service in Chapter 3, the actors and their expectations, and therefore the quality metrics, differ. First, a stationary *ESS* is available round-the-clock and its usage is not limited to a certain duration throughout the day. Second, the goal and usage of the stationary battery capacity are different compared to *EV* batteries. An *ESS* is typically used in conjunction with *PV* systems to store surplus energy for later self-consumption, however, there is no requirement for a minimum charge. Moreover, the *MILP* problem is used for offline flexibility planning, which guarantees that the *EV* is charged with its required energy within the available charging time by the constraint in Equation (4.5). Therefore, the *QoS* and *QoE* metrics from Chapter 3 are not listed

as part of the metrics for the resource provider. Furthermore, the main goal of the *EMS* is to maximize energy *autarky* and *self-consumption* by objective O_1 in Equation (4.16), which contradicts the goal of the online charging service allocation in Chapter 3.

4.3.1 Quality of Service

In the context of this work, the *QoS* of the power flexibility provision service is defined by the following criteria, which consider the view of the resource provider and the resource aggregator. The metric value numbering continues from the last chapter with document-wide unique metric identifiers.

Resource Provider The first aspect the resource provider cares about is to which extent its *EMS* contributes to a power flexibility request to the flexibility pool. Higher participation in flexibility delivery corresponds to a higher payment by the resource aggregator for providing flexibility. Consequently, the individual resource providers may perceive this as a higher service quality.[13] On the contrary, a low flexibility contribution or not being selected at all corresponds to low service quality. The participation of an *EMS* can be measured proportionally to the total flexibility request to the pool as given in Equation (4.40). Absolute flexibility power values are used because a positive flexibility request to the flexibility pool can theoretically result in a combination of positive and negative flexibility requests to the single *EMSs*, which together deliver the flexibility request of the pool.

$$QoS_5 = \frac{\sum_{t\in\mathcal{T}} [\![|P_{flex}(t)|]\!]_e \cdot \Delta_t}{\sum_{d=1}^{K} \sum_{t\in\mathcal{T}} [\![|P_{flex}(t)|]\!]_d \cdot \Delta_t} \in [0,1] \tag{4.40}$$

As the second quality criterion, the resource provider may care about the direct effect of flexibility scheduling on its locally optimal *EMS* schedule. Since the *MILP* performs a hierarchical optimization on a four-dimensional objective vector (O_1, O_2, O_3, O_4) with decreasing priority, the loss of optimality is computed by the difference of the objective value before and after scheduling the flexibility. The metric is calculated as in Equation (4.41) no matter if it is a flexibility request for a single time slot or over multiple time slots. To take care of the objective prioritization, the weights ω_k decrease by one order of magnitude for each objective k ($\omega_1 = 10^3, \omega_2 = 10^2, \omega_3 = 10^1, \omega_4 = 10^0$).

$$QoS_6 = -\sum_{k=1}^{4} \omega_k \cdot (O_k^a - O_k^b) \tag{4.41}$$

If an *EMS* does not participate in the flexibility delivery ($QoS_5 = 0$), there is logically no deviation from the optimal schedule and, hence, $QoS_6 = 0$ as well. On the contrary, any participation in the flexibility request can yield very small to large deviations from the optimal schedule depending on how the flexibility is provided. Ultimately, there even exists flexibility participation that does not impact the objective vector at all. This is the

[13] Compensation should be at least as high as the cost $\gamma(x)$ of the scheduled flexibility request. A merit-order principle could be implemented to determine the final reward for the resource provider.

case when the optimal solution after scheduling the flexibility request resides within the solution space of the original *MILP* problem (without scheduled flexibility) after adding the last hierarchical objective value of O_4 to the constraints.

Resource Aggregator Besides the resource provider, also the resource aggregator can have expectations from the flexibility disaggregation service. The first *QoS* metric of power flexibility provision is the total costs of the disaggregation vector, which is calculated in Equation (4.42). Lower cost of flexibility delivery is seen as higher service quality from the perspective of the resource aggregator and *vice versa*.

$$QoS_7 = -\gamma^{Pool}\left(\left(\left[\!\left[P_{flex}(t)\right]\!\right]_e\right)_{e\in\{1,\dots,K\},t\in\mathcal{T}}\right) \tag{4.42}$$

There exists a nearly linear relation between the optimality metric QoS_6 of a single resource provider and the cost of flexibility delivery since the first objective O_1 is used to determine the cost of the flexibility schedule. However, for QoS_6, the lower-priority objectives may additionally influence the metric value, so that the two metrics express slightly different service expectations. Moreover, QoS_7 measures the total cost of flexibility provided by the flexibility pool, where a single highly non-optimal *EMS* does not necessarily result in a high aggregated flexibility cost.

The resource aggregator further cares about the probability of flexibility delivery, because this impacts the possibility to trade the flexibility on the energy market. The respective metric QoS_8 calculates the mean delivery probability and is given in Equation (4.43).[14] High flexibility participation (high QoS_5) may cause the *EMSs* to operate near their limits, which consequently lowers the delivery probability depending on the *CDF* of the forecast model. However, since QoS_5 is measured proportional to the flexibility request to the pool, high participation can yield a high delivery probability if the total request is rather small.

$$QoS_8 = \frac{1}{K\cdot|\mathcal{T}|}\sum_{e=1}^{K}\sum_{t\in\mathcal{T}}\left[\!\left[\rho\left(P_{flex}(t)\right)\right]\!\right]_e \in [0,1] \tag{4.43}$$

4.3.2 Quality of Experience

In contrast to QoS_5, which measures proportional participation in a flexibility request, the first *QoE* metric refers to the participation capabilities of the single *EMS*. If an *EMS* utilizes a large portion of its local flexibility potential, this corresponds to a high *QoE* for the resource provider. On the contrary, small local flexibility utilization results in a low *QoE* metric value. Thus, the quality metric depends on the size of the flexibility request. The calculation of QoE_4 is given in Equation (4.44), where $h(t)$ measures the request-specific flexibility limits and is defined in Equation (4.45).

$$QoE_4 = \frac{\sum_{t\in\mathcal{T}}|P_{flex}(t)|\cdot\Delta_t}{\sum_{t\in\mathcal{T}}|h(t)|\cdot\Delta_t} \in [0,1] \tag{4.44}$$

[14] The mean value is chosen instead of the multiplication in Equation (4.39) because the resource aggregator may tolerate and compensate a few low delivery probabilities by single *EMSs*.

$$h(t) = \begin{cases} \check{P}_{flex}^{min}(t) & \text{if } P_{flex}(t) \leq 0 \\ \check{P}_{flex}^{max}(t) & \text{else} \end{cases} \tag{4.45}$$

The minimum and maximum flexibility limits of an *EMS* at time t depend on scheduled flexibility at all other time slots of the flexibility request $(P_{flex}(t))_{t\in\mathcal{T}}$. Therefore, $\check{P}_{flex}^{min}(t)$ and $\check{P}_{flex}^{max}(t)$ are re-calculated for each time slot t with flexibility scheduled for all other time slots $t' \in \mathcal{T} \setminus \{t\}$. Note that again the absolute values for power flexibility and flexibility limits are used because a flexibility request may contain both positive and negative flexibility at different times over multiple time slots.

Second, the resource provider is concerned about the achieved degree of *self-consumption*, which is defined as the percentage of the produced *PV* energy that is locally consumed or stored in the *ESS*. The maximum self-consumption of 100 % means that the whole energy produced by the *PV* system is directly consumed and no energy is sold to the grid at any point in time. The metric calculation is defined in Equation (4.46), where the feed-in power of the *EMS* is given by the negative power value $P_G^{sell}(t) = \min(0, P_G(t))$.

$$QoE_5 = 1 - \frac{\sum_{t=1}^{T} \left| P_G^{sell}(t) \right| \cdot \Delta_t}{\sum_{t=1}^{T} P_{PV}(t) \cdot \Delta_t} \tag{4.46}$$

The calculated self-consumption can theoretically be smaller than zero if the battery has a high initial *SoC* and all energy is sold to the grid.[15] This however is avoided by objective O_1 in Equation (4.16), which minimizes the energy sold to the grid.

A third *QoE* metric for *EMSs* is the degree of energy *autarky*, which is defined as the percentage of the local energy consumption that is covered by *PV* generation or discharging the *ESS*. The maximum autarky of 100 % means that local consumption is entirely supplied by *PV* generation or indirectly by stored *PV* energy from the *ESS*, thus no energy from the grid is needed. The autarky metric is calculated in Equation (4.47).

$$QoE_6 = 1 - \frac{\sum_{t=1}^{T} P_G^{buy}(t) \cdot \Delta_t}{\sum_{t=1}^{T} \left(P_{HH}(t) + \sum_{i=1}^{N} P_{EV,i}(t) \right) \cdot \Delta_t} \tag{4.47}$$

Since the *ESS* is not allowed to be charged from the power grid by the constraint in Equation (4.14), the power grid profile can never exceed the local consumption and, hence, $P_G^{buy}(t) \leq P_{HH}(t) + \sum_{i=1}^{N} P_{EV,i}(t)$ for any time slot $t \in \{1, \ldots, T\}$. Lifting this constraint may result in a negative autarky because energy charged from the power grid reduces the energy autarky of the *EMS* in the given time horizon.

Self-consumption and autarky measure the performance of the *MILP* and are orthogonal to the utilization of the local flexibility potential. However, a flexibility request may decrease the percentage of self-consumption or autarky due to sub-optimal local appliance scheduling. The degree of self-consumption and autarky are orthogonal since the first refers to the generation of the *PV* system and the second refers to the consumption of the *EMS*.

There is no *QoE* metric for the resource aggregator because trading of aggregated flexibility on the market is beyond the scope of this thesis and this chapter mainly focuses on the *EMS*

[15] This is only possible if the *EMS* constraints from Equation (4.14) and (4.15) are lifted.

modeling. Nevertheless, a *QoE* metric for the resource aggregator may cover expected and realized earnings from the flexibility market.

4.3.3 Fairness

The algorithms described in the following section disaggregate a flexibility request to the *EMSs* of one flexibility pool. Hence, fairness cannot be measured between different flexibility pools, but only among the *EMSs* of the same flexibility pool. Therefore, the resource aggregator quality metrics QoS_7 and QoS_8 are not considered for fairness analysis. Nevertheless, a fair distribution of the quality metric to the resource providers is in the sense of the resource providers and the aggregator.

Again, the fairness index from Equation (2.3) is used to measure the fairness of *QoS* and *QoE* distributions among the *EMSs* of a flexibility pool. All metrics defined in Equations (4.40) - (4.47), except QoS_6 and QoS_7 in Equations (4.41) and (4.42), are bounded within the interval $[0,1]$.[16] Since $L = 0$ and $H = 1$, the index can be simplified similarly as in Equation (3.13). The maximum value H of QoS_6 is equal to zero, but the minimum value L is unbounded. To have a comparable fairness index, L is determined by the $\min(.)$ function over all metric values from the different *EMSs*. The unbounded metric QoS_7 is not considered for fairness evaluation, as it only reflects a single value for the flexibility pool.

4.4 Methodology

This section discusses two heuristics to assign individual flexibility requests to the *EMSs* of a flexibility pool that together deliver the requested flexibility. The first method is a linear heuristic that schedules the flexibility request to a single time slot after each other. The second method is based on a *Genetic Algorithm* (*GA*), which targets multiple objectives and iteratively optimizes all time slots of the flexibility request at once, starting with an initially generated set of possible assignments.

From the perspective of the resource aggregator, an optimal flexibility provision should yield both low total cost and high delivery probability. Unfortunately, this cannot be achieved using linear optimization, because the objective function would be non-linear (flexibility delivery probability) or can only be computed at the *EMS* (flexibility cost).

4.4.1 Iterative Linear Heuristic

The linear heuristic algorithm iteratively assigns portions of the flexibility request (with a maximum size of S^{max}) to the most appropriate *EMS*, which is determined by a *disaggregation priority function* $p([\![P_{flex}(t)]\!]_e) \rightarrow \mathbb{R}$, until the flexibility request for a specific time slot t is fully scheduled. This priority function must decrease monotonically with increasing assigned power flexibility to ensure correctness. Flexibility requests to multiple time slots are solved iteratively, starting with the earliest time slot in $\mathcal{T}$.

[16] No negative self-consumption and autarky values are measured in all conducted experiments.

Algorithm 4.1: Heuristic algorithm to disaggregate positive flexibility.

Input: Flexibility request to the flexibility pool $P_{flex}^{Pool}(t) > 0$
Data: Disaggregation priority function $p(\llbracket P_{flex}(t) \rrbracket_e) \to \mathbb{R}$,
maximum flexibility portion size $S^{max} > 0$
Output: Disaggregation vector $\left(\llbracket P_{flex}(t) \rrbracket_1, \ldots, \llbracket P_{flex}(t) \rrbracket_K\right)$

1 **if** $P_{flex}^{Pool}(t) < P_{flex}^{Pool,max}(t)$ **then**
2 $\quad \forall e = 1, \ldots, K : \llbracket P_{flex}(t) \rrbracket_e \leftarrow 0$; // Initialize *EMSs*
3 $\quad F \leftarrow \left\{ e = 1, \ldots, K : \left\llbracket P_{flex}^{max}(t) > 0 \right\rrbracket_e \right\}$; // feasible *EMSs*
4 $\quad P_{flex}^{rem}(t) \leftarrow P_{flex}^{Pool}(t)$; // Remaining flexibility
5 $\quad$ **while** $P_{flex}^{rem}(t) > 0$ **do**
6 $\quad\quad p_{max} \leftarrow -\infty$;
7 $\quad\quad$ **foreach** $e \in F$ **do**
8 $\quad\quad\quad S_e \leftarrow \min\left(S^{max}, P_{flex}^{rem}(t), \left\llbracket P_{flex}^{max}(t) - P_{flex}(t) \right\rrbracket_e\right)$;
9 $\quad\quad\quad$ **if** $p(\llbracket P_{flex}(t) \rrbracket_e + S_e) > p_{max}$ **then**
10 $\quad\quad\quad\quad p_{max} \leftarrow p(\llbracket P_{flex}(t) \rrbracket_e + S_e)$;
11 $\quad\quad\quad\quad d \leftarrow e$; // Select *EMS* d
12 $\quad\quad\quad$ **end**
13 $\quad\quad$ **end**
14 $\quad\quad \llbracket P_{flex}(t) \rrbracket_d \leftarrow \llbracket P_{flex}(t) \rrbracket_d + S_d$;
15 $\quad\quad P_{flex}^{rem}(t) \leftarrow P_{flex}^{rem}(t) - S_d$;
16 $\quad\quad$ **if** $\left\llbracket P_{flex}^{max}(t) = P_{flex}(t) \right\rrbracket_d$ **then**
17 $\quad\quad\quad F \leftarrow F \setminus d$;
18 $\quad\quad$ **end**
19 $\quad$ **end**
20 $\quad$ **return** $\left(\llbracket P_{flex}(t) \rrbracket_1, \ldots, \llbracket P_{flex}(t) \rrbracket_K\right)$
21 **end**

The disaggregation algorithm for positive flexibility is given in Algorithm 4.1, negative flexibility is scheduled accordingly. The algorithm starts with initializing required variables (lines 2 - 4), before iteratively assigning flexibility portions (lines 14 - 18) to the *EMS* with the highest priority value, determined in lines 6 - 13. Because the set of feasible *EMSs* F is unordered, an *EMS* is randomly selected if several *EMSs* yield the same priority value. Note that this algorithm assigns only positive requests to the *EMSs* and, hence, no balancing between *EMSs* is performed to optimize the priority function value. Allowing a disaggregation vector with non-uniform signs may also imply a higher total cost for the flexibility because additional negative flexibility must be compensated by the remaining *EMSs*. Furthermore, power distribution grid limitations are not considered explicitly by the disaggregation algorithm, however, the grid losses and voltage levels obtained by a disaggregation vector can be checked with power flow analysis. Alternatively, the distributed *EMSs* should only be allowed to offer limited flexibility, e.g. by restricting the parameters $P_G^{max}(t)$ and $P_G^{min}(t)$ or by implementing concepts like flexibility quotas proposed in [175].

In the following, different policies are discussed together with their priority functions for positive flexibility. Priority functions for negative flexibility are constructed similarly.

Equal (EQUAL) With equal disaggregation, each *EMS* receives an absolute equal share of the flexibility request, unless it is infeasible. In that case, limited *EMSs* will use their maximum possible flexibility, and the remaining *EMSs* have an equal share of flexibility provisioning. The priority function is given in Equation (4.48). This policy implements the *water-filling* algorithm.

$$p^{EQUAL}(\llbracket P_{flex}(t) \rrbracket_e) = -\left|\llbracket P_{flex}(t) \rrbracket_e\right| \tag{4.48}$$

Proportional (PROP) With proportional disaggregation, each *EMS* receives a flexibility request proportional to its maximum possible flexibility. The priority function in Equation (4.49) ensures a proportional fair disaggregation with respect to local capabilities defined by QoE_4 in Equation (4.44).

$$p^{PROP}(\llbracket P_{flex}(t) \rrbracket_e) = -\left\llbracket \frac{P_{flex}(t)}{P_{flex}^{max}(t)} \right\rrbracket_e \tag{4.49}$$

Cost-optimal (COST) With cost-optimal disaggregation, the total flexibility cost, defined in Equation (4.38), is minimized by the priority function in Equation (4.50).

$$p^{COST}(\llbracket P_{flex}(t) \rrbracket_e) = -\llbracket \gamma(P_{flex}(t)) \rrbracket_e \tag{4.50}$$

The continuous relaxation of the *MILP* constraint set is convex as well as the affine-linear objective function O_1. As a result, the solution space is convex and the objective function O_1 increases monotonically with increasing scheduled flexibility. The cost-optimal policy implements a *merit-order* curve for scheduling flexibility to a single *EMS*. The iterative algorithm in Algorithm 4.1 heuristically takes the *EMS* that provides the next flexibility portion with the lowest overall flexibility cost, which is equal to stepping through the combined merit-order curve of the flexibility pool with step size S^{max}.

Probability-optimal (POPT) With probability-optimal disaggregation, the overall probability from Equation (4.39) is maximized. Assuming that the sum of all flexibility probabilities of the *EMSs* is constant, all probabilities must be identical to maximize the product of probabilities according to Lemma 1. Hence, the first objective is to equalize the probabilities between the *EMSs*. Because the product of all probabilities is continuous, it is enough to reach an arbitrary near value to equality in order to obtain a value near the maximum. Following Lemma 2, the maximum possible product of all single probabilities is greater if the sum of the probabilities is greater, and therefore the second objective is to maximize the sum of the probabilities. Both objectives can heuristically be achieved with the priority function in Equation (4.51).

$$p^{POPT}(\llbracket P_{flex}(t) \rrbracket_e) = \llbracket \rho(P_{flex}(t)) \rrbracket_e \tag{4.51}$$

Lemma 1
Let $K \in \mathbb{R}_0^+$, $a_1,\ldots,a_n \in \mathbb{R}_0^+$ *such that* $\sum_{i=1}^n a_i = K$ *and* $f\colon \{\mathbb{R}^n | \sum_{i=1}^n a_i = K\} \to \mathbb{R}\colon (a_1,\ldots,a_n) \mapsto \prod_{i=1}^n a_i$.

$$\forall i = 1,\ldots,n : a_i = \frac{K}{n} \textit{ assumes a maximum in } f(a_1,\ldots,a_n)$$

PROOF Assume there is a vector $(b_1,\ldots,b_n)$ that yields a maximum value for $\prod_{i=1}^n b_i$, but has different values at indices k and m, $k \neq m$. Let $d > 0$, $b_k = \frac{K}{n} - d$ and $b_m = \frac{K}{n} + d$, then $b_k \cdot b_m = \left(\frac{K}{n}\right)^2 - d^2 < \left(\frac{K}{n}\right)^2$, hence $\prod_{i=1}^n b_i$ not maximal. ∎

Lemma 2
Let $K_1, K_2 \in \mathbb{R}_0^+$, $K_1 < K_2$, $a_1,\ldots,a_n,b_1,\ldots,b_n \in \mathbb{R}_0^+$ *such that* $\sum_{i=1}^n a_i = K_1$ *and* $\sum_{i=1}^n b_i = K_2$.

$$\max_{a_1,\ldots,a_n} \prod_{i=1}^n a_i < \max_{b_1,\ldots,b_n} \prod_{i=1}^n b_i$$

PROOF Without loss of generality we assume that $K_2 = K_1 + 1$. Let $a_1,\ldots,a_n$ be values that yield the maximum for $\prod_{i=1}^n a_i$ where $\sum_{i=1}^n a_i = K_1$. Define $\forall i \in \{1,\ldots,(n-1)\}\colon b_i = a_i$, $b_n = a_n + 1$ and $\sum_{i=1}^n b_i = K_2$, then obviously $K_1 < K_2$ and

$$\max_{a_1,\ldots,a_n} \prod_{i=1}^n a_i = a_n \prod_{i=1}^{n-1} a_i < (a_n + 1) \prod_{i=1}^{n-1} a_i = \prod_{i=1}^n b_i \leq \max_{b_1,\ldots,b_n} \prod_{i=1}^n b_i \quad . \qquad \blacksquare$$

Since $\rho(.)$ decreases monotonically with increasing flexibility, selecting the *EMS* with the highest probability will narrow down the differences between *EMSs* as well as possible. On the other hand, if all probabilities are already close, selecting the *EMS* with the highest probability value will keep the sum of probabilities high as well.

The aforementioned flexibility disaggregation procedure works for disaggregating a flexibility request in a single time slot. Flexibility requests that cover multiple time slots $\mathcal{T}$ are disaggregated by Algorithm 4.2, where the linear heuristic starts with the earliest time slot $t_1 \in \mathcal{T}$. Afterward, each *EMS* re-computes its optimal grid profile $(P_G(t))_{t=1,\ldots,T}$, considering the scheduled flexibility at time t_1. Because the aggregation of these new *EMS* grid profiles may deviate from the original aggregation, to which the flexibility was requested, a potential deviation at the second time slot $t_2 \in \mathcal{T}$ must be compensated by adding or subtracting the difference in line 3. This procedure is executed until eventually all time slots of the flexibility request are scheduled. In case a disaggregation is not feasible at any time slot $t \in \mathcal{T}$, the flexibility request $(P_{flex}^{Pool}(t))_{t\in\mathcal{T}}$ cannot be fulfilled.

This iterative approach may not find a globally optimal flexibility disaggregation for all time slots, since a modification of already processed time slots may help to improve the metric function of other time slots. This is because the linear heuristic only ensures a nearly

Algorithm 4.2: Heuristic algorithm to disaggregate flexibility over multiple time slots.

Input: Flexibility request to the flexibility pool $\left(P_{flex}^{Pool}(t)\right)_{t\in\mathcal{T}}$
Data: Original aggregated grid profile of the pool
$\left(P_G^{Pool}(t)\right)_{t\in\mathcal{T}} = \left(\sum_{e=1}^{K} [\![P_G(t)]\!]_e\right)_{t\in\mathcal{T}}$
Output: Disaggregation matrix $\left([\![P_{flex}(t)]\!]_e\right)_{e\in\{1,\dots,K\},t\in\mathcal{T}}$

1 **foreach** $t \in \mathcal{T}$ **do**
2 $\quad \check{P}_G^{Pool}(t) \leftarrow \sum_{e=1}^{K} [\![\check{P}_G(t)]\!]_e$; // Updated grid profile
3 $\quad P_{flex}^{Pool}(t) \leftarrow P_{flex}^{Pool}(t) + \left(\check{P}_G^{Pool}(t) - P_G^{Pool}(t)\right)$; // Compensate difference
4 $\quad \left([\![P_{flex}(t)]\!]_1, \dots, [\![P_{flex}(t)]\!]_K\right) \leftarrow disaggregate(P_{flex}^{Pool}(t))$; // Algorithm 4.1
5 $\quad$ **foreach** $e = 1, \dots, K$ **do**
6 $\quad\quad schedule([\![P_{flex}(t)]\!]_e)$; // Equation (4.26)
7 $\quad\quad optimize(e)$; // Solve *MILP*
8 $\quad$ **end**
9 **end**
10 **return** $\left([\![P_{flex}(t)]\!]_e\right)_{e\in\{1,\dots,K\},t\in\mathcal{T}}$

optimal solution of flexibility disaggregation for one single time slot. When disaggregating multiple time slots at once, Algorithm 4.2 can be executed with different permutations of the time slots in line 1. Nevertheless, the iterative flexibility disaggregation in the correct time order is reasonable when the linear heuristic is used for online flexibility disaggregation, where past flexibility allocations cannot be changed anymore and future flexibility requirements are not known in advance.

4.4.2 Genetic Meta Heuristic

Compared to the linear heuristic, the meta heuristic schedules flexibility requests to multiple time slots at once. Thereby, one possible solution – in the context of *GAs* called *chromosome* – is represented by a matrix as already defined in Equation (4.37), where each column vector models a *gene* of the chromosome. The *GA* targets multiple objectives from different stakeholders at the same time, which are encoded by the fitness function that measures the optimality of a chromosome. The *GA* fitness function includes *(i)* the resource aggregator, who wants to minimize the cost for flexibility while maximizing the probability of delivery; *(ii)* the grid operator, who wants to minimize grid losses; *(iii)* the resource provider (*EMS* owner), who wants to maximize the level of autarky and self-consumption; *(iv)* the fairness service provider, who wants to ensure that among *EMSs* a fair level of autarky and self-consumption is reached. Thereby, the internals of the power grid and the *EMS* optimization are treated as *black boxes*, which allows the flexibility disaggregation service to be easily implemented by the resource aggregator. Moreover, the *GA* allows *a posteriori* selection of a schedule from the Pareto-set [46], which is beneficial if multiple different stakeholders are involved with unknown elasticity to their objectives.

Algorithm 4.3: Initial population sampling for the *GA*. $random(.)$ samples a uniform random value from an interval or a uniform element from a set.

Input: Flexibility request to the flexibility pool $\left(P_{flex}^{Pool}(t)\right)_{t\in\mathcal{T}}$

Output: Initial population set Π_0

1 $\Pi_0 \leftarrow \emptyset$;
2 **while** $|\Pi_0| < \mu$ **do**
3 $\quad C \leftarrow 0_{K,|\mathcal{T}|}$; `// Initialize chromosome matrix`
4 $\quad$ **while** $\exists t \in \mathcal{T}\colon \sum_{e=1}^{K} C_{e,t} < P_{flex}^{Pool}(t)$ **do**
5 $\quad\quad e \leftarrow random(\{1,\ldots,K\})$; `// Select random EMS`
6 $\quad\quad t \leftarrow random(\mathcal{T})$; `// Select random time`
7 $\quad\quad p \leftarrow random(\left[0, P_{flex}^{Pool}(t) - \sum_{e=1}^{K} C_{e,t}\right] \cap \left[\left[\left[P_{flex}^{min}(t)\right]\right]_e, \left[\left[P_{flex}^{max}(t)\right]\right]_e\right])$;
8 $\quad\quad p \leftarrow \min(\max(p, S^{max}), P_{flex}^{Pool}(t) - \sum_{e=1}^{K} C_{e,t})$;
9 $\quad\quad C_{e,t} \leftarrow C_{e,t} + p$
10 $\quad$ **end**
11 $\quad \Pi_0 \leftarrow \Pi_0 \cup \{C\}$
12 **end**
13 **return** Π_0

The *GA* processes chromosomes in population sets Π_p of size μ, where the index $p \in \mathbb{N}_0$ represents the generation number. The initial population Π_0 for a positive flexibility request is generated by Algorithm 4.3, which iteratively selects a random *EMS* $e \in \{1,\ldots,K\}$ (matrix row) and a random time slot $t \in \mathcal{T}$ (matrix column), and samples a random flexibility portion from the feasible flexibility of the selected *EMS* e at time slot t from the interval given in line 7. Thereby, it is ensured that not arbitrarily small portions are assigned by limiting the minimum portion to size S^{max}. The result is added to the already scheduled flexibility until the full flexibility request to the flexibility pool is served. The initial population for negative flexibility requests is constructed similarly. Note that this creates a flexibility disaggregation with identical signs at all *EMS* because inverse flexibility contribution from a single *EMS* would need compensation by the other *EMSs*. Furthermore, it is worth mentioning that the resulting initial population is not uniformly distributed in the solution space. Even though each *EMS* is selected with the same probability, an *EMS* with higher flexibility range will more likely receive an overall higher contribution share.

Starting with the initial population Π_0, subsequent populations Π_{p+1} are created by the *selection*, *modification*, and *insertion* of new chromosomes using *genetic operators* as described in Algorithm 4.4. Thereby, each population first updates the known Pareto-optimal solution set Γ_{Pareto} in line 6 comparing the chromosomes by their fitness values, which encodes the multi-objective criteria. Afterward, the chromosomes are sorted by a *diversity operator* in line 7 to avoid clustering of similar chromosomes. In line 8, the *selection operator* selects μ_{sel} candidates from the population set, and together with newly generated random chromosomes and chromosomes from the *Pareto-optimal solution set*, the next population of size $\mu = \mu_{sel} + \mu_{new} + \mu_{pareto}$ is created. The *crossover operator* and the *mutation operator* in line 10 modify the population to generate chromosome variations with

Algorithm 4.4: *GA* to determine the Pareto-optimal solution set of a flexibility request to a flexibility pool.

Input: Initial population Π_0
Output: Pareto-optimal solution set Γ_{Pareto}

$p \leftarrow 0$;
$\Gamma_{Pareto} \leftarrow \{\}$;
while *stop criteria not met* **do**
  $p \leftarrow p+1$;
  $\Pi_p = \Pi_{p-1}$;
  $\Gamma_{Pareto} \leftarrow update(\Gamma_{Pareto}, \Pi_p)$;
  $\Pi_p \leftarrow diversify(\Pi_p)$;
  $\Pi_p \leftarrow select(\Pi_p) \cup new\_chromosomes() \cup random(\Gamma_{Pareto})$;
  $\Pi_p^c, \Pi_p^m \subset \Pi_p \colon \Pi_p^c \cap \Pi_p^m = \emptyset \wedge \left|\Pi_p^c\right| = \left|\Pi_p^m\right| = \frac{|\Pi_p|}{2}$; `// Equal subsets`
  $\Pi_p \leftarrow crossover(\Pi_p^c) \cup mutate(\Pi_p^m)$;
end
return Γ_{Pareto}

potentially better fitness. This procedure is repeated until the termination criteria are met, *e. g.*, maximum generation count or time limitation. In the end, the *GA* yields the *Pareto-optimal solution set* of the multi-dimensional fitness function from which one chromosome can be selected posteriorly. Compared to the classical *GA*, only one of the two operators crossover and mutation are applied to each chromosome in every generation. This is because both operations can preserve partial fitness values from the old chromosome, which helps to speed up the fitness calculation, as explained in the following sections. Chaining both operations can lead to the loss of most cached fitness values and requires full fitness re-evaluation.

4.4.2.1 Fitness Function

Chromosomes are compared by their fitness values, which are calculated by a fitness function that is maximized. The *GA* can handle non-linear, non-convex, and multi-dimensional fitness functions by design. The objectives of each stakeholder are represented by one dimension of the four-dimensional fitness function in Equation (4.52).

$$f_{fit} \colon \mathbb{R}^{K \times |\mathcal{T}|} \to \mathbb{R}^4 \colon x \to (f_1(x), f_2(x), f_3(x), f_4(x)) \,. \tag{4.52}$$

These multi-dimensional fitness vectors are compared by *Pareto-dominance* $\succ_P$ [169], where the comparison operator is defined in Equation (4.53) for chromosomes $x, y \in \mathbb{R}^{K \times |\mathcal{T}|}$.

$$f_{fit}(x) \succ_P f_{fit}(y) \Leftrightarrow \forall j \colon f_j(x) \geq f_j(y) \wedge \exists k \colon f_k(x) > f_k(y) \tag{4.53}$$

Resource Aggregator The resource aggregator maps flexibility requests to the *EMSs* while aiming for the lowest cost and the highest flexibility delivery probability. Referring

to Section 4.2.3, the cost of a flexibility request x is determined by $\gamma^{Pool}(x)$ and the delivery probability is given by $\rho^{Pool}(x)$. The fitness dimension of the resource aggregator is expressed in Equation (4.54), where $w_1, w_2 \geq 0$ allows for biasing.

$$f_1 \colon \mathbb{R}^{K\times|\mathcal{T}|} \to \mathbb{R} \colon x \to -w_1 \cdot \gamma^{Pool}(x) + w_2 \cdot \rho^{Pool}(x) \tag{4.54}$$

Grid Operator The goal of the grid operator is to reduce grid losses due to resistive heating in grid assets [3]. Therefore, reported grid profiles $P_G(t)$ at the grid connection points of the *EMSs* are used to determine grid losses in Equation (4.55).

$$f_2 \colon \mathbb{R}^{K\times|\mathcal{T}|} \to \mathbb{R} \colon x \to -\sum_{t\in\mathcal{T}} \mathcal{L}_{flow}\left(\left(\llbracket P_G(t)\rrbracket_e\right)_{e\in\{1,\dots,K\}}\right) \tag{4.55}$$

Since according to Assumption A12 required power grid data is available, $\mathcal{L}_{flow} \colon \mathbb{R}^K \to \mathbb{R}$ calculates the grid losses at time t using the power flow equations and measures the grid loss at the transformer. These grid losses are compared to a baseline without applied flexibility, hence only the additional grid losses of the flexibility request are considered.

Resource Provider The resource providers want to maximize their achieved self-consumption and energy autarky. Therefore, this fitness dimension measures the aggregated degree of self-consumption and autarky that are achieved after scheduling the flexibility request. The individual self-consumption and autarky, defined in Equations (4.46) and (4.47), are summed and linearized with weights $w_3, w_4 \geq 0$ in Equation (4.56).

$$f_3 \colon \mathbb{R}^{K\times|\mathcal{T}|} \to \mathbb{R} \colon x \to w_3 \cdot \sum_{e=1}^{K} \llbracket QoE_5\rrbracket_e + w_4 \cdot \sum_{e=1}^{K} \llbracket QoE_6\rrbracket_e \tag{4.56}$$

Service Fairness Provider The last fitness dimension focuses on service fairness among *EMSs*. While f_3 maximizes the linear sum of quality metrics of the *EMS*, it does not consider how even the degree of self-consumption and autarky are distributed among the *EMSs*. Therefore, the service fairness provider aims to achieve similar degrees of self-consumption and autarky among *EMSs*. To not bias the fairness value with the achieved quality metric, the *Fairness Index F* from Equation (2.3) is used. Since self-consumption and autarky are limited in $[0,1]$, the fairness index simplifies to $F(S) = 1 - 2\sigma(S)$, where S is the set of quality metrics of the different *EMSs*. The two fairness indices are linearized with $w_5, w_6 \geq 0$ in Equation (4.57).

$$f_4 \colon \mathbb{R}^{K\times|\mathcal{T}|} \to \mathbb{R} \colon x \to w_5 \cdot F(\{\llbracket QoE_5\rrbracket_e\}_{e=1,\dots,K}) + w_6 \cdot F(\{\llbracket QoE_6\rrbracket_e\}_{e=1,\dots,K}) \tag{4.57}$$

Note that in some cases the fitness function of a chromosome cannot be calculated or is invalid. This is the case when the *MILP* problem of the *EMS* does not contain any solution (hence is not solvable) or the disaggregation schedule causes voltage violations during the power flow analysis in $\mathcal{L}_{flow}$ (with voltage values outside the permitted range of 0.9 p.u. to 1.1 p.u.). These chromosomes are ignored and simply removed from the population set.

Algorithm 4.5: Population fitness clearing.

Input: Population list Π of size μ
Output: Population Π with an updated fitness value

$f_{fit}^{min} \leftarrow (\min_{c\in\Pi}(f_1(c)), \min_{c\in\Pi}(f_2(c)), \min_{c\in\Pi}(f_3(c)), \min_{c\in\Pi}(f_4(c)))$;
$\Pi \leftarrow sorted(\Pi)$; `// Sort population to descending fitness values`
for $i = 0, \ldots, \mu - 1$ **do**
  $winners \leftarrow 1$;
  for $j = i+1, \ldots, \mu - 1$ **do**
    if $d_{chrom}(\Pi[i], \Pi[j]) < \sigma_{clear}$ **then**
      if $winners < \kappa_{clear}$ **then**
        $winners \leftarrow winners + 1$;
      else
        $f_{fit}(\Pi[j]) \leftarrow f_{fit}^{min}$;
      end
    end
  end
end
return Π

4.4.2.2 Genetic Operators

This section details the operators of the *GA* from Algorithm 4.4.

Diversity There is a chance that mutation and crossover operators produce many similar offspring chromosomes, such that the whole population narrows down to a limited region in the search space or optimizes towards a local optimum. *Population clearing* is used to ensure diversity in the population set [161]. This method uses the *Manhattan distance* between two chromosomes x and y, which is defined for fitness values in Equation (4.58) and chromosome matrix representations in Equation (4.59).

$$d_{fit}(x,y) = \sum_{i=1}^{4} |f_i(x) - f_i(y)| \tag{4.58}$$

$$d_{chrom}(x,y) = \sum_{e=1}^{K} \sum_{t\in\mathcal{T}} |x_{e,t} - y_{e,t}| \tag{4.59}$$

The goal is to identify clusters of chromosomes with very similar matrix representations with a maximum distance of $d_{chrom}(x,y) \leq \sigma_{clear}$. The clearing procedure as defined in [133] is detailed in Algorithm 4.5. In each cluster, κ_{clear} chromosomes with the highest fitness values, called the winners, remain unchanged, while the fitness values of the other chromosomes are reduced to the lowest value per dimension in the population. As a result, the winners of a cluster will have a higher chance to be selected in the selection procedure, and similarity clustering is reduced.

Selection The selection operator is based on tournament selection, where a subset of chromosomes $\Pi_p^{sel} \subseteq \Pi_p$ with tournament size $k_s = \left|\Pi_p^{sel}\right|$ is randomly chosen from the current population and the best chromosome c of the subset is copied to the next generation. This process is repeated until μ_{sel} chromosomes are selected. In the case of Pareto-equivalence, the chromosome with a higher linearized fitness value wins.

Pareto-optimal Solutions and Random Chromosomes Because the population size is limited to μ, the known Pareto-optimal solution set is stored separately from the population. To avoid overlooking chromosomes with high fitness, μ_{pareto} chromosomes from the Pareto-optimal solution set are re-added to the population. This implements an *elitist* strategy, where chromosomes with already high fitness may further improve via crossover and mutation. Additionally, to avoid optimizing towards a local optimum, μ_{new} random chromosomes are generated similarly to the initial population and are added to the population.

Crossover The crossover operator takes two parent chromosomes x and y, and crosses them at random locations to create one new chromosome z. Because all genes need to fulfill the requirement from Equation (4.36), the new chromosome matrix z is created by using column vectors either from x or y, which already satisfy this condition. The *splitting-point* method is used, whereby the number of splitting points $k_c \geq 0$ is sampled from a Poisson distribution with λ_c. The unique random splitting points $\{t_1, \ldots, t_{k_c}\}$, for which $\forall i \in \{1, \ldots, k_c\} \colon t_i \in \mathcal{T} \wedge \neg \exists j \neq i \colon t_j = t_i$, define which genes inherit from parent x and which from y. The new chromosome z is created by copying column vectors from x until the next splitting point is reached, followed by columns from y, and so on. In addition to z, the crossover operator automatically creates the inverse chromosome z' which takes the column vectors from the opposite parent chromosomes. Note that the Manhattan distance between the two parent chromosomes x and y is equivalent to the distance between the two offspring chromosomes z and z', which follows from the definition of the distance function in Equation (4.59).

The newly created chromosome is a time-stripped mixture of the two parent chromosomes, where partial results from the steady state power flow calculation in f_2 can be re-used. Therefore, it is guaranteed that the new chromosome will not violate any voltage limitations, however, the *EMSs* may become infeasible.

Mutation The mutation operator takes one chromosome x and creates a variant z of that chromosome. It exploits the fact that the total flexibility per gene (column vector) has to remain constant, hence, only the flexibility distribution among the *EMSs* may change. For each time slot $t \in \mathcal{T}$, the mutation operator first samples a random number of mutations $k_m \geq 0$ to be applied to the corresponding gene from a Poisson distribution with λ_m. Second, for each mutation two random unique *EMS* $i, j \in \{1, \ldots, K\} \colon i \neq j$ (matrix rows) are selected. To increase the probability that the resulting chromosome is feasible from the *EMS* perspective, the flexibility potential of both *EMSs* is evaluated. The range in Equation (4.60) defines the possible flexibility shift from *EMS* i to j, from which a random value is chosen. In total, $k_m \geq 0$ mutations are applied to each gene of chromosome x.

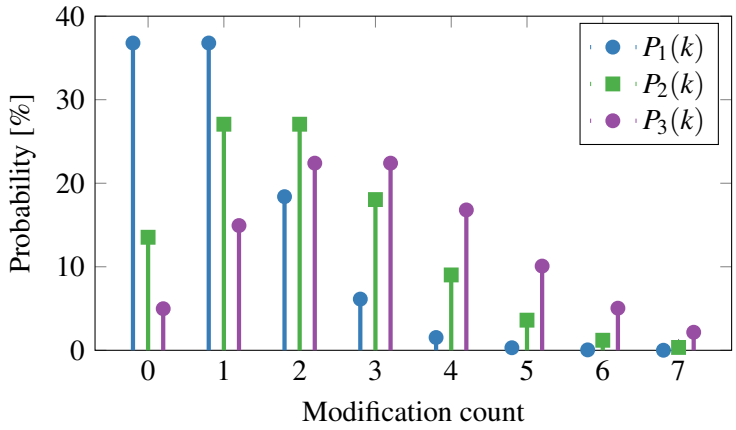


Figure 4.3: Poisson distribution $P_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda}$ with $\lambda = 1,2,3$.

$$[0, \min(\left[\!\left[-P^{min}_{flex}(t)\right]\!\right]_i, \left[\!\left[P^{max}_{flex}(t)\right]\!\right]_j)] \tag{4.60}$$

It can be seen that mutation does not modify all *EMSs*, hence the remaining *EMSs* still have the same cost, probability, self-consumption, and autarky, which can be cached.

Figure 4.3 shows the Poisson distribution of k_c and k_m with different configurations of the parameters λ_c and λ_m. With $\lambda_c = \lambda_m = 2$, this gives an approximate 13.5 % chance for a chromosome to stay unmodified by the crossover operator and a $13.5^{|\mathcal{T}|}$ % chance for the mutation operator. However, depending on the population, their fitness value may get modified by the diversity operator, or the selection operator may drop the chromosome for the next generation.

4.4.3 Discussion

The linear and the meta heuristic target similar objectives, however, the methods to reach these objectives are fundamentally different. For example, minimization of flexibility disaggregation cost is achieved by the disaggregation priority function p^{COST} and objective function f_1 with $w_1 = 1$ and $w_2 = 0$. A second example is the flexibility delivery probability which is optimized by the priority function p^{POPT} and objective function f_1 with $w_1 = 0$ and $w_2 = 1$. Both have in common that the *EMSs* are considered as black boxes, which enables interoperability of different *EMS* formulations. The linear heuristic can only optimize towards a single objective, which is further limited by the required properties of the disaggregation priority function $p(.)$. On the other hand, the meta heuristic is more generalized and supports multiple objective functions with no property restrictions. Since the linear heuristic schedules flexibility for a single time slot after another, it is more suitable for near real-time operation, where scheduled flexibility at past time slots cannot be changed anymore and future flexibility requirements are not known yet. The meta heuristic is better suited for day-ahead optimization, as it can explore the interaction of flexibility disaggregation between multiple time slots.

To disaggregate a flexibility request $P_{flex}^{Pool}(t)$ to *K EMSs* at a single time slot t, the linear heuristic needs to calculate the priority value once for each *EMS*, which is assumed to be constant in $\mathcal{O}(1)$. After assigning a flexibility portion to one *EMS*, only the priority value of that single *EMS* must be updated, other priority values can be cached. The number of required update steps depends on the flexibility request $P_{flex}^{Pool}(t)$ and the configured maximum step size S^{max} and is therefore independent of the number of *EMSs K* in the flexibility pool. The overall computational effort for multiple time slots $\mathcal{T}$ is estimated in Equation (4.61).

$$\mathcal{O}\left(|\mathcal{T}| \cdot \left(K + \left|\frac{P_{flex}^{Pool}(t)}{S^{max}}\right|\right)\right) \tag{4.61}$$

Note that the accuracy of the algorithm to achieve the expectations defined by $p(.)$ depends on S^{max} and the shape of the disaggregation priority function. Smaller step sizes can improve the accuracy, but also impact the computation time.

The meta heuristic processes a population set of μ chromosomes over several generations. The initial population is created once at the beginning of the algorithm and schedules flexibility portions to the *EMSs*, where for each portion the upper and lower flexibility limits of one *EMS* are calculated in assumed $\mathcal{O}(1)$. The expected maximum computation time for μ chromosomes is estimated in Equation (4.62).

$$\mathcal{O}\left(\mu \cdot |\mathcal{T}| \cdot \left|\frac{P_{flex}^{Pool}(t)}{S^{max}}\right|\right) \tag{4.62}$$

Assuming that an upper limit of G generations is executed by the *GA*, the fitness value of each chromosome must be evaluated $\mathcal{O}(\mu \cdot G)$ times. In practice, the genetic operator design allows the fitness function, which can be computationally intense, to be partially re-evaluated and results from the unmodified part of the genes to be reused, which can drastically improve the expected worst case execution time. During each generation, all the genetic operators are executed once. Updating the Pareto-set requires in the worst-case $\mathcal{O}(\mu \cdot \mu \cdot G)$ fitness value comparisons, since each generation out of G may potentially yield μ Pareto-equivalent chromosomes, to which the current generation with size μ is compared. Note that this is a very conservative complexity estimation as not every chromosome will be part of the Pareto-front. The diversity operator must additionally compare the chromosomes of the current population with each other in $\mathcal{O}(\mu \cdot \mu)$. The selection operator samples μ_{sel} tournaments with a maximum of $k_s = \mu$ participants hence runs in $\mathcal{O}(\mu \cdot \mu)$, μ_{new} chromosomes are created similar to the initial population as estimated in Equation (4.62), and adding μ_{pareto} chromosomes from the Pareto-set is limited in $\mathcal{O}(\mu)$. The crossover operator performs a maximum of $k_c = |\mathcal{T}|$ splitting points in $\mathcal{O}(\mu \cdot |\mathcal{T}|)$ and the mutation operator runs in $\mathcal{O}(\mu \cdot |\mathcal{T}| \cdot k_m)$ since for each chromosome and time slot a maximum number of k_m mutations require to calculate the possible flexibility limits of two *MILPs*. The overall complexity of the *GA* is summarized in Equation (4.63). As can be seen, the most impacting parameter is the population size μ, followed by the generation count G. The *GA* is independent of the number of *EMSs* K of the flexibility pool. Comparing Equations (4.61) and (4.63), it can be seen that the *GA* comes with an at minimum $\mu \cdot G$ higher computational complexity, which is driven by the number of chromosomes per

generation and the number of processed generations.

$$\mathcal{O}\left(\mu\left(|\mathcal{T}|\cdot\left|\frac{P_{flex}^{Pool}(t)}{S^{max}}\right|+G\cdot\left(G+\mu+|\mathcal{T}|\left(k_m+\left|\frac{P_{flex}^{Pool}(t)}{S^{max}}\right|\right)\right)\right)\right) \tag{4.63}$$

Finally, it is possible to combine both heuristics, such that the flexibility schedules from the linear heuristic are used to generate a more advanced initial population set of the *GA*. A maximum of $4\cdot|\mathcal{T}|!$ possible variations of the same flexibility request can be created using the four different disaggregation priority functions EQUAL, PROP, COST, and POPT, and different scheduling order of the time slots in line 1 of Algorithm 4.2. Finally, the initial population of the *GA* can be created from a mixture of chromosomes from the linear heuristic with a set of random chromosomes for exploration possibilities.

4.5 Evaluation

This section first describes the hardware setup and scenario description for the evaluation, including the data and forecast models for the *EMSs*, as well as the underlying low-voltage power distribution grid. In the second part, the performance of the *MILP* formulation is evaluated before the effectiveness of the different *GA* operators is demonstrated. Furthermore, the optimal hyper-parameters of the *GA* are determined and fixed by a *one-at-a-time* sensitivity analysis. The linear heuristics as well as the *GA* is applied to the use case scenario and evaluated with regard to *QoS*, *QoE*, and fairness, as well as their impact on the power system. Three additional realistic use case scenarios for flexibility disaggregation, for which the impact of the different flexibility disaggregation policies is analyzed, extend the analysis.

4.5.1 Setup

The *MILP* problem is modeled using the linear programming library *PuLP*[17] in *python*[18] and is solved by the commercial solver *gurobi*[19]. The execution time highly depends on the *MILP* implementation and the used solver, especially because both the linear heuristic and the meta heuristic require iterative optimal schedules of all *EMSs* of the flexibility pool. The *EMS* and the *GA* simulations are executed on a single server setup using 32 threads on an *Intel Xeon* processor, which allows exploiting the parallelism of the *GA* that resulted in a speedup of approximately 19.8 between a single thread and 32-thread execution. This large gap is due to the required synchronization during the *GA* execution, since only the *CPU*-intense tasks, *e. g.*, creation of new random chromosomes, mutation operator, and fitness evaluation, are executed in parallel. Furthermore, a fitness value cache stores partial results of the fitness function, *e. g.*, power flow solutions of individual time slots and optimal schedules of the *EMS*, which results in an additional speedup of approximately 1.2. In total,

[17] https://pypi.org/project/PuLP (version 2.6.0)
[18] https://www.python.org (version 3.9.10)
[19] https://www.gurobi.com (version 9.5.2)

parallelization and fitness caching achieve a speedup of 23.76 on the use case scenario described in the following section.

4.5.1.1 Energy Management System and Data

To perform home *EMS* optimization, forecast models for the *PV* generation $P_{PV}(t)$ and the household load $P_{HH}(t)$ are required. In this study, a *Long-Short Term Memory* (*LSTM*)-based *PV* forecast model and a *similar-day* household load forecast model are used, which are presented by co-authors in [37]. However, the forecast models are interchangeable and the power flexibility provision works with any model that can provide the forecast error distribution function or empirical error data as stated in Assumption A16. The *LSTM* forecast model is trained on 5 different real *PV* systems for the next 24 h using only historical weather data, similar to [2, 99, 110, 145]. A *similar-day* model is applied for household load forecast, which simply yields the average load profile of similar weekdays over the last four weeks. Household load profiles from 10 different real households are used in this study. Both *PV* and load profiles are recorded by *OpenEMS* [54] at real household installations located in Central Europe using the anonymized data that is available on the *Open Energy Platform* [98]. More details on the forecasting models can be found in [37].

The *EV* driving patterns are the same as described in Section 3.5.1.1. The *EV* availability is determined by the arrival and departure times as given in Equation (4.64).

$$\alpha_i(t) = \begin{cases} 1 & if\, t_{arr} \leq t < t_{dep} \\ 0 & else \end{cases} \tag{4.64}$$

The driving distance converts to required energy $E^{req}_{EV,i}$ with an assumed energy consumption of 17 kWh per 100 km and a realistic average battery storage capacity of $E^{max}_{EV,i} = 40\,\text{kWh}$, similar to Chapter 3. All charging processes start with an initial energy level of 10 %. The charging power of the *EV* is limited to values defined by IEC 61851-1 with $P^{min}_{EV,i} = 4.3\,\text{kW}$ and $P^{max}_{EV,i} = 11\,\text{kW}$ and a charging efficiency of $\mu_{EV,i} = 0.9$.

For each *EMS*, the upper and lower grid limits $P^{min}_{G}(t)$ and $P^{max}_{G}(t)$ are fixed to ±20 kW, which represents the fuse limit at the grid connection point. Since nowadays energy tariffs are most often still constant for private customers in Germany, the energy cost $c_{buy}(t)$ is assumed constant.[20] Nevertheless, the *MILP* model works with time-of-use prices as well. A typically sized *ESS*[21] is specified for each *EMS* with $P^{max}_{ESS} = 9\,\text{kW}$, $E^{max}_{ESS} = 12\,\text{kWh}$, $\mu_{ESS} = 0.9$, and an initial energy level of 10 %. Furthermore, with the same reasoning as in Section 3.5.1.1, between 0 and 2 *EV* charging processes from the aforementioned charging patterns are randomly assigned to each *EMS*. If not noted otherwise, all simulations are carried out with one-hour resolution, hence $T = 24$, and $\Delta_t = 1$. In the case of quarter-hourly simulation, the parameters are set to $T = 96$ and $\Delta_t = 0.25$.

[20] The actual value does not matter in the linear objective function in Equation (4.16)

[21] The recorded households in [98] use the FENECON Pro 9-12 storage system with a 12 kWh Lithium iron phosphate (LiFePO4) battery pack and 9 kW peak power. (https://fenecon.de/en/fenecon-pro-9-12)

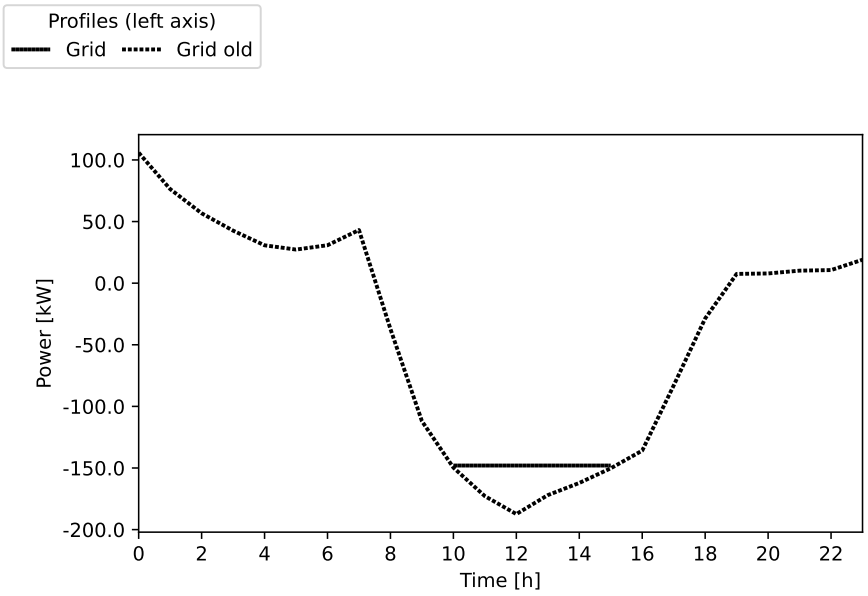


Figure 4.4: Use case scenario for power flexibility scheduling to a flexibility pool. The dotted line shows the original aggregated grid connection profile of 55 *EMS* and the solid line shows the grid profile after peak shaving of reverse power flow between 10 h to 15 h.

4.5.1.2 Power Grid and Simulation Scenario

The use case scenario builds on the IEEE 906 low-voltage test feeder, which is depicted in Figure 3.7 and detailed in Section 3.5.1.2. This low-voltage grid is simulated by the power flow solver *pandapower*[22]. Instead of the 55 static household load profiles as provided by the benchmark grid, one *EMS* is connected to each of the household grid connection points with assumed phase-balanced connection.

During the day, high *PV* generation causes reverse power flow from the low-voltage grid to the upper-level power system. To reduce this reverse power flow, the *DSO* may be interested to schedule flexibility to the flexibility pool that includes all the 55 *EMSs*. The investigated use case scenario aims to perform peak shaving of the reverse power flow to approximately −150 kW, as depicted in Figure 3.15. The required flexibility request is defined by the time horizon $\mathcal{T} = \{10, 11, 12, 13, 14, 15\}$ and the flexibility request given in Equation (4.65), which schedules a total energy demand of more than 106 kWh, which is approximately 16 % of the aggregated *ESS* capacity of the *EMSs*.

$$P_{flex}^{Pool} = (-1\,677, -24\,835, -39\,392, -24\,022, -14\,137, -2\,253) \tag{4.65}$$

[22] https://www.pandapower.org (version 2.10.1)

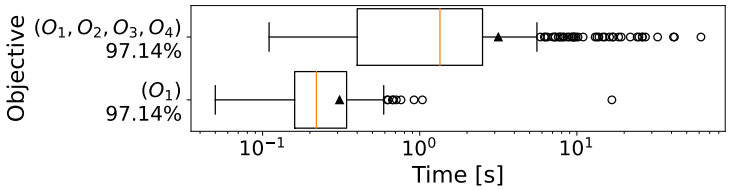


Figure 4.5: Computation time and feasibility ratio of 350 *EMSs* within 2 min. Note that the time axis is on logarithmic scale.

4.5.2 Analysis

This section first evaluates the performance of the *MILP* formulation on the given hardware and validates the efficiency of the genetic operators, including a detailed analysis of the hyper-parameters of the *GA*. Afterward, the linear heuristic and the meta heuristic are compared using the aforementioned use case scenario. The resulting disaggregation schedules of both methods are analyzed on their achieved service quality metrics and fairness among the *EMSs* as defined in Section 4.3.

The fitness function weights are set to $\forall i = 1,\ldots,6\colon\ w_i = 0.5$ by default. To better compare the fitness of chromosomes, each dimension of the fitness value is normalized to the minimum and maximum fitness dimension values of 1000 randomly created chromosomes using Algorithm 4.3. Values of f_1 are within $[-6.087, -1.658]$, $f_2 \in [62279, 62625]$, $f_3 \in [37.838, 38.704]$, and $f_4 \in [0.568, 0.598]$. Consequently, fitness dimensions with negative values evaluate worse than 1000 random chromosomes, and fitness values above 1.0 evaluate better. In the following, f_x refers to the normalized fitness value of dimension x. Finally, to directly compare chromosomes with each other, the multi-dimensional normalized fitness vector (f_1, f_2, f_3, f_4) is linearized with equal weights of 0.25 per dimension to $f_{lin} = 0.25\ f_1 + 0.25\ f_2 + 0.25\ f_3 + 0.25\ f_4$.

4.5.2.1 Performance of the Energy Management System

To evaluate the computation time of the *MILP*, 350 different realistic *EMS* scenarios are sampled from 10 household profiles, 5 *PV* profiles, and 7 different summer days. With quarter-hourly time resolution, the number of decision variables per *EMS* varies with the number of assigned *EVs*. The mean number of decision variables is 1218 (min 584, max 2116), of which 19.8 % (min 16.4 %, max 22.7 %) are binary variables, and the mean number of constraints is 1887 (min 1249, max 2789). As can be seen in Figure 4.5, most *EMS* scenarios can be solved within the time limitation of 2 min using only a single thread of the multi-core *CPU*. Solving objective function O_1 for cost estimation is on average drastically faster than solving all hierarchical objectives, which in turn positively influences the computation time of the COST flexibility disaggregation.

The performance of the *MILP* can be quantified by the degree of self-consumption and energy autarky as defined by the quality metrics QoE_5 and QoE_6 in Equations (4.46) and (4.47). The box plots in Figure 4.6 show the distribution of the metric values, where the

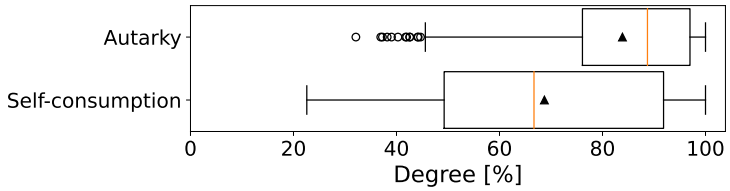


Figure 4.6: Degree of self-consumption and autarky of 350 *EMS* scenarios.

MILP can not always reach the maximum degree of self-consumption and autarky for all *EMS*, because in many *EMS* scenarios, the *PV* generation is smaller than the total consumption and highly misaligned with the consumption profile. In [37], it was demonstrated that the average achieved degree of self-consumption and autarky using specific forecast models (*LSTM* for *PV* generation and *similar day* forecast for household load) is lower compared to assumed perfect forecasts. Consequently, forecast inaccuracies impact the optimality of the planned profiles to a non-negligible degree. This should be kept in mind during the analysis of flexibility provisioning where scheduled power flexibility always needs to be provided.

The *EMS* constraints in Equations (4.14) and (4.15) prevent the *ESS* to charge from the grid or to discharge to the grid. These constraints are required to prohibit the *EMS* to use its *ESS* to participate in the energy market and cause unpredictable situations in the power distribution grid. However, these limitations drastically impact the possible flexibility that can be scheduled to an *EMS*. As can be seen in Figure 4.7, the flexibility range of a flexibility pool consisting of 150 *EMSs* is limited on average to only 30.37 % of the possible flexibility. In addition, *EV* charging provides mostly negative flexibility at night, and *PV* generation provides most of the positive flexibility during the day. As a result, these *EMS* constraints should only be used for local optimization and be neglected when it comes to flexibility scheduling to make maximum use of the local *ESSs*. Therefore, in the following, the *EMS* uses a *MILP* formulation without the constraints in Equations (4.14) and (4.15).

4.5.2.2 Validation of the Genetic Operators

For a *GA* with a multi-dimensional fitness function, not only the linearized fitness value of the best chromosome counts but also the diversity within the Pareto-set. On the one hand, high fitness diversity ensures that the Pareto-front of the optimization problem is well explored. On the other hand, a high chromosome diversity of a population shows that the search space of all possible chromosomes is traversed well. The fitness and chromosome diversity of a set of chromosomes Π are defined based on the mean *Manhattan distance* as defined in Equations (4.58) and (4.59) and are calculated according to Equations (4.66) and (4.67).

$$div_{fit}(\Pi) = \frac{1}{2\,|\Pi|} \sum_{x \in \Pi} \sum_{y \in \Pi} d_{fit}(x,y) \tag{4.66}$$

$$div_{chrom}(\Pi) = \frac{1}{2\,|\Pi|} \sum_{x \in \Pi} \sum_{y \in \Pi} d_{chrom}(x,y) \tag{4.67}$$

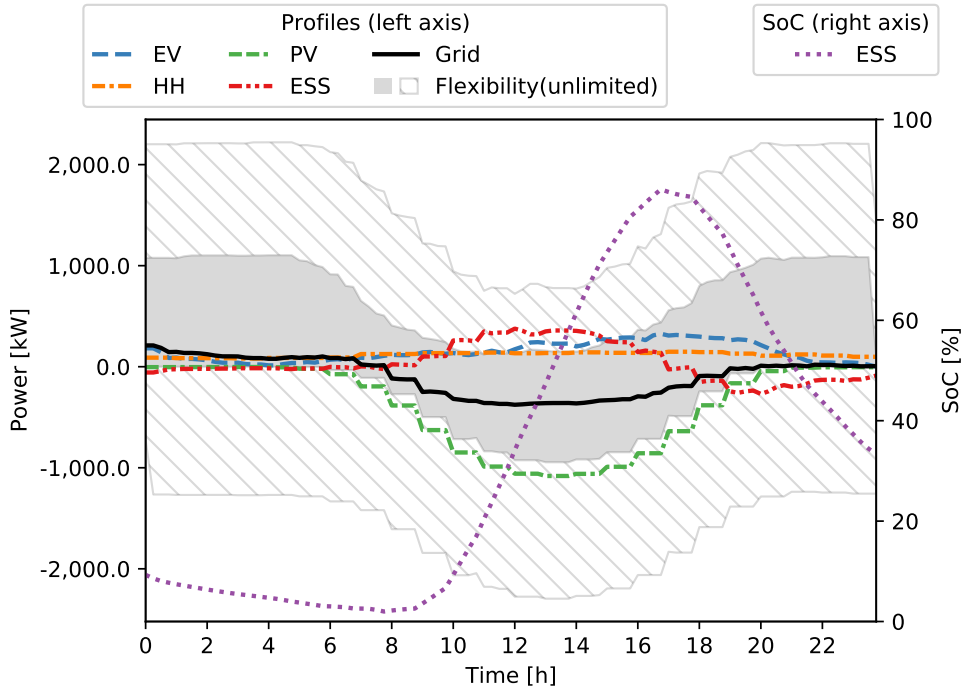


Figure 4.7: Optimal profiles of a flexibility pool with 150 *EMSs* [37].

The chromosomes of a population with small chromosome diversity are located near each other, and the *GA* may miss regions of the search space. On the contrary, large chromosome diversity in a population corresponds to chromosomes being widespread in the search space, which in the case of too large diversity may not result in an overall optimal solution. The fitness diversity of the Pareto-set quantifies the exploration of the multi-objective solution space, where a Pareto-set with larger fitness diversity contains a more widespread Pareto-front, from which one chromosome is selected posteriorly.

Genetic Operators First, the performance of creating new random chromosomes for the use case scenario is evaluated. The creation of 1000 random chromosomes using Algorithm 4.3 takes on average 17.56 s per chromosome on a single thread of the hardware setup. The algorithm yields a very small failure rate of only 0.4 % of invalid chromosomes and the resulting population set contains 66 dominating chromosomes as defined in Equation (4.53). According to Table 4.2, the computation time is larger compared to the other genetic operators because iterative calculations of the *MILP* variable boundaries are required. Nevertheless, new chromosomes are mainly created once before entering the generations' loop of the *GA* and, therefore, the time required for generating new chromosomes has only a small effect on the overall execution time of the *GA*. The chromosome diversity of the random population Π_{rand} shows how well the chromosomes are distributed in the search space. Assuming that only equally signed flexibility is scheduled to the *EMSs*, like it is done in Algorithm 4.3, the maximum possible Manhattan distance between two chromosomes is

Table 4.2: Genetic operator evaluation including the quality of the offspring chromosomes and the resulting Pareto-set of 1000 initial chromosomes.

Operator	Failure rate	Parent p			Population			Time
		$f_{fit}(x) \succ_P f_{fit}(p)$	Pareto equivalent	$f_{fit}(p) \succ_P f_{fit}(x)$	Pareto-front size	$div_{fit}(\Gamma_{Pareto})$	$div_{chrom}(\Pi_{rand})$	
	[%]	[%]	[%]	[%]	#			[s]
Random	0.4	-	-	-	66	0.77	164 288	17.56
Crossover ($\lambda_c = 1$)	53.0	0.8	45.7	0.5	38	0.87	164 332	0.01
Crossover ($\lambda_c = 2$)	76.6	0.8	22.0	0.6	31	0.69	164 065	0.01
Crossover ($\lambda_c = 3$)	90.0	0.4	9.3	0.3	16	0.79	163 831	0.01
Mutate ($\lambda_m = 1$)	5.0	13.0	77.7	4.3	80	0.87	184 856	6.19
Mutate ($\lambda_m = 2$)	19.1	11.6	68.8	0.5	68	0.99	203 396	7.51
Mutate ($\lambda_m = 3$)	34.4	9.5	56.1	0.0	46	1.24	221 227	9.58

limited by the flexibility request vector $(P_{flex}^{Pool}(t))_{t\in\mathcal{T}}$, and is given by $div_{chrom}^{max} = 212.632$. The chromosome diversity $div_{chrom}(\Pi_{rand})$ covers 77 % of div_{chrom}^{max}, which means that the search space is well explored and the proposed creation of random chromosomes is suitable for exploration of the given search space.

To demonstrate that the crossover and the mutation operator create new feasible and improved versions of their parent chromosomes, an empirical analysis is conducted that applies each of the operators to the same population set of 1000 random chromosomes. According to Table 4.2, the crossover operator configured with $\lambda_c = 1$ produces 47 % feasible chromosomes. The relatively high infeasible rate, which even increases with higher values for λ_c, is explained by the simplistic crossover procedure that does not utilize *EMS*-specific logic. Independent of λ_c, most of the feasible offspring chromosomes are Pareto-equivalent to their parents, which shows that the crossover operator aims for *exploitation* [30]. Since the crossover operator only combines parts of the parent chromosomes, the Manhattan distance of the offspring chromosomes is identical to the distance between its parents. Consequently, the chromosome diversity of the resulting population $div_{chrom}(\Pi)$ stays nearly the same compared to the random population, with only small variations due to the infeasible chromosomes.

The mutation operator configured with $\lambda_m = 1$ creates feasible chromosomes with 95.0 % probability, which is due to the definition of the mutation procedure that only shifts flexibility between *EMSs* within their flexibility limits. However, since only the upper and lower boundaries of the flexibility potential are considered, *EVs* that require a minimum charging rate may still cause the chromosome to become infeasible, especially when multiple mutations are applied at once with $\lambda_m = 2$ and $\lambda_m = 3$. With $\lambda_m = 1$, around 4.3 % of the feasible offspring yields worse fitness values, while 13.0 % dominate their parents. The

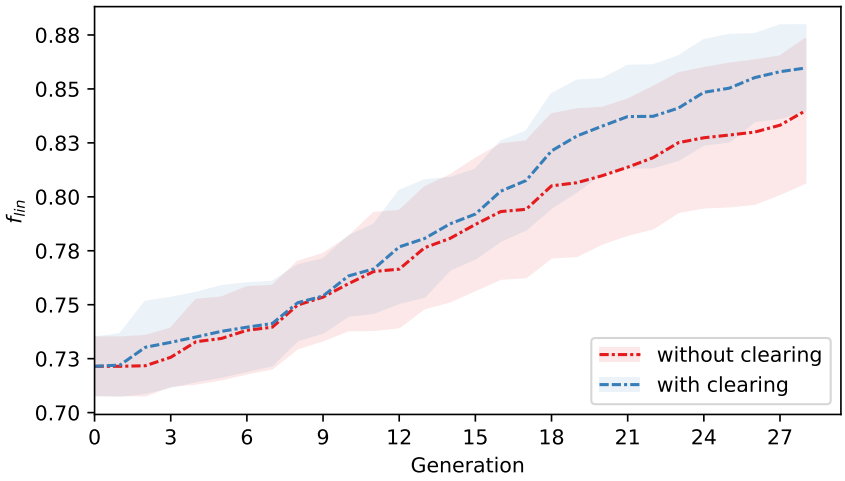


Figure 4.8: Impact of population clearing using the clearing operator. The line shows the mean linearized fitness value and the shaded area represents the 95 % confidence interval.

size of the Pareto-front is slightly larger with better fitness diversity compared to the random population. The mutation operator creates diverse chromosomes (*exploration*), which is underlined by the high probability of Pareto-dominating and Pareto-equivalent offspring and the increased chromosome diversity of the new population, especially with higher λ_m values. Note that with $\lambda_m = 3$, the chromosome diversity even overshoots div_{chrom}^{max}, since the mutation operator also produces many chromosomes with opposite signs of flexibility at the *EMSs*.

The following experiment highlights the impact of the diversity operator on the population throughout the generations of the *GA* execution. Therefore, the *GA* is executed six times with random initial populations of size $\mu = 50$ and a time limit of 2 h. Figure 4.8 shows the mean linearized fitness value of the best chromosome that was found until the given generation on the x-axis and its 95 % confidence interval. It can be seen, that fitness clearing with the default parameters from Table 4.3 helps to slightly improve the performance of the *GA*. Note that this graph is cut at the minimum generation count of the random executions, hence depicts the worst-case performance. However, the mean linearized fitness value of the last generation of each *GA* execution shows similar results, where without clearing $f_{lin} = 0.850(\pm 0.028)$ and with clearing $f_{lin} = 0.861(\pm 0.030)$ is achieved.

Finally, the proposed *GA* approach, where crossover and mutation are applied to only half of the population, is compared to the classical *GA*, where the crossover operator is first applied to the entire generation before invoking the mutation operator. Figure 4.9 shows the mean fitness development and its 95 % confidence interval of six *GA* runs with $\mu = 50$ and a time limit of 1 h. As can be seen, the minimum number of processed populations is approximately 17.6 % larger with the modified approach compared to the classical *GA*. Nevertheless, each operator is applied less often, but in total, a larger number of chromosomes is processed during the *GA* execution. Second, the achieved quality of the chromosomes improves drastically compared to the classical *GA*, which can be explained by the

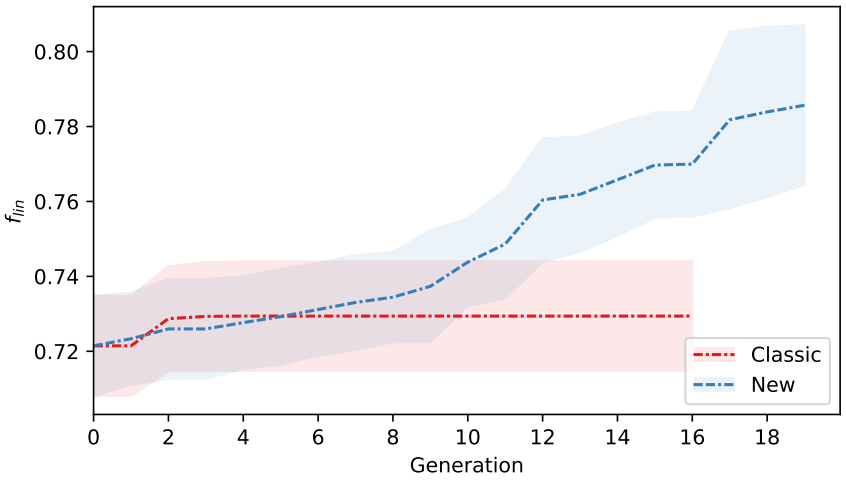


Figure 4.9: Classic *GA* compared to the alternating operator approach comparing six *GA* runs over 1 h.

more fine granular search space traversal, since by design only one of the two operators is applied at once. In the classical *GA*, the probability to apply only one of the two operators is much lower. Note that Figure 4.9 again shows only the minimum generation count from the six simulation runs. The final linearized fitness value is given by $0.729(\pm 0.016)$ and $0.811(\pm 0.023)$, respectively.

Impact of Randomness and Pareto-set Exploration The chromosomes should cover a large region of the fitness function optima to allow *a posteriori* selection of the multi-objective criteria. The scatter plots in Figure 4.10 compare the fitness dimensions from the Pareto-optimal set of six random populations with size $\mu = 1000$ to assess the fitness distribution for randomly created chromosomes. As can be seen, all six random Pareto sets cover mostly the same normalized fitness regions, which indicates that their fitness values are well distributed. With multi-dimensional fitness functions, a higher fitness value in one dimension typically tends to decrease the value of another dimension [39], which can be seen in the majority of the scatter plots. However, this does not hold for the fitness dimensions f_1 and f_3, where the fitness values of the random chromosomes correlate. That is reasonable, because with the *MILP* formulation as described in Section 4.2 minimizing the cost of flexibility provision is highly related to the level of autarky and self-consumption according to the cost function in Equation (4.16). Nevertheless, both fitness dimensions are treated separately since in theory, the delivery probability (part of f_1) depends only on the flexibility limits and the forecast error distributions. Results for random chromosomes of population size $\mu = 100$ are similarly well distributed and are provided in Figure A.4.

In addition to the random chromosome creation, also the crossover and the mutation operator rely on random processes. Figure 4.11 shows a 24 h simulation of the *GA* using six random initial populations of size $\mu = 50$. The dotted lines represent the mean fitness value of the best chromosome found after each generation with a shaded 95 % confidence

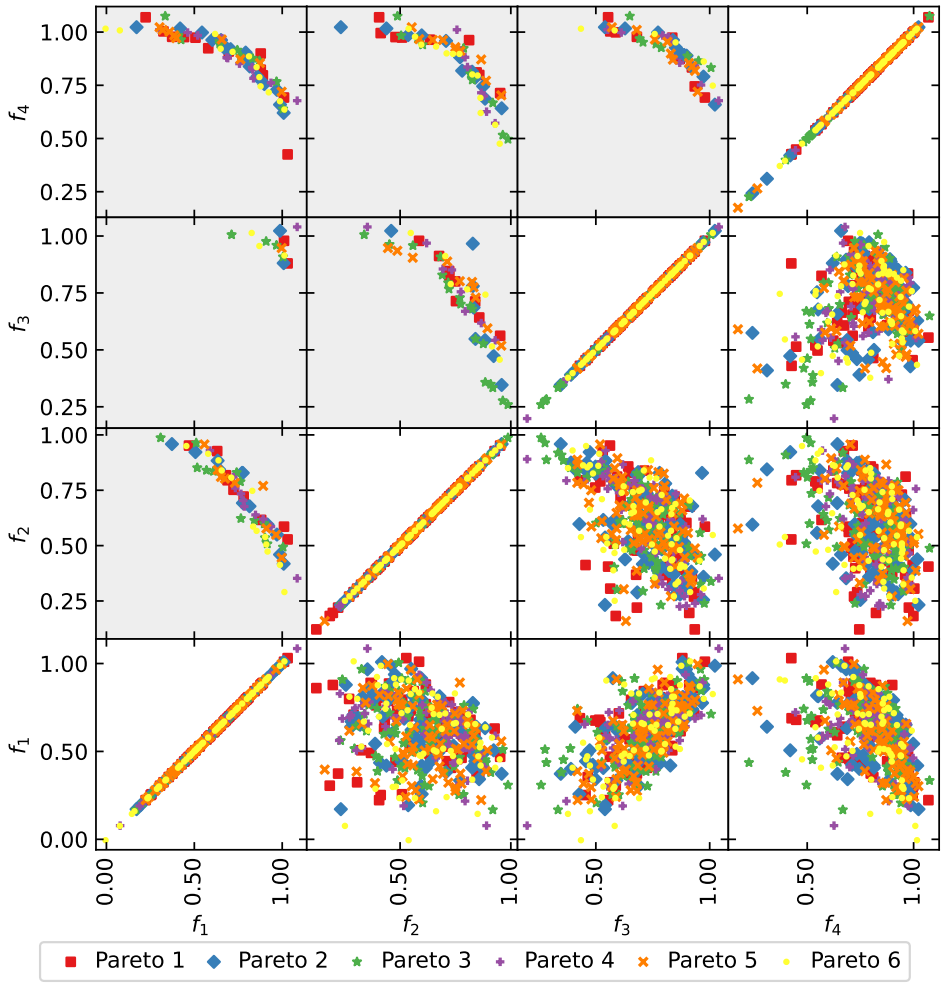


Figure 4.10: Comparison of the fitness dimensions of f_{fit} from the Pareto-optimal solution set of six random population sets of size 1000. Gray-shaded subplots show only the Pareto-dominant chromosomes considering the two dimensions compared.

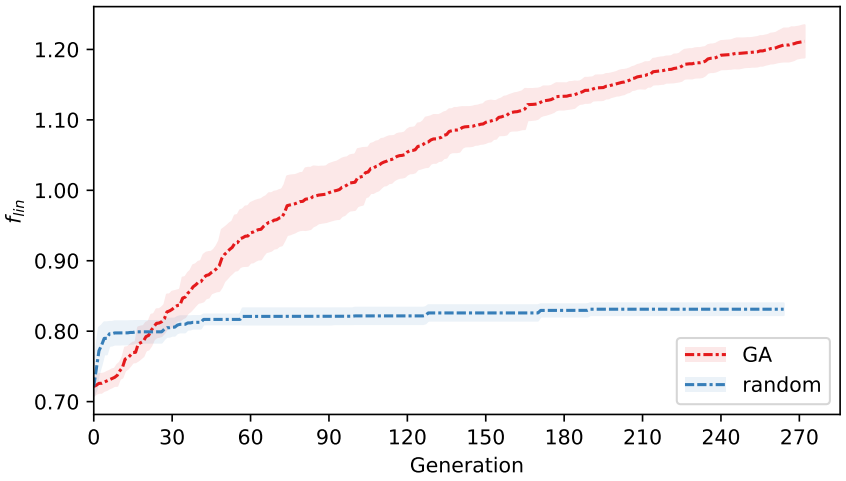


Figure 4.11: Development of the linearized chromosome fitness of six random *GA* runs over 24 h compared with purely random chromosome sampling in each generation.

interval, once for the *GA* execution and once for only sampling new random chromosomes in each generation. The fitness value of the *GA* gradually improves over time but does not converge to an optimum within 24 h. However, the *GA* outperforms the purely random sampling strategy, which is limited to unidirectional flexibility disaggregation following Algorithm 4.3. Since the 95 % confidence interval of the best chromosome after each generation is very narrow, it can be concluded that the randomness introduced by the genetic operators does not impact the performance of the *GA*.

Finally, Figure 4.12 illustrates the resulting Pareto-optimal solution sets of the six *GA* runs that cover most of the optimal regions between any two fitness dimensions in the majority of cases, which offers diverse solutions for *a posteriori* selection. The fitness dimensions f_1 and f_3 have not improved much over 1000 random chromosomes as shown in Figure 4.10, which is because shifting flexibility during surplus *PV* generation is cost-neutral most of the time and has only a small impact on self-consumption and autarky. Furthermore, an increase in delivery probability, self-consumption, or autarky at one *EMS* typically causes a decrease at another *EMS*, which finally yield very similar aggregated values. Whereas, fitness values of f_2 and f_4 have improved by factors of 3.5 and 1.5, respectively. This improvement comes from a better allocation of the flexibility request in the power grid to reduce grid losses and a more fair disaggregation to the *EMSs* of the flexibility pool. Note that this improvement comes with a decrease in fitness in other dimensions. Overall, the *GA* can move the Pareto-front to better results in all fitness dimensions.

Genetic Algorithm Hyper-Parameters As described in Section 4.4.2, the *GA* has plenty of parameters for fine-tuning the algorithm. Testing all combinations of the different parameter configurations is not possible in a reasonable time. Therefore, a *one-at-a-time* sensitivity analysis is performed, where only one parameter is altered while the others stay with their default values as marked bold in Table 4.3.

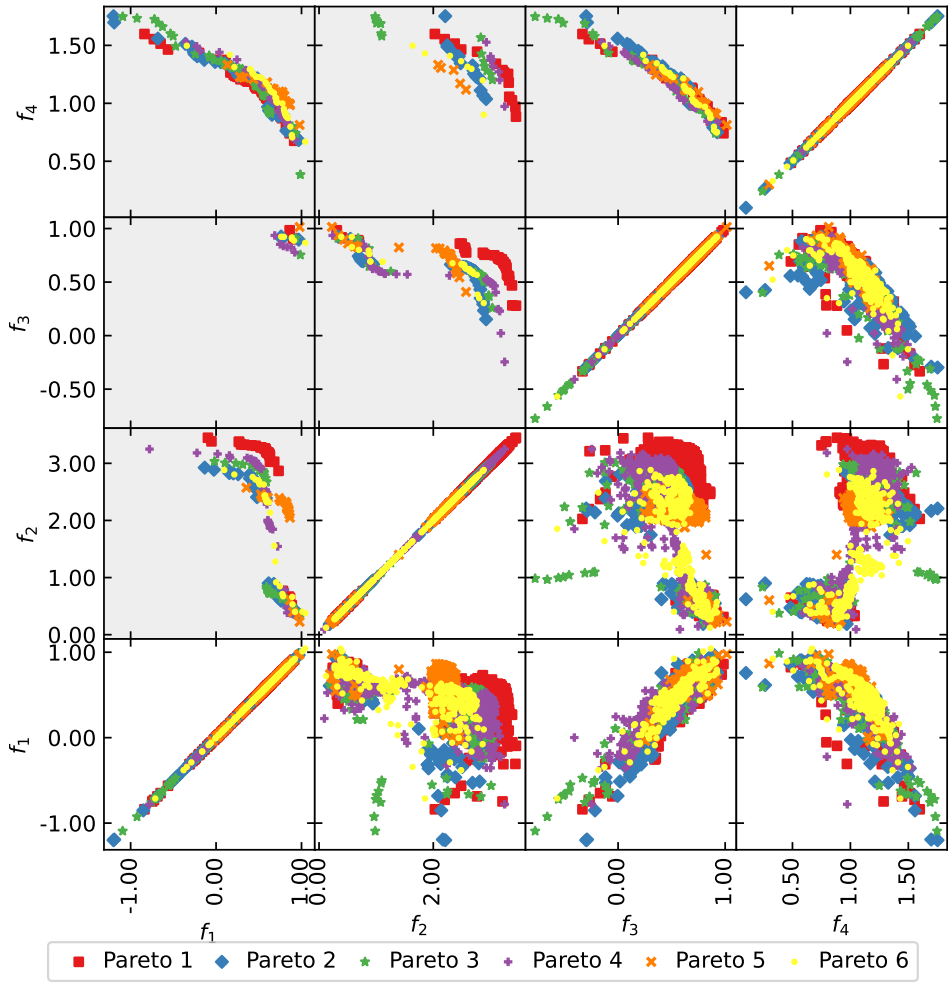


Figure 4.12: Comparison of the fitness dimensions of f_{fit} from the Pareto-optimal solution set after six 24 h *GA* executions with $\mu = 50$. Gray-shaded subplots show only the Pareto-dominant chromosomes considering the two dimensions compared.

Table 4.3: Hyper-parameter selection for the one-at-a-time analysis, where bold parameters show the default configuration. The default parameters are determined by a first guess and multiple one-at-a-time sensitivity analyses where in each step the most influential parameter is fixed to its best configuration in the order of this table.

Hyper-parameter	Symbol	Values
Time limit		60 min
Population capacity	μ	10, **50**, 100, 500 ($\mu_{new} = 0.1\mu$, $\mu_{pareto} = 0.1\mu$)
Mutation Poisson	λ_m	1, **2**, 3
Crossover Poisson	λ_c	1, **2**, 3
Clearing cluster size	κ_{clear}	5 %,**10 %**,15 %,20 % of μ
Clearing distance	σ_{clear}	0.5 %, **1 %**, 3 %, 5 % of div^{max}_{chrom}
Tournament size	k_s	**5**, 10, 15

Table 4.4: Impact of population size on the *GA* performance, comparing the mean linearized fitness values of ten random 1 h *GA* executions. The values in brackets show the 95 % confidence interval and ↑ is the improvement rate compared to the initial population.

μ	# Gen	$\sum$ chrom	Initial fitness	Final fitness	↑ [%]	Worst fitness
10	52.6	526	0.688 ($\pm$ 0.016)	0.718 ($\pm$ 0.020)	4.4	0.666
50	29.4	1470	0.729 ($\pm$ 0.019)	0.843 ($\pm$ 0.023)	15.6	0.790
100	17.1	1710	0.753 ($\pm$ 0.021)	0.817 ($\pm$ 0.017)	8.5	0.771
500	3.3	1650	0.792 ($\pm$ 0.012)	0.794 ($\pm$ 0.011)	0.3	0.779

The population capacity μ influences not only the number of chromosomes per generation but also the computational cost of each generation as determined in Equation (4.63). Together with a time limit, this leads to a trade-off between population capacity μ and generation count G. This relation holds pretty well in Table 4.4, where the number of processed chromosomes is nearly equal for different population sizes, and doubling the population capacity μ nearly halves the number of generations processed within the time limit of 1 h. There is only one exception for the very small population size of $\mu = 10$ where not all threads are fully utilized. With a larger population capacity μ, the best linearized fitness value of the initial population is higher. However, since fewer generations are processed on average, the finally obtained linearized fitness value is worse compared to lower population capacities, where multiple crossover and mutation operations generate better chromosomes. This experiment shows that a population capacity of $\mu = 50$ is reasonable for the use case scenario on the hardware setup described in Section 4.5.1.

Setting λ_c and λ_m to 1 would keep chromosomes unaltered with a probability of $37.5\,\% + 37.5^{|\mathcal{T}|}\%$. On the other side, parameter values of 3 or higher cause multiple modifications at once. This produces larger steps between the old and the new chromosome but also leads to a higher chance of infeasible chromosomes as shown in Table 4.2. Small mutation λ_m equal to 1 and 2 perform similarly well, however increasing λ_m to 3 produces a 12.4 % lower linear fitness value after 1 h of *GA* execution. Since $\lambda_m = 2$ yields better chromosome

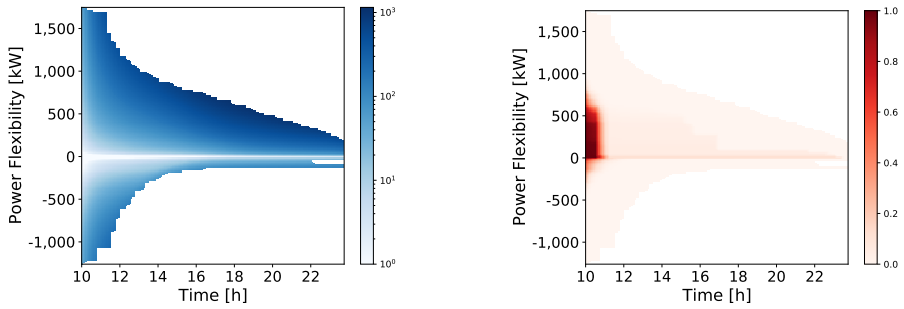
and fitness diversity according to Table 4.2, the default decision of $\lambda_m = 2$ looks reasonable for the use case scenario. A one-hour *GA* execution with $\lambda_c = 2$ produces a 5.3 % better fitness value than with $\lambda_c = 3$ and an even 11.3 % better value than with $\lambda_c = 1$.

Population clearing is configured by the clearing distance σ_{clear} and the cluster size κ_{clear}, whereby σ_{clear} depends on the flexibility request, more precisely on the maximum Manhattan distance div_{chrom}^{max} between two chromosomes, and κ_{clear} depends on the population size μ. Large σ_{clear} will group more chromosomes to the same cluster, from which the best κ_{clear} will keep their fitness value. Experiments show that the parametrization of κ_{clear} has a higher impact on the linearized fitness value with variations in the range of 11.1 % and a peak at $\kappa_{clear} = 0.1$, whereas σ_{clear} shows a slightly smaller impact of only 8.5 % with a peak at $\sigma_{clear} = 0.01$. Finally, the tournament size shows fitness variations of only 1.5 % with the best values for $k_s = 5$.

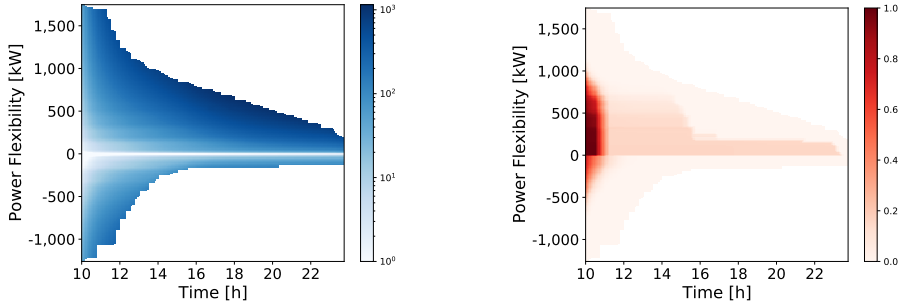
4.5.2.3 Performance of Power Flexibility Provision

To demonstrate the performance of the linear heuristic for iterative scheduling, the proposed disaggregation policies are compared by analyzing a flexibility pool consisting of 150 *EMS*. Thereby, flexibility with different power levels is scheduled starting from time 10 until the flexibility pool cannot provide the required power flexibility anymore. Figure 4.13 shows the cost and probability of the scheduled flexibility for each policy. The cost develops quite similarly for all policies except for COST, which has a lower total cost for small power flexibility at the beginning. Bigger differences can be seen in the probability of flexibility delivery, where POPT outperforms EQUAL and COST. Only the PROP policy achieves comparable results because the disaggregation vector is very similar to POPT due to the similarity of the *CDF* curves. Consequently, cost-optimal disaggregation, as widely proposed in the literature, is not necessarily the best choice when a high delivery probability from distributed *EMSs* is desired.

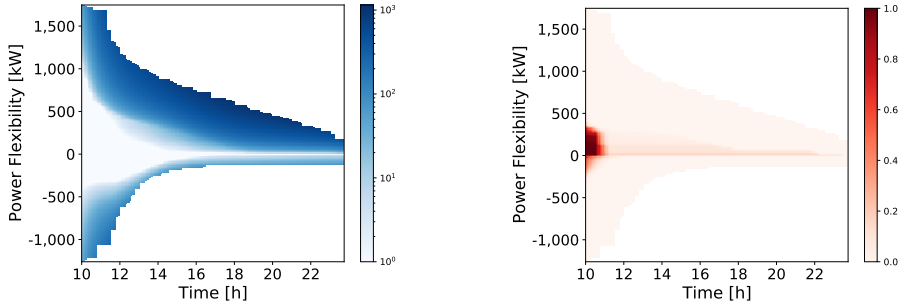
Next, the proposed *GA* is compared with the linear heuristic using the disaggregation policy COST. Therefore, the fitness function is configured to the same convex cost-only objective by setting $w_1 = 1$ and $\forall i = 2,\ldots,6\colon w_i = 0$, such that $f_{fit} = -\gamma^{Pool}(x)$. Despite the fact of this much simpler fitness function, the *GA* cannot compete with the linear heuristic even with a 12 h run, as shown in Figure 4.14. The linear heuristic creates a flexibility schedule with a non-normalized fitness value of -0.233, whereas the best solution found by the *GA* evaluates to -1.613, gradually improving from values around -4 of the random initial population. On the other hand, when initializing the *GA* with known good solutions coming from the linear heuristic, the best chromosome improves after more than 320 generations to a slightly better value of -0.206. This can be explained by the mutation operator which shifts continuous flexibility between *EMSs*, while the linear heuristic only allocates multiples of S^{max}. The two main observations are: First, the linear heuristic can find a nearly optimal solution and, second, the proposed meta heuristics cannot compete with the specialized linear heuristic in terms of performance and efficiency when problem-specific knowledge can be utilized. However, the *GA* has its main strength to solve multiple objectives and can generate the Pareto-set of the problem.



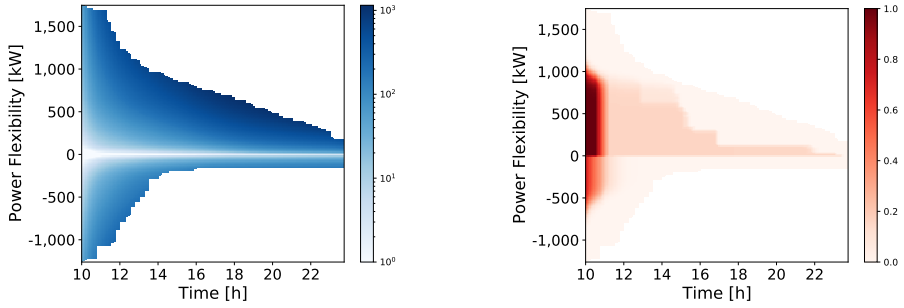
(a) EQUAL.



(b) PROP.



(c) COST.



(d) POPT.

Figure 4.13: Aggregated cost (left, blue) and probability (right, red) of energy flexibility of a pool consisting of 150 *EMS* with $S^{max} = 1$kW [37]. Note that cost graphs use log scale in the color map.

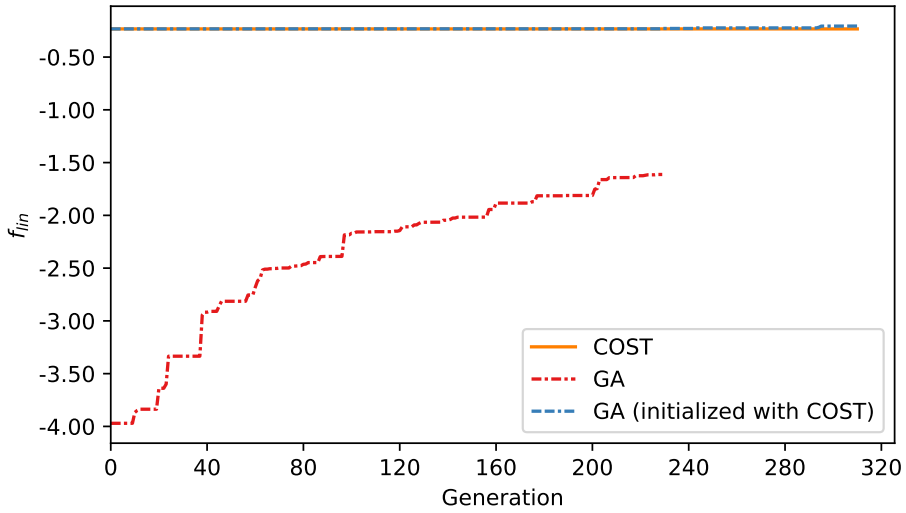


Figure 4.14: Best cost-only fitness compared to the problem-specific linear heuristic.

The resulting disaggregation schedule of the linear heuristic can be mapped to the chromosome representation on which the fitness values can be evaluated. Figure 4.15 shows the Pareto-front of a 12 h execution of the *GA* together with the results from the four disaggregation policies of the linear heuristic. It can be seen that the *GA* tends to optimize towards the objectives of fitness functions f_2 and f_4, which outperform the linear heuristic. On the other hand, some linear heuristic policies can find better solutions concerning f_1 and f_3, which both are directly or indirectly part of their priority functions.

Figure 4.16 summarizes the absolute scheduled energy flexibility for each of the four disaggregation policies of the linear heuristic and the chromosomes from the Pareto-front with the best linearized and the best fitness value per dimension. As can be seen, the absolute scheduled energy flexibility differs quite a lot. On the one hand, the rebound effect of flexibility requires additional flexibility to be scheduled to the flexibility pool as explained in Section 4.4.1. On the other hand, additional flexibility may come with no cost or can be utilized to improve the optimization objective, which for example is the case for optimizing power grid losses by fitness dimension f_2. It is worth mentioning that only the best chromosome for fitness dimension f_3 schedules solely negative flexibility requests to the *EMSs* with no additional flexibility requirement. This fitness function maximizes the degree of self-consumption and autarky, which both are reduced by a flexibility rebound that must be compensated by other *EMSs*.

4.5.2.4 Quality of Service/Experience and Fairness

The obtained metric values for the quality of service and experience of the different disaggregation policies and the *GA* are analyzed in this section. Figure 4.17 shows box plots over 55 *EMSs*, where the mean metric values are indicated by black circles. Below the box plots, the fairness index F is depicted for each metric and policy, where $L = -3090.91$ and

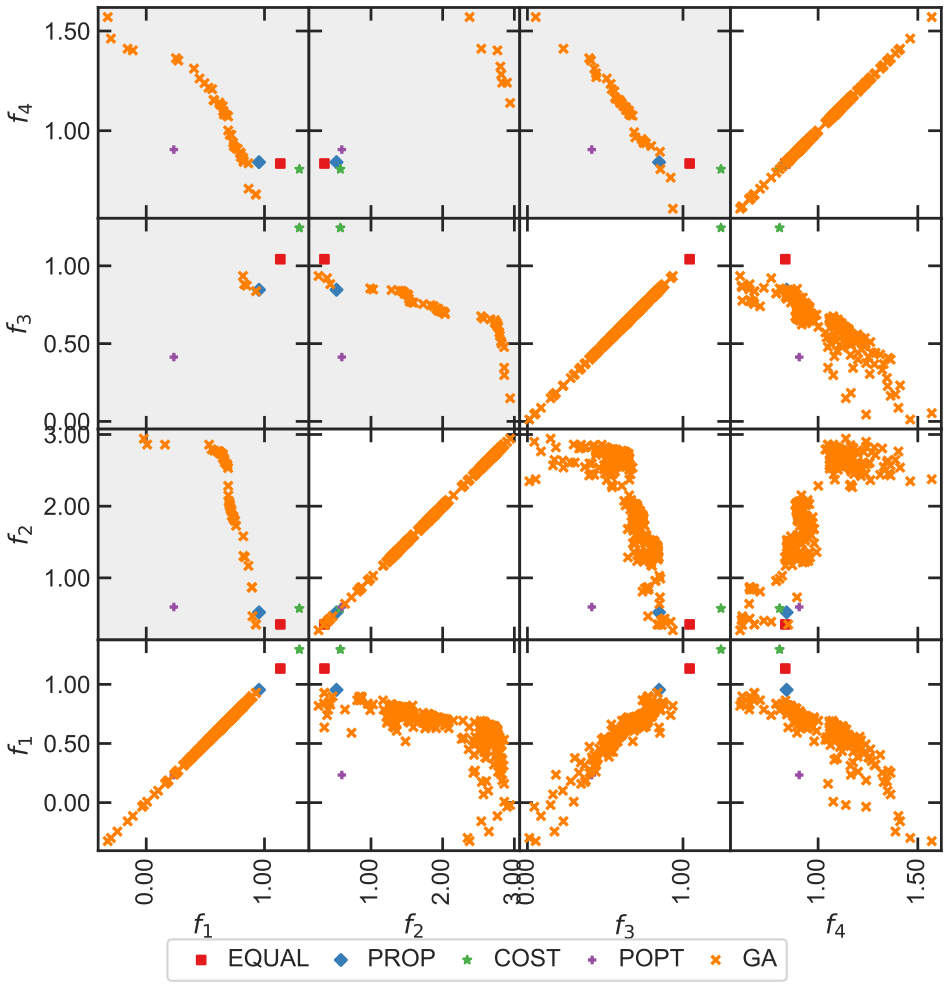


Figure 4.15: Pareto-set of the *GA* of a 12 h execution compared with the results from the linear heuristic.

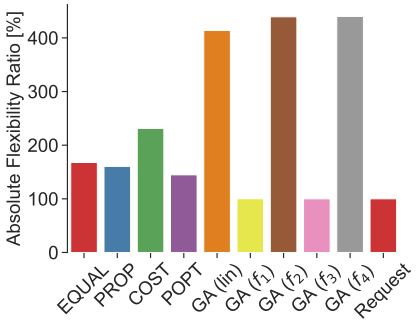


Figure 4.16: Absolute schedule flexibility to the flexibility pool, which differs between the disaggregation policies and optimization objectives.

$H = 0$ for QoS_6. Note that since QoS_7 and QoS_8 yield only one value for the flexibility pool, no fairness index can be computed.

The QoS_5 metric measures the relative participation of an *EMS* to the flexibility request. By definition, the mean QoS_5 value among the method must be equal, since the number of participating *EMSs* is constant. However, the individual participation ratio slightly differs among the *EMSs*, where the policy EQUAL achieves the fairest participation rates by design. Since COST iteratively schedules flexibility to the *EMS* with the lowest cost, many *EMS* receive at least a minimum, often nearly equal flexibility participation share as well. All the other disaggregation schedules suffer from outliers where up to 30 % of the flexibility request is delivered by one single *EMS*.

The metric QoS_6 measures the deviation from the initial optimal schedule without scheduling a flexibility request. This is best achieved by the COST policy and the *GA* with f_1 and f_3 since all three minimize directly or indirectly the flexibility cost, which is defined based on the highest priority objective function O_1 in Equation 4.29. Again, the EQUAL policy performs sufficiently well since all *EMSs* participate in the delivery of the flexibility request with a small share, which typically comes with only a small deviation from the optimal grid profile. Visually, most quality metrics of the *EMSs* with *GA* (f_3) are very close to each other, however many outliers with drastic changes of lower priority objectives cause the fairness index to drop.

There is only a single value for the two metrics QoS_7 and QoS_8, which measure the quality of the flexibility resource aggregator. As can be expected, COST yields the smallest flexibility cost and therefore receives the best QoS_7 metric score, followed by EQUAL and PROP which distribute the flexibility request to all *EMSs*. This results in a usually small total cost since a small flexibility request can be provided by any *EMS* with usually very small or no cost at all. Similar to Figure 4.14, the *GA* with f_1 receives a slightly lower quality score than COST but outperforms the chromosomes that are chosen for the other fitness dimensions. Only *GA* (f_3) benefits from scheduling the minimum amount of flexibility and reaches a comparable metric value.

As can be expected, POPT achieves the best quality score for the mean delivery probability measured by QoS_8. The PROP policy reaches a high metric score as well, since the flexibility is scheduled proportional to the flexibility limits of the *EMSs*, and the error probabilities of the different *EMSs* are similar to each other. From the chromosomes of the *GA*, again fitness dimensions f_1 and f_3 perform better since f_1 aims to maximize the delivery probability together with the flexibility cost and f_3 schedules no additional flexibility, which keeps a larger safety margin towards the flexibility limits of all *EMSs*.

With QoE_4, the ratio of flexibility participation is measured against the flexibility limits of the *EMSs*. As can be expected, the flexibility schedules that utilize additional flexibility achieve a larger flexibility utilization at the *EMSs*. This is the case for *GA* (lin), *GA* (f_2), and *GA* (f_4). On the other side, *GA* (f_1), *GA* (f_3), and POPT, with relatively small or no additional flexibility yield much lower utilization values, but reach better fairness indices.

The resulting degree of self-consumption and autarky, measured by QoE_5 and QoE_6, do not differ much between the disaggregation policies. When comparing the box plots of self-consumption (QoE_5) with Figure 4.6, it can be seen that the mean value drops by

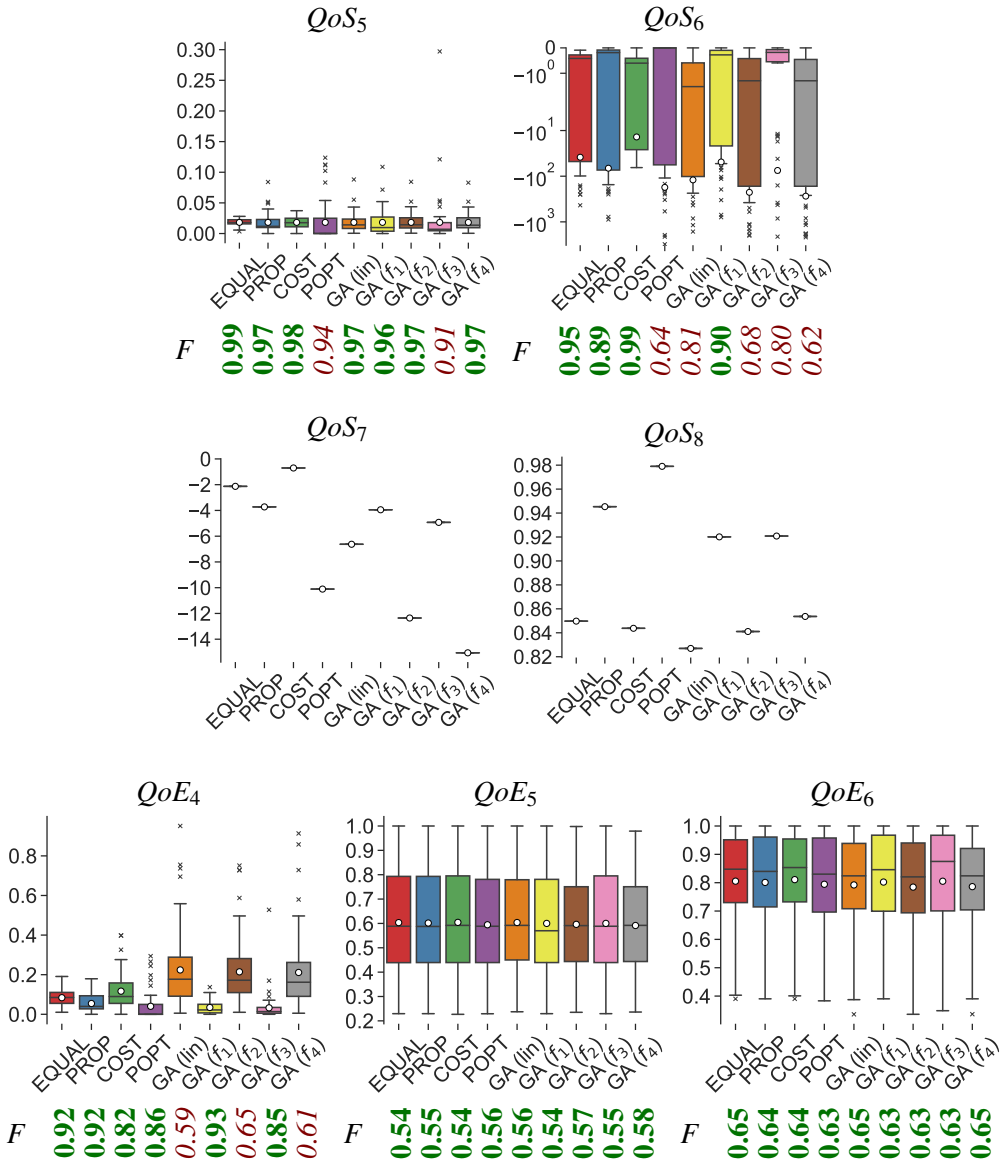


Figure 4.17: *QoS* and *QoE* metrics of the different disaggregation policies for 55 *EMSs*. The box plots show the resulting distribution among the *EMSs* and the circle denotes the mean value. Below the box plots, the achieved fairness index F is given. Note that for QoS_6, the box blot uses a logarithmic scale to better capture the outliers.

Table 4.5: Impact of flexibility disaggregation on the power grid.

Disaggregation mechanism	Cost (QoS_7)	Mean probability (QoS_8) [%]	Min [V]	Max [V]	Grid Losses [%]
EQUAL	**–2.13**	*84.97*	**229.95**	**234.66**	*4.28*
PROP	**–3.73**	**94.53**	**229.95**	**234.72**	*4.06*
COST	**–0.70**	*84.38*	**229.94**	**234.81**	*3.99*
POPT	*–10.11*	**97.90**	**229.96**	**234.69**	*3.96*
GA (lin)	*–6.63*	*82.69*	**229.89**	*235.13*	**1.16**
GA (f_1)	**–3.95**	**92.01**	**229.96**	**234.65**	*4.28*
GA (f_2)	*–12.35*	*84.11*	**229.88**	*235.13*	**0.89**
GA (f_3)	*–4.93*	**92.08**	**229.96**	**234.60**	*4.39*
GA (f_4)	*–15.06*	*85.36*	**229.90**	*235.12*	**1.63**
Baseline	-	-	229.90	234.99	0.00

approximately 9 % points compared to the 350 *EMSs* without scheduled flexibility. Nevertheless, the statistical range of the values is nearly similar. The degree of autarky (QoE_6) does not show any significant differences as well. It can be concluded that, first, scheduling flexibility in the investigated use case scenario does only slightly impact the degree of self-consumption and autarky and, second, that all different disaggregation schedules yield a very similar metric score distribution among the *EMSs* in the investigated use case scenario.

From the different disaggregation policies and *GA* fitness dimensions, PROP is among the best disaggregation schedules with both a high mean metric score and fairness index in most of the quality of service and experience metrics on the given use case scenario.

4.5.2.5 Impact on the Low-Voltage Power Grid

The *DSO* cares about the voltage levels in the power grid and the overall grid losses of the flexibility request, which the *GA* already considers by minimizing the grid losses in fitness dimension f_2 and by dropping chromosomes that produce invalid voltage values. The nine different schedules from the analysis above are applied to the 55 *EMSs* located in the simulated low-voltage power grid, where the resulting grid losses and voltage levels are listed in Table 4.5. The minimum voltage level at any bus in the low-voltage power grid is in a reasonable range and only differs by less than 0.06 V, which can be neglected. From the given *PV* peak shaving scenario, it can be expected that the maximum voltage level is reduced since the *PV* injection at the distributed *EMSs* is lowered after flexibility scheduling. Thus, most voltage values improve by around 0.2 V to 0.4 V, however, this is not the case for all disaggregation schedules. The three *GA* schedules *GA* (lin), *GA* (f_2), and *GA* (f_4) shift a high amount of additional flexibility between the *EMSs*, which even increases the maximum measured voltage level by approximately 0.12 V. This drastically impacts the grid losses of the flexibility request that drop from around 4 % of the scheduled flexibility

to values below 1.6 %, where the best value of 0.89 % is achieved by *GA* (f_2). Consequently, different flexibility schedules do not have a major impact on the voltage levels, but grid losses can be reduced by shifting flexibility between the *EMSs*. This however comes with additional flexibility cost (QoS_7) and finally leads to a trade-off between avoiding grid losses and flexibility cost.

4.5.3 Sensitivity of Use Case Scenarios

The results discussed so far are based on the use case scenario defined in Section 4.5.1. This section investigates the impact of flexibility disaggregation on three further use case scenarios as described in the following.

- First, the same peak shaving is applied during *PV* surplus but with the capacity threshold set to −180 kW. This results in a much smaller flexibility request between 11:00 and 14:00 as given in Equation (4.68) which sums up to only 22.25 kWh.

$$P_{flex}^{Pool} = (-2\,835, -17\,392, -2\,022) \tag{4.68}$$

  This very small flexibility request yields a large solution space for different disaggregation schedules. Additional flexibility is mainly utilized for the optimization of grid losses and fairness by the fitness dimensions f_2 and f_4. This ultimately results in negative grid losses of −31.4 %[23], which however comes with a high cost in terms of objective O_1 at the *EMSs*. On the contrary, the cost metric QoS_7 is very close to zero for all other scheduling policies, where COST incurs nearly no cost at all. Similarly, *GA* (f_4) schedules additional flexibility to improve the fairness index on achieved self-consumption (QoE_5) and autarky (QoE_6) by 0.03 points with a negligible reduction in both metric values. Unlike all the other schedules, POPT distributes the flexibility request to only 8 out of 55 *EMSs*, each of which can provide flexibility with high delivery probability. However, due to the small flexibility volume, PROP and *GA* (f_1) achieve similar high delivery probabilities. The remaining flexibility service metrics look very similar to the use case scenario discussed in the last section. Related graphs are shown in Figure A.5.

- In a completely different scenario, positive flexibility is scheduled to counteract the expected evening peak load of residential households. The corresponding *SLP* has an evening peak between 19:00 and 22:00 as depicted by the red area in Figure 4.18. The required peak shaving is highlighted by the red dotted area and can be accomplished by a flexibility request similar to Equation (4.69), which sums up to 230 kWh.

$$P_{flex}^{Pool} = (80\,000, 100\,000, 50\,000) \tag{4.69}$$

  Since this positive flexibility request is scheduled in the evening, the *ESSs* are almost fully charged by the *PV* surplus and, hence, the flexibility is provided by discharging

[23] A negative grid loss ratio means that, compared to the scheduled flexibility request, a larger flexibility reaction can be expected at the transformer.

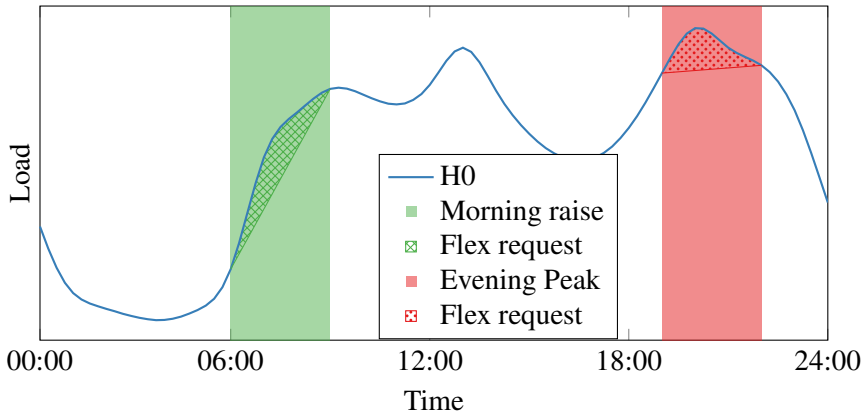


Figure 4.18: *SLP* H0 for German households on a winter working day provided by the *German Association of Energy and Water Industries* (*BDEW*). This plot intentionally does not have y-axis ticks since it only visualizes the use case scenario from which numerical flexibility requests are derived.

the *ESSs* and additionally shifting flexible *EV* charging processes towards the night hours. Thereby, only *GA* (f_2) utilizes additional flexibility to improve the grid losses from around −0.5 % to −1.14 %, which corresponds to less than 1 kWh. The effect is much lower compared to the above scenario since the transformer that connects the resource providers is only slightly loaded during that time according to Figure 4.4. The total cost for the evening peak shaving flexibility request, measured by QoS_7, is comparably high because most flexibility is provided by discharging the *ESSs* which results in a 12 % lower mean battery *SoC* at the end of the day. Thereby, COST outperforms the other policies by a factor of 1.60 to 3.27. Since the grid profiles of the *EMSs* are close to zero during the flexibility request, the mean delivery probability is high for all scheduling policies, where QoS_8 yields values of 0.975 and above. Again, self-consumption and autarky are not influenced much by the different disaggregation schedules of the flexibility request. Related graphs are shown in Figure A.6.

- Finally, the residential load behavior tends to increase very rapidly in the morning as highlighted by the green area in Figure 4.18. A flexibility request similar to the marked green area can be scheduled to smooth the load profile between 6:00 and 9:00 as given in Equation (4.70). This positive flexibility request sums up to 200 kWh.

$$P_{flex}^{Pool} = (50\,000, 100\,000, 50\,000) \tag{4.70}$$

Since not many *EV* charging processes are pending in the morning, the *ESSs* are precharged from the grid during the night and will reach 0 % *SoC* after delivering the flexibility request. Due to this limited flexibility potential, consequently, not much additional flexibility is scheduled by the different policies, especially by *GA* (f_2). EQUAL, PROP, and COST reach nearly even proportional flexibility participation among the *EMSs* (QoS_5), while the other schedules vary to a greater extent, however,

no large outliers are recorded. Similar to the use case of evening peak shaving, the positive flexibility comes with a higher total flexibility cost (QoS_7) than the negative flexibility of *PV* generation shaving. However, the differences between the scheduling policies are much smaller due to the limited flexibility provided by *EV* charging processes. The mean delivery probability measured by QoS_8 is similarly high as with the evening peak shaving with values above 0.975, except for *GA* (f_2) and *GA* (f_4) which yield reduced metric scores of 0.89 and 0.94. Finally, the number of *EMSs* with very high self-consumption (QoE_5) is reduced. This consequently leads to some *EMSs* with very low autarky, especially with *GA* (f_2). Related graphs are shown in Figure A.7.

From these three additional scenarios, it follows that *GA* (f_2) and *GA* (f_4), which aim for minimum grid losses and maximum fairness index, highly contradict to flexibility cost and delivery probability. Consequently, there needs to be a trade-off between the objectives of the resource aggregator and the *DSO*. This is possible with the *GA* which allows the selection of a suitable chromosome from the Pareto-set posteriorly.

4.6 Applicability

This section deals with the practical implementation of flexibility disaggregation to distributed *EMSs* using the linear heuristic from Section 4.4.1 and the *GA* from Section 4.4.2. Thereby, possible real-life application scopes are introduced. Moreover, technical and legal challenges, as well as limitations are discussed. Finally, an extension towards *ASs* in the form of reserve power disaggregation is presented.

4.6.1 Scope of Application

The flexibility pools allow aggregating small flexibility capabilities from distributed *EMSs* to market-sized flexibility potential, while the disaggregation of flexibility requests to the flexibility pool can reach a certain quality of service and fairness for the resource providers. This offers the potential for different applications on the energy market or direct flexibility activation by the *DSOs*, where three ideas are discussed in the following.

- By scheduling flexibility to distributed *EMSs*, a *Balance Responsible Party* (*BRP*) can **optimize its balancing group**. Thereby, flexibility is scheduled to compensate for supply and demand mismatch during the planning horizon, *e. g.*, a day ahead. Unpredictable rebound effects that can happen before and after the flexibility request must be considered by the *BRP*, however, the rebound will not create critical new peaks due to objective O_3 of the *EMSs*. Furthermore, if a direct shift of energy from one time slot to another is desired, multiple opposite flexibility requests can be placed at once which most likely results in the desired energy shift, as given by the example in Figure 4.19. Not only the ability to request flexibility helps the *BRP* to better plan its balancing group, but also the feedback from the *EMSs* with its optimal operational grid profiles is more accurate than using *SLPs*.

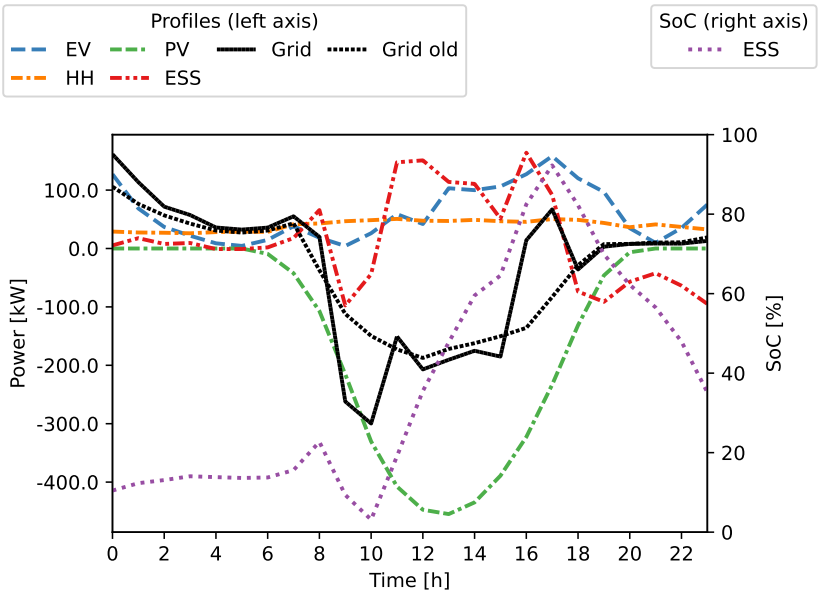


Figure 4.19: Energy shift of 300 kWh from 9:00/10:00 to 16:00/17:00 using two opposite flexibility requests.

- The resource aggregator can **dynamically assign flexibility** to the flexibility pool during the day. Since the linear heuristic schedules the flexibility to one time slot after another, it is most suitable for dynamic intraday flexibility disaggregation. This allows the resource aggregator to interact with the intraday markets and to control flexibility delivery in hourly or quarter-hourly blocks on the fly. Using the *receding horizon* approach for the *MILP* of the *EMSs*, it is further possible to keep a future flexibility horizon of for example 24 h and also possible to integrate the actual uncertainty realization into the control loop.

- The *DSO* may request flexibility from distributed *EMSs* to **shave load and generation peaks, smooth fast grid utilization changes, or solve grid congestion** in the power distribution grid. These possible application scenarios are similar to the use case scenarios from Sections 4.5.1 and 4.5.3. From the power grid perspective, only fast changes and critical peak loads are relevant, hence the rebound effect of the flexibility provision can be neglected as long as it does not cause further grid issues. Grid congestion relates to the overloading of transformers or certain power lines, which can be solved by scheduling opposite flexibility to the connected regions similar to nowadays applied re-dispatch actions.. Request from the power grid operator can either be issued via a local flexibility market or as a direct control signal if the *DSO* faces congestion problems. Especially in this application scenario, it is important to provide flexibility with minimum or at least reduced grid losses because grid losses put additional stress on the power distribution grid.

4.6.2 Technical Challenges

The main technical limitation for applying the proposed methods in the field is the lack of computational power for *MILP* at the distributed *EMSs*. Most systems run on low-performance hardware such as the single-board computer Raspberry Pi, which is based on a 4-core ARM chip. Compared to the x86 instruction set, which most desktop computers and servers use, there are not many *MILP* solvers that efficiently run on ARM hardware. However, the ARM architecture is getting more popular and since 2020 one of the top-ranked supercomputers *Fugaku* uses ARM chips[24]. Meanwhile, heavy optimization runs of the *MILP* problem can be performed in cloud computing environments since there is anyhow no need for real-time control at the local *EMS*. Alternatively, a relaxation by lifting the constraints for minimum charging current for *EV* charging will turn the *MILP* problem into a much simpler linear problem. To avoid infeasible solutions, the *ESS* parameters may be restricted to always be able to provide compensation flexibility for the relaxation, which however will reduce the flexibility potential of the *EMS*.

Many manufacturers in the domain of *PV* inverters, *ESSs*, and home automation offer their own implementation of a local home *EMS* with a variety of local energy optimization strategies. In addition to that, several software-based open source alternatives exist, such as OpenEMS.[25] One open technical challenge is to connect a heterogeneous set of *EMSs* to a flexibility pool. This is only possible with a well-defined *Application Programming Interface* (*API*), which in the context of this thesis requires a set of methods. These methods include *(i)* retrieval of the optimal grid profile for the planning horizon, *(ii)* retrieval of the maximum and minimum possible flexibility for a single time slot, *(iii)* scheduling of a flexibility request, and *(iv)* calculating the flexibility cost, delivery probability, self-consumption, and autarky. Any *EMS* implementation that offers these functionalities can be grouped into a flexibility pool as described in this thesis. In the context of the research project *EMSIG* – funded by the *Bavarian Ministry of Economic Affairs, Regional Development, and Energy*, and by the *Zentrum Digitalisierung.Bayern* –, a simplified version of the *MILP* problem is implemented within the OpenEMS platform and is available under open source license.[26]

Since both the linear heuristic and the *GA* require a lot of interaction with the *MILP* problem of the distributed *EMS*, it is most reasonable to encapsulate the *MILP* problem (or a virtual copy) into a black box and provide this black box to the disaggregation service of the flexibility pool. In this way, the details of the *EMSs*, *e. g.*, the number of *EVs*, are kept private from the flexibility aggregator.

4.6.3 Legal and Energy Market Framework

Resource aggregators, in the role of a *BRP*, can trade energy on the energy market. In Central Europe, the day-head and intraday markets are organized by the *EPEX Spot* market

[24] Top 500 list November 2022 (https://www.top500.org/lists/top500/2022/11/)

[25] https://openems.io/

[26] The proof of concept was implemented by the project partner FENECON and is available as a pull request https://github.com/OpenEMS/openems/pull/1466 (visited on Feb. 3, 2023)

in Paris, where energy products of different time scales, *e. g.*, 15 minutes or one hour, are traded with a minimum capacity of 0.1 MWh. Consequently, to be able to place reasonable orders, a flexibility pool must contain a minimum of 100 *EMSs*, each of which provides 1 kWh. Especially the possibility to place *loop blocks* to the day-ahead market – a combination of two blocks that are either executed or rejected together – allows for offering flexibility with rebound effects as demonstrated in Figure 4.19.

In the case of grid-supportive flexibility for the *DSO*, there are two possible frameworks. First, similar to the *EV* charging-as-a-service from Section 3.6.3, the *Energy Industry Law* (§ 14a EnWG) can be applied in Germany to reduce the grid usage fees for flexible appliances. In turn, it allows the *DSO* to request flexibility. In this way, reduced grid usage fees are granted all over the time even if no flexibility is required by the *DSO*. A second option is the newly introduced *EPEX localflex market* [50] or similar local flexibility markets, which connect flexibility providers with *Transmission System Operators* (*TSOs*) and *DSOs* via a market platform to trade long-term and short-term flexibility.

4.6.4 Possible Extension of Reserve Power Flexibility

The *MILP* formulation of the energy optimization problem of an *EMS* can also schedule *ASs* in the form of *aFRR* under Assumption A17. These power reserves are usually traded on larger time blocks of four hours and require specific hardware controllers in place, *e. g.*, a droop controller that reacts on under- and over-frequency events. Therefore, in this *EMS* model, only the stationary *ESS* is considered for reserve power, because it is always available and may provide the required control functionality.

Reserve Power Request In contrast to power flexibility, scheduling reserve power does not directly influence the grid profile at the *EMS*, but only adds additional constraints to the decision variables, such that the reserve power can be provided if activated. The upper and lower reserve power bounds can be determined by maximizing and minimizing the *ESS* usage as in Equations (4.71) and (4.72). The possible reserve power is the difference between the optimal *ESS* profile $P_{ESS}(t)$ and its limits.

$$P_{ESS}^{max}(t) = \max_{M_{EV,i}(t),P_{EV,i}(t),M_{ESS}(t),P_{ESS}^{+}(t),P_{ESS}^{-}(t)} P_{ESS}(t) \tag{4.71}$$

$$P_{ESS}^{min}(t) = \min_{M_{EV,i}(t),P_{EV,i}(t),M_{ESS}(t),P_{ESS}^{+}(t),P_{ESS}^{-}(t)} P_{ESS}(t) \tag{4.72}$$

To provide reserve power of $P_{res}(t)$, the utilization of the *ESS* in terms of power capacity $P_{ESS}(t) = P_{ESS}^{+}(t) - P_{ESS}^{-}(t)$ must be restricted to have a safety margin of $P_{res}(t)$ towards its maximum or minimum possible value depending on the direction of the requested reserve power. This safety margin for reserve power capacity is added to the *ESS* by the constraint in Equation (4.73). Besides the *ESS*, also the grid profile is similarly restricted by the constraint in Equation (4.74) to be able to provide reserve power in case of activation.

$$-P_{ESS}^{max} + P_{res}(t) \leq P_{ESS}(t) \leq P_{ESS}^{max} + P_{res}(t) \tag{4.73}$$

$$P_G^{min}(t) + P_{res}(t) \leq P_G(t) \leq P_G^{max}(t) + P_{res}(t) \tag{4.74}$$

Finally, the battery *SoC* is restricted such that the energy required for the reserve power can be buffered or provided for the requested duration. In general, *aFRR* is traded in four-hour blocks in Europe, where the respective reserve power must be provided for the full four hours. Since multiple *EMSs* participate in a *VPP*, a single *EMS* may only need to deliver reserve power in a shorter duration, and other *EMSs* of the *VPP* can compensate for the remaining power during the four-hour block. However, to avoid rebound effects during the same *aFRR* window, each *EMS* must reserve energy capacity at the *ESS* for the remaining four-hour block, such that even in case of *aFRR* activation, the initially planned *ESS* schedule can be performed. That will delay potential rebound effects after reserve power activation to blocks, where no *aFRR* is scheduled, which allows the *VPP* operator to compensate for the required energy rebound by its local generation sources.[27]

For each four-hour block, where positive or negative reserve power is scheduled, the following constraints are added to restrict the usable battery *SoC*. First, the required energy for each time slot of the *aFRR* block is calculated for positive and negative reserve power in Equations (4.75) and (4.77). These two equations simply sum up the required energy within every four-hour block if reserve power is scheduled at any time slot t'. In both cases, the initial energy reserve is equal to zero $E_{res}^{+}(t-1) = E_{res}^{-}(t-1) = 0$. Consecutive reserve blocks use the energy reserve from the former block as the starting point, because there is no chance to compensate for the rebound effect in between.

$$E_{res}^{+}(t) = \begin{cases} E_{res}^{+}(t-1) + \min(0, P_{res}(t)) \cdot \Delta_t \cdot \frac{1}{\mu_{ESS}} & \text{if cond. (4.76)} \\ 0 & \text{else} \end{cases} \tag{4.75}$$

$$\exists t' \in \left\{ \left\lfloor \frac{t \cdot \Delta_t}{4} \right\rfloor + i \colon 0 \leq i < \frac{4}{\Delta_t} \right\} : P_{res}(t') > 0 \tag{4.76}$$

$$E_{res}^{-}(t) = \begin{cases} E_{res}^{-}(t-1) + \max(0, P_{res}(t)) \cdot \Delta_t \cdot \mu_{ESS} & \text{if cond. (4.78)} \\ 0 & \text{else} \end{cases} \tag{4.77}$$

$$\exists t' \in \left\{ \left\lfloor \frac{t \cdot \Delta_t}{4} \right\rfloor + i \colon 0 \leq i < \frac{4}{\Delta_t} \right\} : P_{res}(t') < 0 \tag{4.78}$$

Second, the storage capacity in the *ESS* is limited by the constraint in Equation (4.79).

$$E_{res}^{+}(t) \leq E_{ESS}(t) \leq E_{ESS}^{max} - E_{res}^{-}(t) \tag{4.79}$$

Note that the energy calculation in Equations (4.75) and (4.77) are an approximation of the exact power reserve scheduling problem because they neglect the fact that reserve power that is provided by reduced (dis-)charging power does not yield charging losses that are modeled by the factor μ_{ESS}. Hence, the *SoC* of the *ESS* is slightly more restricted.

[27] The *VPP* may trade a baseload of the aggregated profiles on the day-ahead market and keep some oversupply capacity for compensating the rebound effect, similar to the approach proposed in [117]. If the *aFRR* is not activated, this capacity can be shifted by power flexibility requests to the *EMSs*.

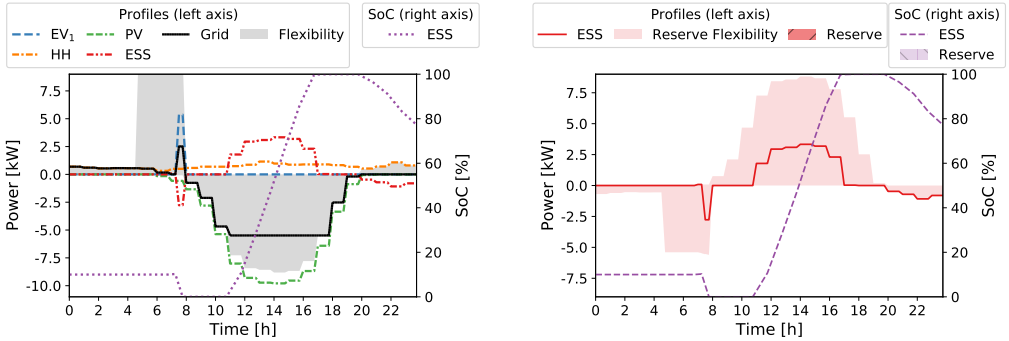
The following example is based on the same *EMS* configuration as in Figure 4.1 with *EMS* constraints from Equations (4.14) and (4.15) enabled. Figure 4.20(a) shows the optimal profile for the *EMS* in the left diagram and the *ESS* flexibility before the scheduling of reserve power in the right diagram. In the morning, only negative reserve power is possible during *EV* availability. At times with *PV* oversupply, the optimal *ESS* operation is to charge the battery with medium charging power, which offers the possibility for both negative and positive reserve power. In Figure 4.20(b), a reserve power of −6 kW is scheduled between 12:00 and 13:00. First, the remaining *ESS* charging power flexibility in the right diagram is limited during that time. Furthermore, the battery *SoC* is restricted to keep the battery below approximately 60 % during the reserve power block to be able to buffer the reserve power for at least four hours. This causes the optimal schedule (left diagram) to alter from the original optimal one, and the *ESS* charging is partially shifted outside the four-hour block. In the case of *aFRR* activation, the *ESS* provides the scheduled reserve power and the optimal *EMS* profile alters again in Figure 4.20(c). Note that requirement of activation is first known at 12:00, hence the *EMS* sticks to the planned optimal profile before. Because the rebound effect of the *ESS* can only happen outside the four-hour block, the optimal profile between 13:00 and 16:00 stays the same as well.

Reserve Power Cost and Uncertainty Compared to power flexibility, the reserve power generates two cost types, once at the time of scheduling the reserve power and once for its activation. This fits the reserve power market, which has a capacity price and an energy price, as explained in Section 2.1.2. First, the additional constraints on the *ESS* for power and energy reservation can cause sub-optimal operational profiles for the *EMS* at the time of reserve power scheduling. This results in the direct costs for the power reservation $\gamma_{res}^{sched}(x)$, which are calculated similarly to power flexibility scheduling in Equation (4.29). Second, reserve power activation may result in sub-optimal battery charging or discharging schedules, where the cost $\gamma_{res}^{act}(x)$ can be estimated by comparing the reserved scenario (O_1^{sched}) with the activated scenario (O_1^{sched}). The total cost for reserve power is given in Equation (4.80), where $\upsilon \in [0,1]$ is the expected activation probability.

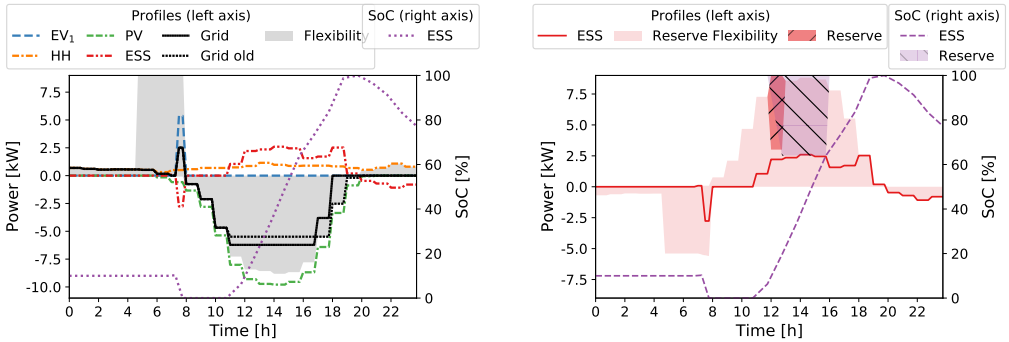
$$\gamma_{res}(x) = \gamma_{res}^{sched}(x) + \upsilon \cdot \gamma_{res}^{act}(x) \tag{4.80}$$

Reserve power delivery from the *ESS* is guaranteed by the added constraints during reserve power scheduling. However, the uncertainty arising from load and *PV* forecast can affect the grid profile and, hence, cause a violation of the grid limitation constraint. In the case of reserve power activation, a higher safety margin between the activated grid profile and the grid limitations can help to tolerate larger forecast errors. Consequently, the probability for reserve power delivery is calculated the same as with power flexibility requests in Equation (4.33), where $x = P_{res}(t)$.

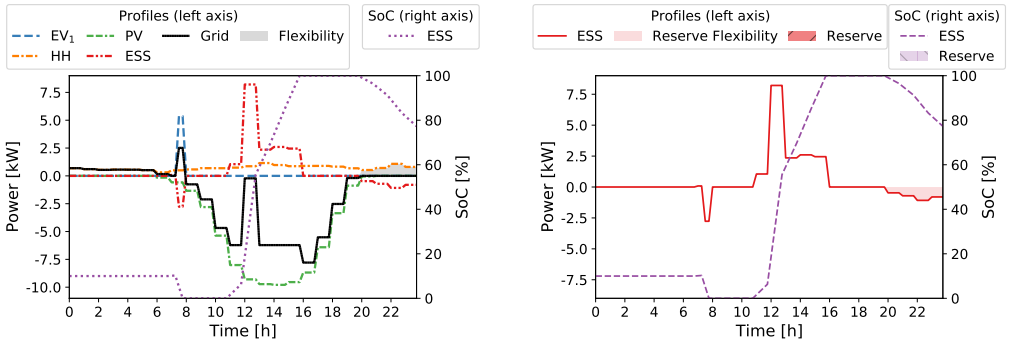
Reserve power aggregation and disaggregation are modeled similarly to power flexibility, including the flexibility request $(P_{res}(t))_{t\in\mathcal{T}}$, the minimum and maximum reserve power limits $P_{res}^{min,Pool}(t)$ and $P_{res}^{max,Pool}(t)$, and the disaggregation of a reserve power as a matrix $(\llbracket P_{res}(t) \rrbracket_e)_{e\in\{1,\dots,K\},t\in\mathcal{T}}$. Finally, the two discussed heuristics for flexibility provision in Section 4.4 can also be applied to reserve power requests.



(a) Optimal profile before scheduling reserve power.



(b) Optimal profile after scheduling −6 kW power reserve between 12:00 and 13:00.



(c) Optimal profile after activation of the scheduled power reserve.

Figure 4.20: The same exemplary *EMS* from Figure 4.1 with available 4.20(a), scheduled 4.20(b), and activated 4.20(c) reserve power. The left diagram of each sub-figure shows the optimal profiles of the *EMS* and its flexibility. The right diagram shows the *ESS* optimal profile and its reserve power constraints.

4.7 Chapter Summary

This chapter describes a *MILP* model for local home *EMS* that can utilize flexible appliances for local optimization and provision of external flexibility requests. Thereby, the main asset is the stationary *ESS*, from which the expectations of the resource provider are reflected by a set of *QoS* and *QoE* metrics. Multiple of these *EMSs* are grouped into a flexibility pool to reach market-sized flexibility potential, which is disaggregated to the distributed *EMSs*.

Two flexibility disaggregation approaches to calculate the disaggregation vector of a flexibility request to the flexibility pool are discussed. The first applies a linear heuristic that iterates over time and splits the flexibility request into equally sized portions, which are iteratively assigned to the best suitable *EMS*. Discussed disaggregation policies consider among others cost-minimization and delivery probability maximization. The second approach implements a *GA* that assigns multiple time slots at once while maximizing a multi-dimensional fitness function that encodes the objectives of the resource aggregator, the resource provider, and the power grid operator. Starting from a set of random solutions, improved populations are created using a crossover and a mutation operator, which by design are cache-optimized to speed up the execution of the algorithm. From the resulting Pareto-optimal solution set, one suitable flexibility disaggregation is chosen posteriorly considering the trade-off between the fitness dimensions.

These flexibility disaggregation methods are analyzed in depth using one realistic flexibility request scenario based on the IEEE 906 low-voltage test feeder with regard to *QoS*, fairness, and power quality aspects. Three additional flexibility request scenarios underpin the findings. Finally, this chapter discusses different application scopes on the energy market and gives an outlook for extending the flexibility allocation towards reserve power. The main findings of this chapter are as follows.

- One of the simplest disaggregation policies PROP is among the best disaggregation schedules with both a high mean metric score and fairness index in most of the *QoS* and *QoE* metrics. Nevertheless, when specific flexibility delivery requirements must be met, more advanced disaggregation policies should be applied, *e. g.*, cost or delivery probability-optimal.
- The experiments of all use case scenarios show that there is a trade-off between grid aspects (in terms of grid losses and voltage levels) and service quality metrics (cost, probability, and participation ratio).
- Finally, the degree of self-consumption and energy autarky is not affected much by delivering flexibility throughout the day nor major differences between the applied disaggregation policies can be expected.

CHAPTER 5

Conclusions and Future Work

This chapter concludes the thesis and highlights the main contributions and results in Section 5.1. Furthermore, the limitations of the thesis are discussed in Section 5.2 together with an outlook on how to deal with them.

5.1 Main Contributions and Results

The restructuring of the power distribution grid from central electricity supply to distributed generation, which is necessary for a successful energy transition, will undeniably cause power distribution grids to reach their capacity limits at peak load times. Since required grid expansion is not always possible in a reasonable time, intelligent control of *EVs* and integration of flexibility from distributed *EMSs* seem to be promising solutions to avoid asset overloading and power quality issues. However, when interfering with assets owned and operated by users, one needs to consider the customers' requirements and objectives during smart control. This is comparable with Internet-based services which have service-based requirements for optimal resource allocation, *e. g.*, downloading a file from a server has different requirements compared to video conferences.

The two smart grid applications under investigation are *charging-as-a-service* for distributed *EV* home charging and *flexibility-provision-as-a-service* to distributed home *EMS*. These two are chosen since both impair high loads to the power distribution system, which are a multiple of the usual peak load, and offer decent flexibility for reasonable operational control. The focus of this thesis is to identify technical *QoS* and *QoE* aspects that measure the expected and received service quality by the users, which is only partially discussed in the literature. Furthermore, the thesis proposes different scalable and responsive resource allocation mechanisms for both domains that offer efficient power grid utilization, conform with legacy control capabilities, and consider quality of service and fairness aspects.

In the domain of *EV charging-as-a-service*, a standardized charging service model is introduced in Chapter 3, from which a comprehensive set of *QoS* and *QoE* metrics is derived. These metrics measure among others the ratio of charged energy to the requested energy, assess the continuity of charging power rate, reflect the battery *SoC* at departure, and consider the ability to reach the next destination. In addition to the quality metrics, fairness is discussed as the deviation of received service quality metrics among the distributed charging services using a fairness index that is decoupled from the metric scores.

The efficient and scalable online resource allocation to distributed *EV* charging processes is performed via a charging packet queuing protocol that takes charging process parameters for prioritization and offers not only temporal charging slot allocation but also variable charging rates, complying with legacy control capabilities. The second allocation algorithm is inspired by probabilistic *MAC* methods from the networking context and implements a *try-error* access schema to the power grid that is extended to variable charging rates. Thereby, the charging processes sample random times from an exponential backoff window to defer their charging process in the case of grid asset overloading. *QoS* control is integrated by configuring additional waiting times and window sizes using charging service parameters.

The proposed solutions are evaluated with a co-simulation of the radial IEEE 906 low-voltage test feeder using realistic *EV* driving patterns from a mobility survey in Germany. In contrast to simpler queuing and probabilistic allocation strategies, *QoS*-aware allocation that considers laxity are superior in finishing all *EV* charging service in time. However, none of the applied approaches, namely *First-Come-First-Served* (*FCFS*), *Earliest-Departure-First* (*EDF*), *Least-Laxity-First* (*LLF*), *Proportional* (*PROP*), *Weighted Fair Queuing* (*WFQ*), *Dynamically Weighted Fair Queuing* (*DWFQ*), *Distributed Coordination Function* (*DCF*), and *Enhanced Distributed Channel Access* (*EDCA*), outperforms in all *QoS* metrics. Nevertheless, the proposed *DWFQ* performs well in most metrics and additionally achieves good impact values on power grid losses and the power quality in terms of voltage level. A high fairness level is desired to keep customers confident in the smart grid application, which is not necessarily measured only at the end of the charging service but can also be assessed throughout charging. This is highly important when some *EVs* need to leave earlier than their planned departure. Finally, the experiments show that intelligent *EV* charging control can limit the additional load for *EV* charging service to the peak load of the investigated low-voltage distribution grid without *EV* charging, while with smart resource allocation, all requirements of the *EV* driver can be fulfilled with high service quality.

In the domain of *flexibility-provision-as-a-service* to distributed home *EMS*, a *MILP* problem formulation is introduced in Chapter 4 that models flexible appliances at private households and hierarchically performs an offline optimization of the flexibility provided by *EV* charging processes and a stationary *ESS* using four objective functions. First, the energy consumption from the power grid is minimized, while the *EV* and *ESS* are utilized to directly consume local *PV* generation or store energy for later use. Second, the *EV* charging profiles are blocked such that no repetitive on-off switching of the charging services occurs. Finally, the maximum and minimum grid profile is reduced and the *ESS* usage is optimized to be able to compensate for forecast uncertainty. The solution space of the *MILP* allows scheduling external power flexibility requests to the *EMS*, which are then provided by the flexible assets. Multiple distributed *EMSs* are grouped to reach market-relevant flexibility potential. External flexibility requests change the optimal operation of the *EMSs* and may therefore impose an impact on the expectations of the user. This thesis defines a set of *QoS* and *QoE* metrics that capture the received and perceived service quality of flexibility disaggregation, which among others include metrics to measure the participation ratio to a flexibility request, the achieved level of self-consumption and energy autarky, and the total cost and flexibility delivery probability for the resource aggregator.

To disaggregate an external flexibility request to a set of distributed *EMSs*, two heuristics are proposed that consider several aspects of service quality during disaggregation. The linear heuristic iteratively assigns portions of the flexibility request to the single *EMSs* time slot after time slot, and can heuristically optimize different objectives using one of the four disaggregation policies: *Proportional* (*PROP*), *equal-share (EQUAL)*, *cost-optimal (COST)*, and *probability-optimal (POPT)*. In contrast to the linear heuristic, the meta heuristic can solve multiple objectives over multiple time slots at once and the proposed *GA* implementation returns the Pareto-optimal solution set, from which one solution can be chosen *a posteriori*. The *GA* operators are designed in a way to consider problem-specific constraints and produce valid offspring chromosomes with high probability. Furthermore, due to the iterative application of crossover and mutation, values of the multi-dimensional fitness function can be cached between generations, which consequently improves the performance of the *GA*.

Evaluations of the flexibility disaggregation on the same low-voltage test feeder with a flexibility pool of 55 *EMSs* show that both methods perform reasonably well in terms of calculation performance and targeted objectives. However, the performance in terms of *QoS* differs between the disaggregation policies and multi-objective weighting. The simple PROP disaggregation policy achieves both high quality metric scores and fairness index in most *QoS* metrics. However, the best disaggregation strategy highly depends on the specific use case, *e. g.*, cost-optimal and probability-optimal can achieve much better results for the resource aggregator with reduced service quality for the resource provider. The evaluation results further show that for flexibility disaggregation there is a trade-off between *QoS* to the user and power grid aspects in terms of grid losses and voltage levels. Finally, the investigated use case scenarios and their flexibility request do not have a large impact on the overall self-consumption and energy autarky of the single *EMS* and there are no big differences between different flexibility disaggregations.

5.2 Limitations and Outlook

The most important shortcoming of the thesis is that only technical *QoS* parameters are discussed for the two smart grid applications. However, non-technical quality aspects, which may have a large impact on the users' perception as well, are not considered. Examples include the billing process for *EV* charging and flexibility provision, or the simplicity to provide the required charging parameters to the system. In addition, the defined *QoS* and *QoE* metrics are based on a literature review and the expected perception of service quality by the author of the thesis. Extended user surveys are required to confirm these theories and understand the different importance of the *QoS* and *QoE* metrics. Only after including real users in the loop, one can make a valid judgment of the performance of *charging-as-a-service* algorithms on the *QoS* metrics perceived by the user. Nevertheless, this thesis already demonstrates that with the integration of service-oriented smart grid control, user satisfaction with technical *QoS* and *QoE* metrics can be improved.

This thesis relies on assumptions on *EV* charging, the power distribution grid, and distributed *ESSs*. The author of this thesis believes that when moving from private *EV* home

charging to workplace or public charging, some adaptions to the *QoS* metrics as well as the smart grid application algorithms are required. For example, users at public charging stations will most likely not care much about charging rate variations or the *SoC* at departure[28] but are interested more in the charged energy, which is the most important fact for short charging stops. On the other side, workplace charging is usually managed by load management software at the grid connection point, which can be integrated into the proposed system by modeling the load management as a *SU* that limits the charging current of the parking lot. Future *EV* charging control protocols with potential *V2X* support will allow a more fine-granular charging control, which may eliminate the minimum charging current and enable more advanced flexibility integration of the *EV* into home *EMSs*. This will require an adoption of the *MILP* at home *EMS*, however, have only a limited impact on the flexibility disaggregation.

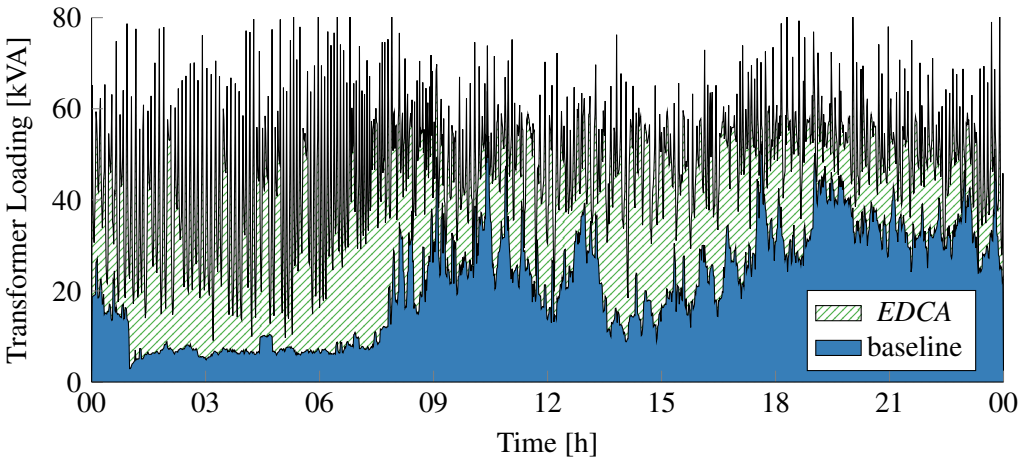
This thesis discusses two possible smart grid applications where users may get actively involved in future smart grid operations by providing flexibility to the power grid. In addition to that, further smart grid applications are worth investigating, *e. g.*, control of heat pumps which enables sector-coupling, or other consumer products like washing machines and air dryers. Furthermore, the potential interaction of parallel smart grid applications in the same power distribution grid remains an open field for research. This includes the research questions on *how to prioritize different applications that share the same power grid resources* and *how to utilize decentralized flexibility to support other smart grid applications in a* QoS*-aware manner*.

[28] *EV* drivers will most likely finish charging at around 80 % *SoC*, after which the charging rate will drop due to the saturation phase of the battery.

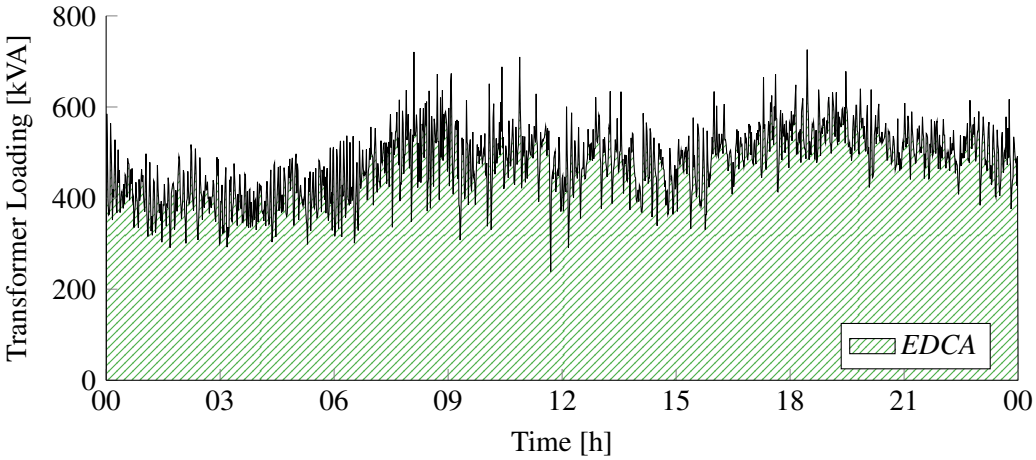
APPENDIX A

Additional Material

A.1 Electric Vehicle Home Charging Service



(a) Baseline and probabilistic methods using *EDCA*.



(b) Aggregated transformer loading using *EDCA* from ten independent simulation runs.

Figure A.1: Transformer loading during a simulated day using *EDCA*.

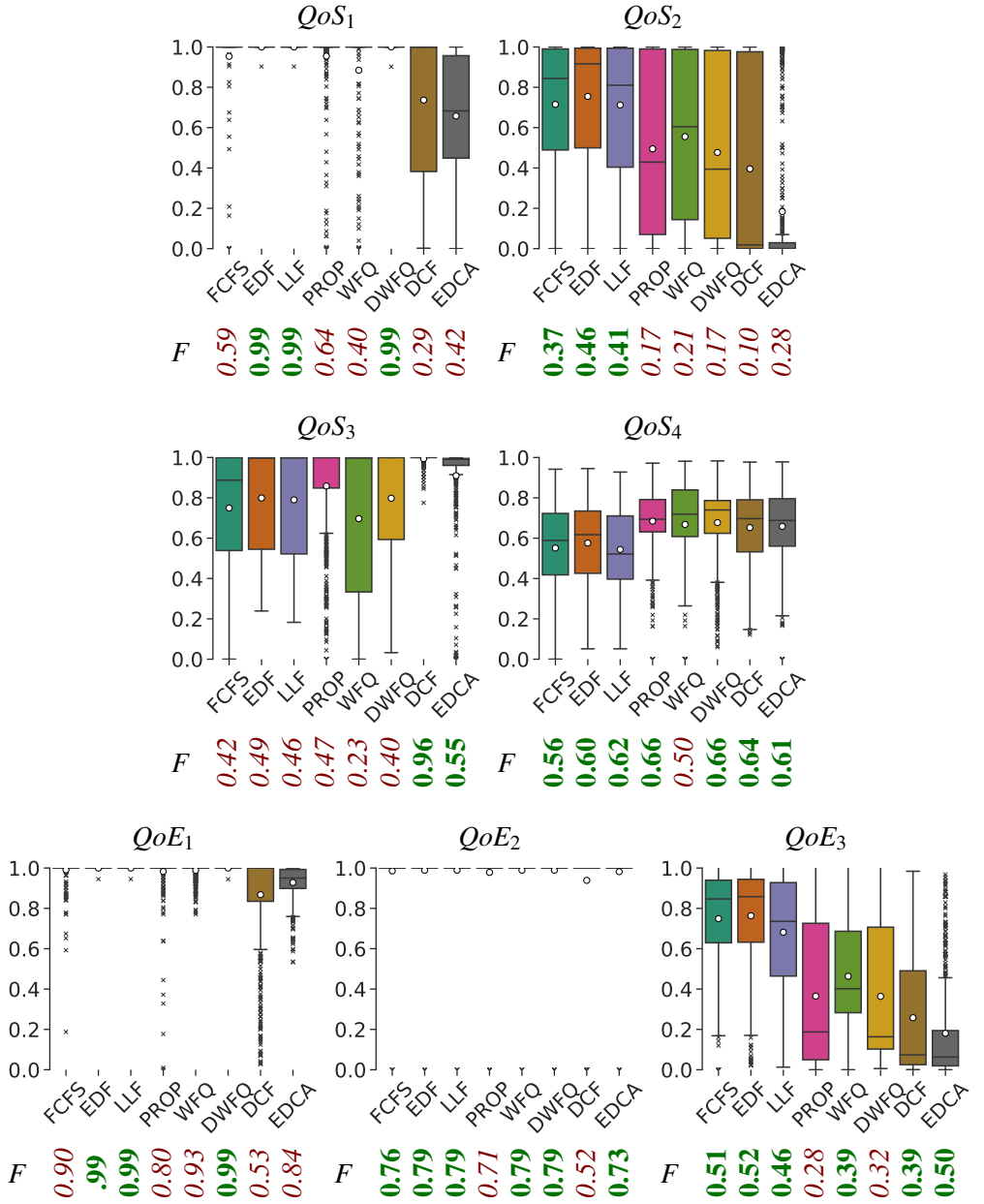


Figure A.2: *QoS* and *QoE* metrics of the different allocation policies for all 557 charging services. The box plots show the resulting distribution among the charging services and the circle denotes the average value. Below the box plots, the achieved fairness index F is given. Note that for QoS_1, QoE_1 and QoE_2 most of the charging service have very high service quality, hence the boxes are very near to 1.0. 11 kW wallboxes, compared to 22 kW wallboxes as in Figure 3.12

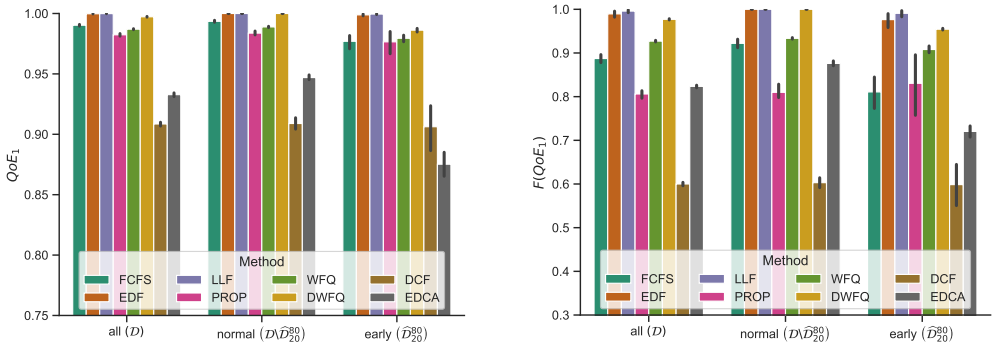


Figure A.3: Impact of $\widehat{\mathcal{D}}_{20}^{80}$ when 20 % of the electric vehicles depart at 80 % of their planned charging stay on the charging service quality (left) and fairness (right).

A.2 Power Flexibility from Battery Storage of Home Energy Management Systems

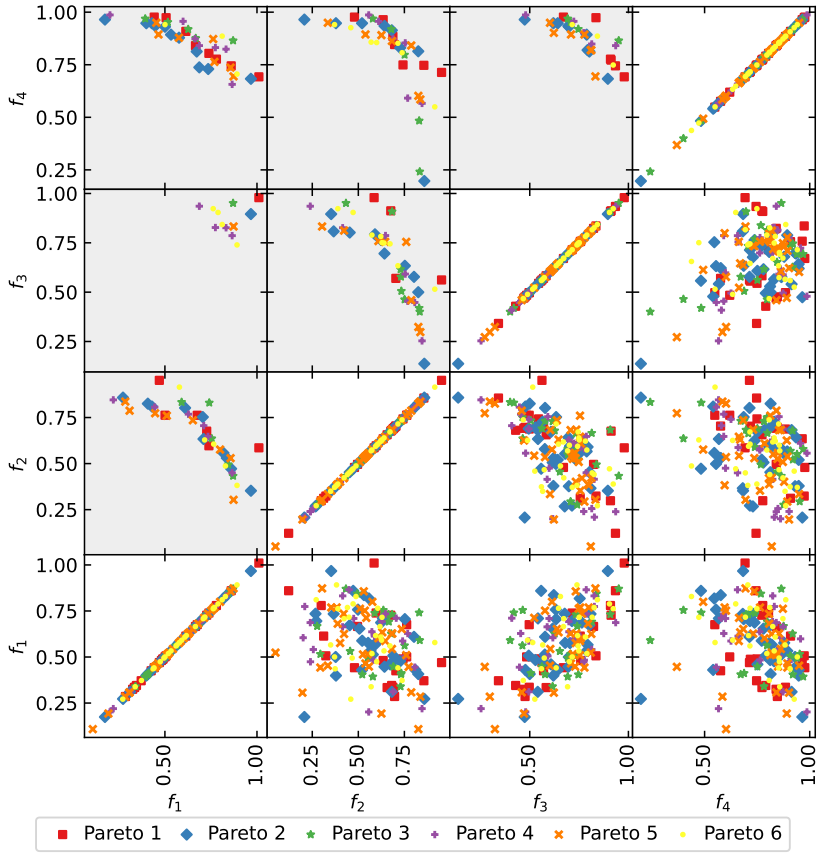
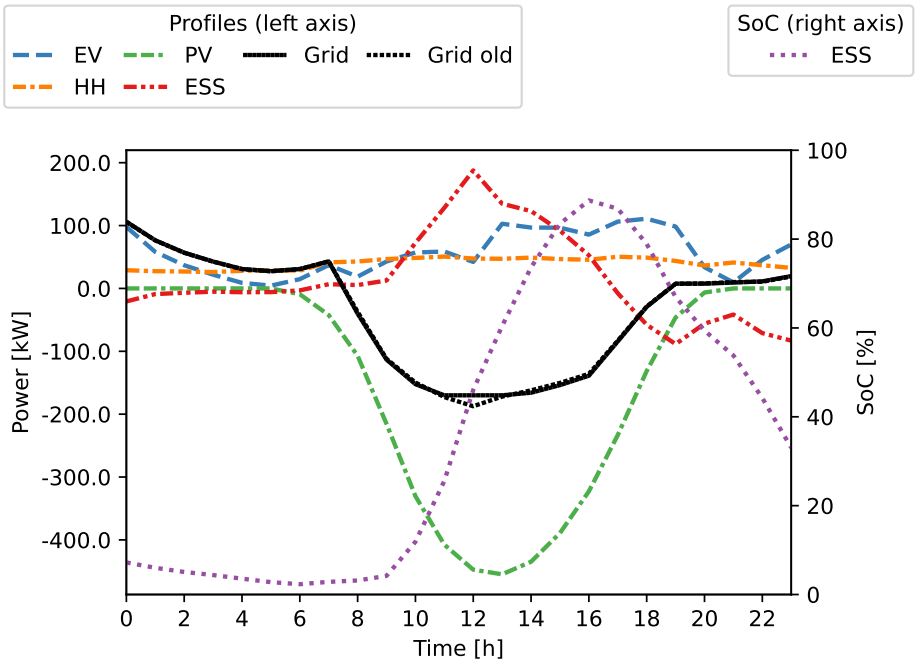
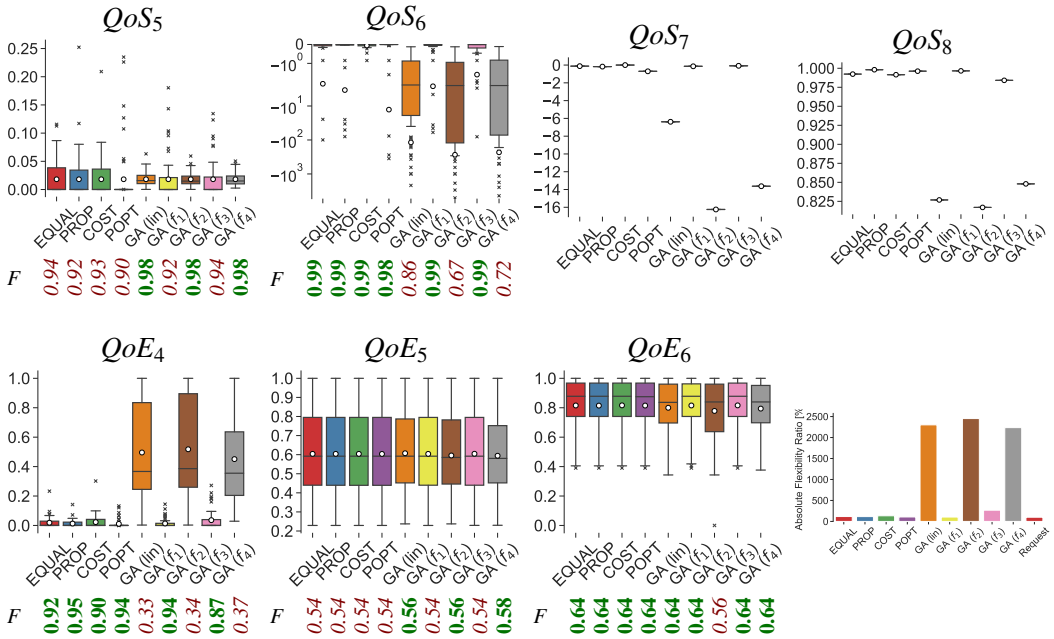


Figure A.4: Comparison of the fitness dimensions of f_{fit} from the Pareto-optimal solution set of six initial population sets of size 100.

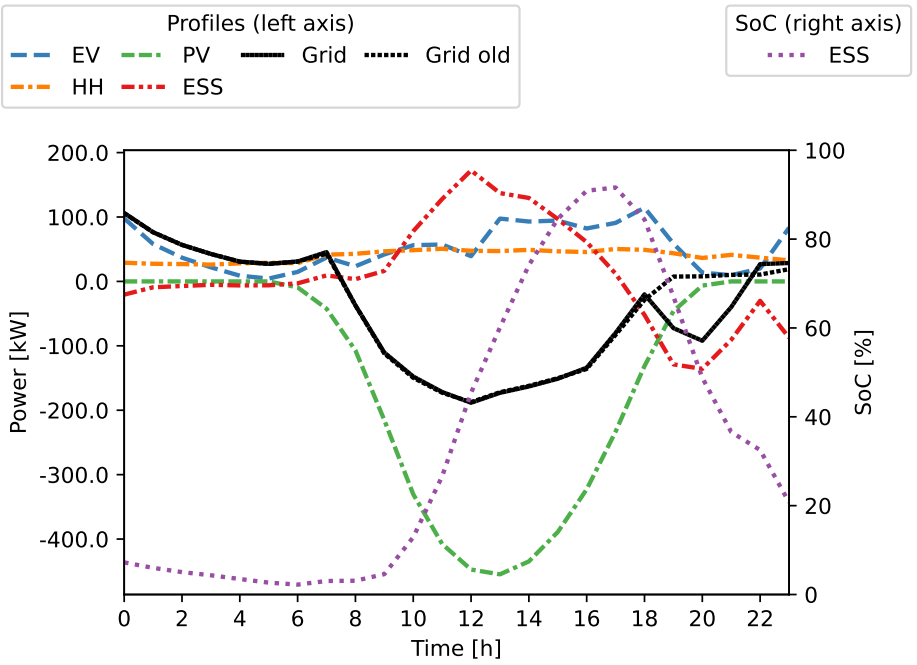


(a) Flexibility request and flexible resource scheduling of the flexibility pool using *PROP*.

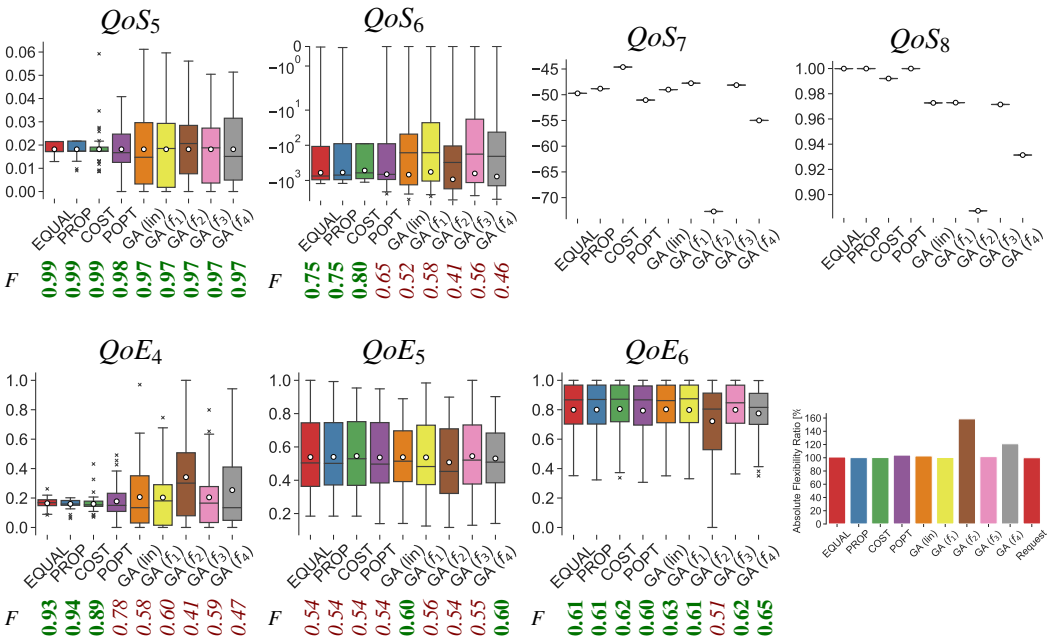


(b) Service quality metrics, fairness indices, and absolute scheduled flexibility.

Figure A.5: Flexibility disaggregation of small peak shaving.

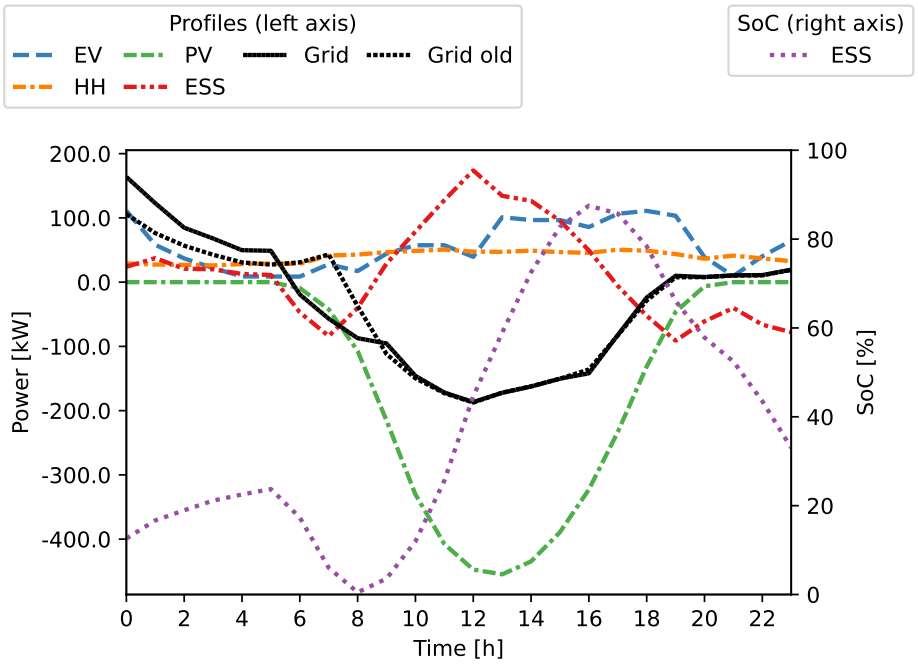


(a) Flexibility request and flexible resource scheduling of the flexibility pool using *PROP*.

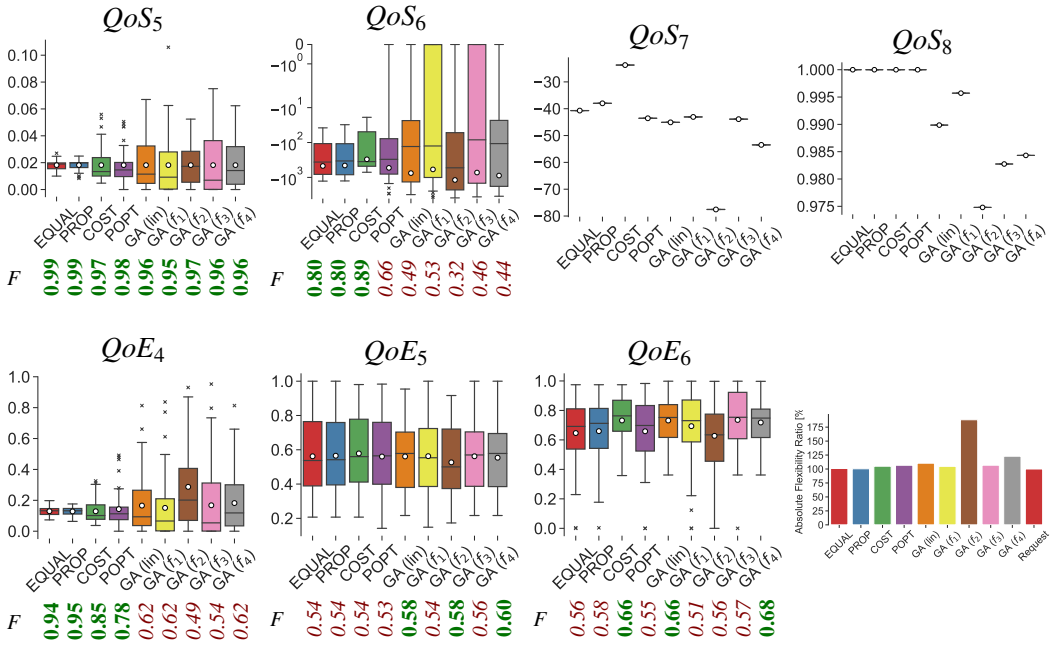


(b) Service quality metrics, fairness indices, and absolute scheduled flexibility.

Figure A.6: Flexibility disaggregation of evening peak shaving.



(a) Flexibility request and flexible resource scheduling of the flexibility pool using *PROP*.



(b) Service quality metrics, fairness indices, and absolute scheduled flexibility.

Figure A.7: Flexibility disaggregation of morning raise flattening.

References

[1] Laura Abraham et al. “enera - Projektkompendium”. Oldenburg: EWE Aktiengesellschaft, Feb. 2021, pages 1–436.

[2] Mohamed Abuella and Badrul Chowdhury. “Solar power forecasting using artificial neural networks”. In: *2015 North American Power Symposium (NAPS)*. Charlotte, NC, USA: IEEE, Oct. 2015, pages 1–5.

[3] George Aburn and Merlyn Hough. *Implementing EPA’s clean power plan: A menu of options*. Technical report. National Association of Clean Air Agencies (NACAA), May 2015. URL: https://www.eesi.org/files/NACAA_Menu_of_Options_LR.pdf (visited on Oct. 25, 2021).

[4] Michael Achatz et al. *D8.4 - Report on the final results of experiments and demonstrations*. Technical report. ELECTRIFIC Project Nº 713864, Aug. 2019. URL: https://cordis.europa.eu/project/id/713864/results (visited on July 13, 2022).

[5] Diego Acuña, Andres Ferragut, Fernando Paganini, and Enrique Briglia. “Symmetrical components analysis for managing phase imbalance in ev charge scheduling”. In: *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*. e-Energy ’22. Virtual Event: Association for Computing Machinery, 2022, pages 401–405.

[6] Abdullah Al Zishan, Moosa Moghimi Haji, and Omid Ardakanian. “Reputation-based fair power allocation to plug-in electric vehicles in the smart grid”. In: *2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS)*. Sydney, NSW, Australia: IEEE, Apr. 2020, pages 63–74.

[7] Zunaib Ali et al. “Multiobjective optimized smart charge controller for electric vehicle applications”. In: *IEEE Transactions on Industry Applications* 58.5 (2022), pages 5602–5615.

[8] Monica Alonso, Hortensia Amaris, Jean Gardy Germain, and Juan Manuel Galan. “Optimal charging scheduling of electric vehicles in smart grids by heuristic algorithms”. In: *Energies* 7.4 (Apr. 2014), pages 2449–2475.

[9] Jorge Nájera Álvarez, Katarina Knezović, and Mattia Marinelli. “Analysis and comparison of voltage dependent charging strategies for single-phase electric vehicles in an unbalanced Danish distribution grid”. In: *2016 51st International Universities Power Engineering Conference (UPEC)*. Coimbra, Portugal: IEEE, Sept. 2016, pages 1–6.

[10] Saeed Alyami, Abdulaziz Almutairi, and Omar Alrumayh. “Novel flexibility indices of controllable loads in relation to EV and rooftop PV”. In: *IEEE Transactions on Intelligent Transportation Systems* 24.1 (Jan. 2023), pages 923–931.

[11] Ammar Alyousef, Dominik Danner, Friederich Kupzog, and Hermann de Meer. “Enhancing power quality in electrical distribution systems using a smart charging architecture”. In: *Energy Informatics* 1.1 (Oct. 2018). Article 28, Proceedings of the 7th DACH+ Conference on Energy Informatics, Oldenburg, Germany, Best Paper Award.

[12] Ammar Alyousef and Hermann de Meer. “Design of a TCP-like smart charging controller for power quality in electrical distribution systems”. In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy ’19. Phoenix, AZ, USA: Association for Computing Machinery, June 2019, pages 128–138.

[13] Muhammad Bashar Anwar, Hassan W. Qazi, Daniel J. Burke, and Mark J. O’Malley. “Harnessing the flexibility of demand-side resources”. In: *IEEE Transactions on Smart Grid* 10.4 (July 2019), pages 4151–4163.

[14] Omid Ardakanian, Srinivasan Keshav, and Catherine Rosenberg. “Real-time distributed control for smart electric vehicle chargers: From a static to a dynamic study”. In: *IEEE Transactions on Smart Grid* 5.5 (Sept. 2014), pages 2295–2305.

[15] Omid Ardakanian, Catherine Rosenberg, and Srinivasan Keshav. “Distributed control of electric vehicle charging”. In: *Proceedings of the Fourth International Conference on Future Energy Systems*. e-Energy ’13. Berkeley, California, USA: ACM, Jan. 2013, pages 101–112.

[16] Omar Isaac Asensio, M. Cade Lawson, and Camila Z. Apablaza. “Electric vehicle charging stations in the workplace with high-resolution data from casual and habitual users”. In: *Scientific Data* 8.1 (July 2021). Article 168.

[17] Iason I. Avramidis, Florin Capitanescu, and Geert Deconinck. “From smart to sustainable to grid-friendly: A generic planning framework for enabling the transition between smart home archetypes”. In: *IEEE Transactions on Sustainable Energy* 12.3 (July 2021), pages 1684–1694.

[18] Chenhui Bai, Jiarui Wang, Jiewen Deng, Yu Liu, and Yi Wang. “A day-ahead scheduling strategy for electric vehicles based on converted economical objective”. In: *2022 IEEE 5th International Electrical and Energy Conference (CIEEC)*. Nangjing, China: IEEE, May 2022, pages 2565–2569.

[19] I. Safak Bayram et al. “Local energy storage sizing in plug-in hybrid electric vehicle charging stations under blocking probability constraints”. In: *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. Brussels, Belgium: IEEE, Oct. 2011, pages 78–83.

[20] Islam Safak Bayram, Ali Tajer, Mohamed Abdallah, and Khalid Qaraqe. “Capacity planning frameworks for electric vehicle charging stations with multiclass customers”. In: *IEEE Transactions on Smart Grid* 6.4 (July 2015), pages 1934–1943.

[21] J.C.R. Bennett and Hui Zhang. “WF/sup 2/Q: Worst-case fair weighted fair queueing”. In: *Proceedings of IEEE INFOCOM ’96. Conference on Computer Communications*. Volume 1. San Francisco, CA, USA: IEEE, Mar. 1996, pages 120–128.

[22] T. Bonald and A. Proutière. “Insensitive bandwidth sharing in data networks”. In: *Queueing Systems* 44.1 (May 2003), pages 69–100.

[23] Morris Brenna et al. “Automatic distributed voltage control algorithm in smart grids applications”. In: *IEEE Transactions on Smart Grid* 4.2 (June 2013), pages 877–885.

[24] Kjell Brunnström et al. *White paper on definitions of quality of experience*. Technical report. Qualinet, Mar. 2013. URL: https://hal.archives-ouvertes.fr/hal-00977812/document.

[25] BSI. *Technische Richtlinie BSI TR-03109-1: Anforderungen an die Interoperabilität der Kommunikationseinheit eines intelligenten Messsystems, Version 1.1*. Technical report. Bundesamt für Sicherheit in der Informationstechnik, Sept. 2021. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03109/TR03109-1.pdf (visited on July 1, 2022).

[26] Niangjun Chen, Christian Kurniawan, Yorie Nakahira, Lijun Chen, and Steven H. Low. “Smoothed least-laxity-first algorithm for electric vehicle charging: Online decision and performance analysis with resource augmentation”. In: *IEEE Transactions on Smart Grid* 13.3 (May 2022), pages 2209–2217.

[27] Yu-Wen Chen, Xiuxing Chen, and Nicholas Maxemchuk. “The fair allocation of power to air conditioners on a smart grid”. In: *IEEE Transactions on Smart Grid* 3.4 (Dec. 2012), pages 2188–2195.

[28] Ching-Yen Chung, Joshua Chynoweth, Chi-Cheng Chu, and Rajit Gadh. “Master-slave control scheme in electric vehicle smart charging infrastructure”. In: *The Scientific World Journal* 2014 (May 2014). Article ID 462312.

[29] Andres Cortés and Sonia Martínez. “A hierarchical algorithm for optimal plug-in electric vehicle charging with usage constraints”. In: *Automatica* 68 (June 2016), pages 119–131.

[30] Matej Črepinšek, Shih-Hsi Liu, and Marjan Mernik. “Exploration and exploitation in evolutionary algorithms: A survey”. In: *ACM Computing Survey* 45.3 (June 2013). Article 35.

[31] Dominik Danner, Ammar Alyousef, Philipp Danner, Wolfgang Duschl, and Hermann de Meer. “Towards grid-friendly electric vehicle charging: Architectural concept and field trials”. In: *3rd E-Mobility Power System Integration Symposium*. Dublin, Republic of Ireland: Energynautics, Oct. 2019.

[32] Dominik Danner, Wolfgang Duschl, and Hermann de Meer. “Fair charging service allocation for electric vehicles in the power distribution grid”. In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy ’19. Poster. Phoenix, AZ, USA: Association for Computing Machinery, June 2019, pages 406–408.

[33] Dominik Danner, Robin Huwa, and Hermann de Meer. “Multi-objective flexibility disaggregation to distributed energy management systems”. In: *ACM SIGEnergy Energy Informatics Review* 2.2 (June 2022), pages 1–12.

[34] Dominik Danner and Hermann de Meer. “State estimation in the power distribution system”. In: *ACM SIGMETRICS Performance Evaluation Review* 46.3 (Dec. 2018). ePerf Workshop: Performance and Modelling of Energy Systems, Toulouse, France, pages 86–88.

[35] Dominik Danner and Hermann de Meer. “Quality of service and fairness for electric vehicle charging as a service”. In: *Energy Informatics* 4.3 (Sept. 2021). Article 16, Proceedings of the 10th DACH+ Conference on Energy Informatics, Virtual.

[36] Dominik Danner and Hermann de Meer. “Max-consensus protocol to determine the regulated node in distributed voltage regulation”. In: *Energy Informatics* 5.1 (Sept. 2022). Article 15, Proceedings of the 11th DACH+ Conference on Energy Informatics, Freiburg, Germany.

[37] Dominik Danner, Jan Seidemann, Michael Lechl, and Hermann de Meer. “Flexibility disaggregation under forecast conditions”. In: *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. e-Energy ’21. Virtual Event, Italy: Association for Computing Machinery, June 2021, pages 27–38.

[38] Philipp Danner, Wolfgang Duschl, Dominik Danner, Ammar Alyousef, and Hermann De Meer. “Flexibility reward scheme for grid-friendly electric vehicle charging in the distribution power grid”. In: *Proceedings of the Ninth International Conference on Future Energy Systems*. e-Energy ’18. 3rd Workshop on Electric Vehicle Systems, Data, and Applications. Karlsruhe, Germany: Association for Computing Machinery, June 2018, pages 564–569.

[39] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. “A fast and elitist multiobjective genetic algorithm: NSGA-II”. In: *IEEE Transactions on Evolutionary Computation* 6.2 (Apr. 2002), pages 182–197.

[40] Merkebu Zenebe Degefa, Iver Bakken Sperstad, and Hanne Sæle. “Comprehensive classifications and characterizations of power system flexibility resources”. In: *Electric Power Systems Research* 194 (May 2021). Article 107022.

[41] Sara Deilami, Amir S. Masoum, Paul S. Moses, and Mohammad A. S. Masoum. “Real-time coordination of plug-in electric vehicle charging in smart grids to minimize power losses and improve voltage profile”. In: *IEEE Transactions on Smart Grid* 2.3 (Sept. 2011), pages 456–467.

[42] Alan Demers, Srinivasan Keshav, and Scott Shenker. “Analysis and simulation of a fair queueing algorithm”. In: *ACM SIGCOMM Computer Communication Review* 19.4 (Sept. 1989), pages 1–12.

[43] Youjun Deng, Yongxi Zhang, Fengji Luo, and Gianluca Ranzi. “Many-objective HEMS based on multi-scale occupant satisfaction modelling and second-life BESS utilization”. In: *IEEE Transactions on Sustainable Energy* 13.2 (Apr. 2022), pages 934–947.

[44] Esther Dudek. “The flexibility of domestic electric vehicle charging: The electric nation project”. In: *IEEE Power and Energy Magazine* 19.4 (Aug. 2021), pages 16–27.

[45] E.ON. *E.ON grids will be upgraded for 100 percent e-mobility*. May 2019. URL: https://www.eon.com/en/about-us/media/press-release/2019/eon-board-member-thomas-koenig-presents-new-study.html (visited on Mar. 4, 2021).

[46] Matthias Ehrgott. "Multicriteria optimization". 2nd edition. Springer, Berlin, Heidelberg, May 2005.

[47] Christine Eisenmann, Bastian Chlond, Tim Hilgert, Sascha von Behren, and Peter Vortisch. *Deutsches Mobilitätspanel (MOP)-Wissenschaftliche Begleitung und Auswertungen Bericht 2016/2017: Alltagsmobilität und Fahrleistung*. Technical report. Karlsruher Institut für Technologie, Mar. 2018. URL: https://fops.de/wp-content/uploads/2020/08/70.923_Schlussbericht_MOP_Jahresbericht-2016-2017.pdf (visited on Feb. 1, 2020).

[48] ENTSO-E. *Market committee - Balancing and ancillary services markets*. July 2021. URL: https://www.entsoe.eu/about/market/#balancing-and-ancillary-services-markets (visited on July 20, 2021).

[49] ENTSO-E. *System separation in the Continental Europe synchronous area on 8 January 2021 – 2nd update*. Jan. 2021. URL: https://www.entsoe.eu/news/2021/01/26/system-separation-in-the-continental-europe-synchronous-area-on-8-january-2021-2nd-update/ (visited on Mar. 4, 2021).

[50] EPEX. *New trading platform boosts EPEX SPOT's Localflex offer*. Nov. 2021. URL: https://www.epexspot.com/en/news/new-trading-platform-boosts-epex-spots-localflex-offer (visited on Mar. 7, 2023).

[51] Melike Erol-Kantarci, Jahangir H. Sarker, and Hussein T. Mouftah. "Quality of service in plug-in electric vehicle charging infrastructure". In: *2012 IEEE International Electric Vehicle Conference*. Greenville, SC, USA: IEEE, Mar. 2012, pages 1–5.

[52] Zhong Fan. "A distributed demand response algorithm and its application to PHEV charging in smart grids". In: *IEEE Transactions on Smart Grid* 3.3 (Sept. 2012), pages 1280–1290.

[53] Yingjia Fang et al. "Optimal home energy management with distributed generation and energy storage systems". In: *2021 31st Australasian Universities Power Engineering Conference (AUPEC)*. Perth, Australia: IEEE, Sept. 2021, pages 1–6.

[54] Stefan Feilmeier et al. "OpenEMS 2021.1.0". OpenEMS, Jan. 2021.

[55] Zahra Foroozandeh, Sérgio Ramos, João Soares, and Zita Vale. "Goal programming approach for energy management of smart building". In: *IEEE Access* 10 (Feb. 2022), pages 25341–25348.

[56] Oliver Frendo, Nadine Gaertner, and Heiner Stuckenschmidt. "Real-time smart charging based on precomputed schedules". In: *IEEE Transactions on Smart Grid* 10.6 (Nov. 2019), pages 6921–6932.

[57] Daniel Gebbran, Sleiman Mhanna, Yiju Ma, Archie C. Chapman, and Gregor Verbič. “Fair coordination of distributed energy resources with Volt-Var control and PV curtailment”. In: *Applied Energy* 286 (Mar. 2021). Article 116546.

[58] Yasin Zabihinia Gerdroodbari, Reza Razzaghi, and Farhad Shahnia. “Decentralized control strategy to improve fairness in active power curtailment of PV inverters in low-voltage distribution networks”. In: *IEEE Transactions on Sustainable Energy* 12.4 (Oct. 2021), pages 2282–2292.

[59] S.J. Golestani. “A self-clocked fair queueing scheme for broadband applications”. In: *Proceedings of INFOCOM ’94 Conference on Computer Communications*. Volume 2. Toronto, ON, Canada: IEEE, June 1994, pages 636–646.

[60] Hessam Golmohamadi, Reza Keypour, Birgitte Bak-Jensen, and Jayakrishnan Radhakrishna Pillai. “Optimization of household energy consumption towards day-ahead retail electricity price in home energy management systems”. In: *Sustainable Cities and Society* 47 (May 2019). Article 101468.

[61] Markus Graebig and Michael Wolfram. “WindNODE Jahrbuch 2020 - Das Schaufenster für intelligente Energie”. Berlin: 50Hertz Transmission GmbH, 2020, pages 1–266.

[62] Albert G. Greenberg and Neal Madras. “How fair is fair queuing”. In: *Journal of the ACM* 39.3 (July 1992), pages 568–598.

[63] David Micheal Greenwood et al. “Frequency response services designed for energy storage”. In: *Applied Energy* 203 (Oct. 2017), pages 115–127.

[64] Zunaib Maqsood Haider et al. “Water-filling algorithm based approach for management of responsive residential loads”. In: *Journal of Modern Power Systems and Clean Energy* 6.1 (Jan. 2018), pages 118–131.

[65] Zunaib Maqsood Haider et al. “Optimal management of a distribution feeder during contingency and overload conditions by harnessing the flexibility of smart loads”. In: *IEEE Access* 9 (Mar. 2021), pages 40124–40139.

[66] Azhar Ul-Haq, Concettina Buccella, Carlo Cecati, and Hassan A. Khalid. “Smart charging infrastructure for electric vehicles”. In: *2013 International Conference on Clean Electrical Power (ICCEP)*. Alghero, Italy: IEEE, June 2013, pages 163–169.

[67] Tuncer Haslak. “Weighted fair queuing as a scheduling algorithm for deferrable loads in smart grids”. In: *Advances in Energy System Optimization*. Cham: Springer International Publishing, 2020, pages 123–141.

[68] Felix Heider, Amra Jahic, Maik Plenz, Konstantin Tröger, and Detlef Schulz. “A generic EV charging model extracted from real charging behaviour”. In: *2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET)*. Arad, Romania: IEEE, 2022, pages 393–398.

[69] Emil Hillberg et al. *Flexibility needs in the future power system*. Technical report. ISGAN, Mar. 2019. URL: https://www.iea-isgan.org/wp-content/uploads/2019/03/ISGAN_DiscussionPaper_Flexibility_Needs_In_Future_Power_Systems_2019.pdf (visited on Feb. 27, 2023).

[70] Christian Hinrichs and Michael Sonnenschein. “A distributed combinatorial optimisation heuristic for the scheduling of energy resources represented by self-interested agents”. In: *International Journal of Bio-Inspired Computation* 10.2 (Jan. 2017), pages 69–78.

[71] Md Murshadul Hoque, Mohsen Khorasany, Reza Razzaghi, Mahdi Jalili, and Hao Wang. “Network-aware coordination of aggregated electric vehicles considering charge-discharge flexibility”. In: *IEEE Transactions on Smart Grid* (June 2022). Early Access.

[72] Vahid Hosseinnezhad, Miadreza Shafie-khah, Pierluigi Siano, and João P.S. Catalão. “Optimal home energy management for electric flexibility provision”. In: *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. Bucharest, Romania: IEEE, Oct. 2019, pages 1–6.

[73] Tobias Hoßfeld, Lea Skorin-Kapov, Poul E. Heegaard, and Martín Varela. “Definition of QoE fairness in shared systems”. In: *IEEE Communications Letters* 21.1 (Jan. 2017), pages 184–187.

[74] Tobias Hoßfeld, Lea Skorin-Kapov, Poul E. Heegaard, and Martín Varela. “A new QoE fairness index for QoE management”. In: *Quality and User Experience* 3.1 (Feb. 2018), page 4.

[75] Qian Hu, Siqi Bu, and Vladimir Terzija. “A distributed P and Q provision based voltage regulation scheme by incentivized EV fleet charging for resistive distribution networks”. In: *IEEE Transactions on Transportation Electrification* 7.4 (Dec. 2021), pages 2376–2389.

[76] Akhtar Hussain and Petr Musilek. “Fairness and utilitarianism in allocating energy to EVs during power contingencies using modified division rules”. In: *IEEE Transactions on Sustainable Energy* 13.3 (July 2022), pages 1444–1456.

[77] Shahid Hussain, Mohamed A. Ahmed, and Young-Chon Kim. “Efficient power management algorithm based on fuzzy logic inference for electric vehicles parking lot”. In: *IEEE Access* 7 (2019), pages 65467–65485.

[78] Shahid Hussain, Mohamed A. Ahmed, Ki-Beom Lee, and Young-Chon Kim. “Fuzzy logic weight based charging scheme for optimal distribution of charging power among electric vehicles in a parking lot”. In: *Energies* 13.12 (June 2020). Article 3119.

[79] Shahid Hussain, Ki-Beom Lee, Mohamed A. Ahmed, Barry Hayes, and Young-Chon Kim. “Two-stage fuzzy logic inference algorithm for maximizing the quality of performance under the operational constraints of power grid in electric vehicle parking lots”. In: *Energies* 13.18 (Sept. 2020). Article 4634.

[80] IEA. *Global EV outlook 2020*. Technical report. Paris: IEA, June 2020. URL: `https://www.iea.org/reports/global-ev-outlook-2020` (visited on May 11, 2021).

[81] IEA. *Annual energy storage deployment, 2013-2019*. Technical report. Paris: IEA, Oct. 2022. URL: https://www.iea.org/data-and-statistics/charts/annual-energy-storage-deployment-2013-2019-2 (visited on Feb. 27, 2023).

[82] IEA. *Global electric car stock, 2010-2021*. Technical report. Paris: IEA, Oct. 2022. URL: https://www.iea.org/data-and-statistics/charts/global-electric-car-stock-2010-2021 (visited on Feb. 27, 2023).

[83] IEA. *Heat pumps*. Technical report. Paris: IEA, Sept. 2022. URL: https://www.iea.org/reports/heat-pumps.

[84] Muhandiram Arachchige Subodha Tharangi Ireshika et al. "Optimal power tracking for autonomous demand side management of electric vehicles". In: *Journal of Energy Storage* 52 (Aug. 2022). Article 104917.

[85] Md Shariful Islam, Nadarajah Mithulananthan, and Kwang Y. Lee. "Suitability of PV and battery storage in EV charging at business premises". In: *IEEE Transactions on Power Systems* 33.4 (July 2018), pages 4382–4396.

[86] ITU. *ITU-T E.800 - Series E: Overall network operation, telephone service, service operation and human factors*. Technical report. ITU-T, Sept. 2008, page 30. URL: https://www.itu.int/rec/T-REC-E.800-200809-I/en.

[87] J. Jaffe. "Bottleneck flow control". In: *IEEE Transactions on Communications* 29.7 (July 1981), pages 954–962.

[88] Raj Jain, Dah-Ming Chiu, and William Hawe. *Dec-TR-301 - A quantitative measure of fairness and discrimination for resource allocation in share computer system*. Technical report. Eastern Research Lab, Sept. 1984. URL: https://arxiv.org/abs/cs/9809099 (visited on Feb. 27, 2023).

[89] Raka Jovanovic, Sertac Bayhan, and Islam Safak Bayram. "Optimization of electric vehicle charge scheduling with consideration of battery degradation". In: *2022 24th European Conference on Power Electronics and Applications (EPE'22 ECCE Europe)*. Hanover, Germany: IEEE, Sept. 2022, pages 1–11.

[90] Kai-Philipp Kairies et al. "Market and technology development of PV home storage systems in Germany". In: *Journal of Energy Storage* 23 (June 2019), pages 416–424.

[91] Frank Kelly. "Charging and rate control for elastic traffic". In: *European Transactions on Telecommunications* 8.1 (Feb. 1997), pages 33–37.

[92] Srinivasan Keshav. "An engineering approach to computer networking: ATM networks, the Internet, and the telephone network". 9th edition. Reading, MA, USA: Addison-Wesley Longman Publishing Co., Inc., May 1997, pages 1–660.

[93] KFW. *Ladestationen für Elektroautos – Wohngebäude*. June 2021. URL: https://www.kfw.de/s/deiBGYew (visited on June 25, 2021).

[94] Fanxin Kong, Xue Liu, Zhonghao Sun, and Qinglong Wang. "Smart rate control and demand balancing for electric vehicle charging". In: *2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)*. Vienna, Austria: IEEE, Apr. 2016, pages 1–10.

[95] Eleftherios O. Kontis, Georgios C. Kryonidis, Andreas I. Chrysochos, Charis S. Demoulias, and Grigoris K. Papagiannis. “Effect of load modelling in coordinated active power curtailment of distributed renewable energy sources”. In: *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016)*. Belgrade, Serbia: Institution of Engineering and Technology, Nov. 2016, pages 1–8.

[96] Georgios C. Kryonidis, Charis S. Demoulias, and Grigoris K. Papagiannis. “A new voltage control scheme for active medium-voltage (MV) networks”. In: *Electric Power Systems Research* 169 (Apr. 2019), pages 53–64.

[97] Georgios C. Kryonidis et al. “Distributed reactive power control scheme for the voltage regulation of unbalanced LV grids”. In: *IEEE Transactions on Sustainable Energy* 12.2 (Apr. 2021), pages 1301–1310.

[98] Michael Lechl and Stefan Feilmeier. *EMSIG: Energy management system data*. OpenEMS Association e.V. Apr. 2021. URL: `https://openems.io/research/emsig/` (visited on Nov. 2, 2021).

[99] Donghun Lee and Kwanho Kim. “Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information”. In: *Energies* 12.2 (Jan. 2019). Article 215.

[100] Zachary J. Lee, Tongxin Li, and Steven H. Low. “ACN-Data: Analysis and applications of an open EV charging dataset”. In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy ’19. Phoenix, AZ, USA: Association for Computing Machinery, June 2019, pages 139–149.

[101] Zachary J. Lee, Sunash Sharma, Daniel Johansson, and Steven H. Low. “ACN-Sim: An open-source simulator for data-driven electric vehicle charging research”. In: *IEEE Transactions on Smart Grid* 12.6 (Nov. 2021), pages 5113–5123.

[102] Zachary J. Lee et al. “Adaptive charging networks: A framework for smart electric vehicle charging”. In: *IEEE Transactions on Smart Grid* 12.5 (Sept. 2021), pages 4339–4350.

[103] Juri Lelli, Claudio Scordino, Luca Abeni, and Dario Faggioli. “Deadline scheduling in the Linux kernel”. In: *Software: Practice and Experience* 46.6 (June 2016), pages 821–839.

[104] Antoine Lesage-Landry, Han Wang, Iman Shames, Pierluigi Mancarella, and Joshua A. Taylor. “Online convex optimization of multi-energy building-to-grid ancillary services”. In: *IEEE Transactions on Control Systems Technology* 28.6 (Nov. 2020), pages 2416–2431.

[105] Hanoch Levy, Benjamin Avi-Itzhak, and David Raz. “Principles of fairness quantification in queueing systems”. In: *Network Performance Engineering: A Handbook on Convergent Multi-Service Networks and Next Generation Internet*. Lecture Notes in Computer Science, vol 5233. Berlin, Heidelberg: Springer, 2011, pages 284–300.

[106] Yang Li, Meng Han, Zhen Yang, and Guoqing Li. “Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle charging station: A bi-level approach”. In: *IEEE Transactions on Sustainable Energy* 12.4 (Oct. 2021), pages 2321–2331.

[107] Fynn Liegmann, Alen Murtovi, Michael Kelker, and Jens Haubrock. “Analysis of user behaviour for modelling an electric vehicle loading profile generator”. In: *PESS 2021; Power and Energy Student Summit*. online: VDE, Nov. 2021, pages 1–5.

[108] Shengbo Liu, Liqun Fu, and Wei Xie. “Hidden-node problem in full-duplex enabled CSMA networks”. In: *IEEE Transactions on Mobile Computing* 19.2 (Feb. 2020), pages 347–361.

[109] J. A. Pecas Lopes, F. J. Soares, and P. M. Rocha Almeida. “Identifying management procedures to deal with connection of Electric Vehicles in the grid”. In: *2009 IEEE Bucharest PowerTech*. Bucharest, Romania: IEEE, July 2009, pages 1–8.

[110] Elke Lorenz, Johannes Hurka, Detlev Heinemann, and Hans Georg Beyer. “Irradiance forecasting for the power prediction of grid-connected photovoltaic systems”. In: *IEEE Journal of selected topics in applied earth observations and remote sensing* 2.1 (Mar. 2009), pages 2–10.

[111] Peter D. Lund, Juuso Lindgren, Jani Mikkola, and Jyri Salpakari. “Review of energy system flexibility measures to enable high levels of variable renewable electricity”. In: *Renewable and Sustainable Energy Reviews* 45 (May 2015), pages 785–807.

[112] Davye Mak and Dae-Hyun Choi. “Smart home energy management in unbalanced active distribution networks considering reactive power dispatch and voltage control”. In: *IEEE Access* 7 (Oct. 2019), pages 149711–149723.

[113] Yuri V. Makarov, Clyde Loutan, Jian Ma, and Phillip de Mello. “Operational impacts of wind generation on California power systems”. In: *IEEE Transactions on Power Systems* 24.2 (May 2009), pages 1039–1050.

[114] J. Markkula, V. Tikka, and P. Järventausta. “Local versus centralized control of flexible loads in Power grid”. In: *CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution*. Volume 2021. Online Conference: IET, Sept. 2021, pages 2294–2298.

[115] Sergejus Martinenas, Katarina Knezović, and Mattia Marinelli. “Management of power quality issues in low voltage networks using electric vehicles: Experimental validation”. In: *IEEE Transactions on Power Delivery* 32.2 (Apr. 2017), pages 971–979.

[116] Jian Meng et al. “Dynamic frequency response from electric vehicles in the Great Britain power system”. In: *Journal of Modern Power Systems and Clean Energy* 3.2 (May 2015), pages 203–211.

[117] Michael Merten. “Participation of battery storage systems in the automatic frequency restoration reserve market based on machine learning”. PhD thesis. Rheinisch-Westfälische Technische Hochschule Aachen, July 2020.

[118] Hyeon-Jin Moon, Ah-Yun Yun, Eung-Sang Kim, and Seung-Il Moon. “An analysis of energy storage systems for primary frequency control of power systems in South Korea”. In: *Energy Procedia* 107 (Feb. 2017). 3rd International Conference on Energy and Environment Research, ICEER 2016, 7-11 September 2016, Barcelona, Spain, pages 116–121.

[119] Pedro Moura, Greta K.W. Yu, and Javad Mohammadi. “Management of electric vehicles as flexibility resource for optimized integration of renewable energy with large buildings”. In: *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. The Hague, Netherlands: IEEE, Oct. 2020, pages 474–478.

[120] Yunfei Mu, Jianzhong Wu, Janaka Ekanayake, Nick Jenkins, and Hongjie Jia. “Primary frequency response from electric vehicles in the great britain power system”. In: *IEEE Transactions on Smart Grid* 4.2 (June 2013), pages 1142–1150.

[121] Fabian L Müller, Jácint Szabó, Olle Sundström, and John Lygeros. “Aggregation and disaggregation of energetic flexibility from distributed energy resources”. In: *IEEE Transactions on Smart Grid* 10.2 (Mar. 2019), pages 1205–1214.

[122] Christopher Neuman, Andrew Meintz, and Myungsoo Jun. “Workplace charging data”. United States, Jan. 2022.

[123] Duong Tung Nguyen and Long Bao Le. “Joint optimization of electric vehicle and home energy scheduling considering user comfort preference”. In: *IEEE Transactions on Smart Grid* 5.1 (Jan. 2014), pages 188–199.

[124] Hussam Nosair and François Bouffard. “Energy-centric flexibility management in power systems”. In: *IEEE Transactions on Power Systems* 31.6 (Nov. 2016), pages 5071–5081.

[125] Hussam Nosair and Franois Bouffard. “Flexibility envelopes for power system operational planning”. In: *IEEE Transactions on Sustainable Energy* 6.3 (July 2015), pages 800–809.

[126] Álvaro Rodríguez del Nozal, Eleftherios O. Kontis, Juan M. Mauricio, and Charis S. Demoulias. “Provision of inertial response as ancillary service from active distribution networks to the transmission system”. In: *IET Generation, Transmission & Distribution* 14.22 (Nov. 2020), pages 5123–5134.

[127] Open Charge Alliance. *Open Charge Point Protocol 2.0.1*. Technical report. Open Charge Alliance, Oct. 2020. URL: `https://www.openchargealliance.org/downloads/` (visited on June 25, 2021).

[128] Sara Ostovar, Moein Moeini-Aghtaie, and Mohammad Behzad Hadi. “Developing a new flexibility-based algorithm for home energy management system (HEMS)”. In: *2020 10th Smart Grid Conference (SGC)*. Kashan, Iran: IEEE, Dec. 2020, pages 1–6.

[129] Konstantinos Oureilidis et al. *D5.2 - Report presenting the portfolio of ancillary services*. Technical report. EASY-RES Project Nº 764090, Feb. 2019. URL: `https://cordis.europa.eu/project/id/764090/results` (visited on Aug. 1, 2022).

[130] Torben Bach Pedersen, Laurynas Šikšnys, and Bijay Neupane. “Modeling and managing energy flexibility using FlexOffers”. In: *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. Aalborg, Denmark: IEEE, Oct. 2018, pages 1–7.

[131] Maria Perez-Ortega et al. *D9.6 - Final exploitation framework: Market analysis, project impact and sustainability plan*. Technical report. ELECTRIFIC Project Nº 713864, Sept. 2019. URL: https://cordis.europa.eu/project/id/713864/results (visited on July 13, 2022).

[132] Dorothee Peters, Rasmus Völker, Frank Schuldt, and Karsten von Maydell. “Are standard load profiles suitable for modern electricity grid models?” In: *2020 17th International Conference on the European Energy Market (EEM)*. Stockholm, Sweden: IEEE, Sept. 2020, pages 1–6.

[133] Alain Petrowski. “A clearing procedure as a niching method for genetic algorithms”. In: *Proceedings of IEEE International Conference on Evolutionary Computation*. Nagoya, Japan: IEEE, May 1996, pages 798–803.

[134] Matthew Rabin. “Incorporating fairness into game theory and economics”. In: *The American economic review* 83.5 (Dec. 1993), pages 1281–1302. URL: https://www.jstor.org/stable/2117561.

[135] Sasan Rafii-Tabrizi, Jean-Régis Hadji-Minaglou, Frank Scholzen, and Florin Capitanescu. “Optimal operation of nearly zero energy buildings using mixed integer linear programming”. In: *2019 International Conference on Smart Energy Systems and Technologies (SEST)*. Porto, Portugal: IEEE, Sept. 2019, pages 1–6.

[136] Xiaohong Ran and Shipeng Leng. “Enhanced robust index model for load scheduling of a home energy local network with a load shifting strategy”. In: *IEEE Access* 7 (Jan. 2019), pages 19943–19953.

[137] Kalle Rauma, Alexander Funke, Toni Simolin, Pertti Järventausta, and Christian Rehtanz. “Electric vehicles as a flexibility provider: Optimal charging schedules to improve the quality of charging service”. In: *Electricity* 2.3 (June 2021), pages 225–243.

[138] Frank Reyer et al. *Continental Europe Synchronous area separation on 8 January 2021 - Final report*. Technical report. ENTSO-E, July 2021. URL: https://eepublicdownloads.azureedge.net/clean-documents/SOC%20documents/SOC%20Reports/entso-e_CESysSep_Final_Report_210715.pdf (visited on July 15, 2021).

[139] Frank Reyer et al. *Continental Europe synchronous area separation on 8 January 2021 - Interim report*. Technical report. ENTSO-E, Feb. 2021. URL: https://eepublicdownloads.azureedge.net/clean-documents/Publications/Position%20papers%20and%20reports/entso-e_CESysSep_interim_report_210225.pdf (visited on Mar. 29, 2021).

[140] Pooya Rezaei, Jeff Frolik, and Paul D.H. Hines. “Packetized plug-in electric vehicle charge management”. In: *IEEE Transactions on Smart Grid* 5.2 (Mar. 2014), pages 642–650.

[141] Hannah Ritchie, Max Roser, and Pablo Rosado. *CO_2 and Greenhouse Gas Emissions*. 2020. URL: https://ourworldindata.org/emissions-by-sector (visited on Mar. 2, 2021).

[142] Jose Rivera, Christoph Goebel, and Hans-Arno Jacobsen. "A distributed anytime algorithm for real-time EV charging congestion control". In: *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*. e-Energy '15. Bangalore, India: ACM, June 2015, pages 67–76.

[143] Daniel Römer and Johannes Salzgeber. *Verkehrswende in Deutschland braucht differenzierte Ansätze in Stadt und Land*. Technical report 363. KFW Research, Jan. 2022. URL: https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-2022/Fokus-Nr.-363-Januar-2022-Verkehrswende.pdf (visited on June 20, 2022).

[144] Roman Rudnik et al. "Real-time control of an electric vehicle charging station while tracking an aggregated power setpoint". In: *IEEE Transactions on Industry Applications* 56.5 (Oct. 2020), pages 5750–5761.

[145] Aminmohammad Saberian, Hashim Hizam, Mohd Amran Mohd Radzi, Zainal Kadir, and Maryam Mirzaei. "Modelling and prediction of photovoltaic power output using artificial neural networks". In: *International Journal of Photoenergy* 2014 (Apr. 2014). Article ID 469701.

[146] Mushfiqur R. Sarker, Miguel A. Ortega-Vazquez, and Daniel S. Kirschen. "Optimal coordination and scheduling of demand response via monetary incentives". In: *IEEE Transactions on Smart Grid* 6.3 (May 2015), pages 1341–1352.

[147] Benjamin Schaden, Thomas Jatschka, Steffen Limmer, and Günther Robert Raidl. "Smart charging of electric vehicles considering SoC-dependent maximum charging powers". In: *Energies* 14.22 (Nov. 2021). Article 7755.

[148] Jonas Schlund. "Electric vehicle charging flexibility for ancillary services in the german electrical power system". PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, Nov. 2021. URL: https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/17600 (visited on Feb. 27, 2023).

[149] Jonas Schlund, Marco Pruckner, and Reinhard German. "Flexability - Modeling and maximizing the bidirectional flexibility availability of unidirectional charging of large pools of electric vehicles". In: *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*. e-Energy '20. Virtual Event, Australia: Association for Computing Machinery, June 2020, pages 121–132.

[150] Rico Schrage, Paul Hendrik Tiemann, and Astrid Niesse. "A multi-criteria metaheuristic algorithm for distributed optimization of electric energy storage". In: *ACM SIGEnergy Energy Informatics Review* 2.4 (Dec. 2022), pages 44–59.

[151] Benyun Shi and Jiming Liu. "Decentralized control and fair load-shedding compensations to prevent cascading failures in a smart grid". In: *International Journal of Electrical Power & Energy Systems* 67 (May 2015), pages 582–590.

[152] M. Shreedhar and George Varghese. “Efficient fair queuing using deficit round-robin”. In: *IEEE/ACM Transactions on Networking* 4.3 (June 1996), pages 375–385.

[153] Laurynas Šikšnys and Torben Bach Pedersen. “Dependency-based FlexOffers: Scalable management of flexible loads with dependencies”. In: *Proceedings of the Seventh International Conference on Future Energy Systems*. e-Energy ’16. Article 11. Waterloo, Ontario, Canada: Association for Computing Machinery, June 2016.

[154] Laurynas Šikšnys, Emmanouil Valsomatzis, Katja Hose, and Torben Bach Pedersen. “Aggregating and disaggregating flexibility objects”. In: *IEEE Transactions on Knowledge and Data Engineering* 27.11 (Nov. 2015), pages 2893–2906.

[155] P. Silva et al. “Two-stage optimal operation of smart homes participating in competitive electricity markets”. In: *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I CPS Europe)*. Bari, Italy: IEEE, Sept. 2021, pages 1–6.

[156] Åse Lekang Sørensen, Karen Byskov Lindberg, Igor Sartori, and Inger Andresen. “Residential electric vehicle charging datasets from apartment buildings”. In: *Data in Brief* 36 (Apr. 2021). Article 107105.

[157] James C. Spall. “Introduction to stochastic search and optimization”. John Wiley & Sons, Inc, Mar. 2003.

[158] Rafal Stankiewicz, Piotr Cholda, and Andrzej Jajszczyk. “QoX: What is it really?” In: *IEEE Communications Magazine* 49.4 (Apr. 2011), pages 148–158.

[159] State Grid Corporation of China and China Electricity Council. *White paper of Chaoji EV charging technology (technical part)*. White Paper. State Grid Corporation of China and China Electricity Council, June 2020. URL: `https://www.cec.org.cn/upload/1/editor/1594869131179.pdf` (visited on Aug. 1, 2022).

[160] Naomi Stringer, Navid Haghdadi, Anna Bruce, and Iain MacGill. “Fair consumer outcomes in the balance: Data driven analysis of distributed PV curtailment”. In: *Renewable Energy* 173 (Aug. 2021), pages 972–986.

[161] Dirk Sudholt. “The benefits of population diversity in evolutionary algorithms: A survey of rigorous runtime analyses”. In: *Theory of Evolutionary Computation: Recent Developments in Discrete Optimization*. Cham: Springer International Publishing, 2020, pages 359–404.

[162] Xianzhuo Sun, Jing Qiu, Yuechuan Tao, Yuan Ma, and Junhua Zhao. “Coordinated real-time voltage control in active distribution networks: An incentive-based fairness approach”. In: *IEEE Transactions on Smart Grid* 13.4 (July 2022), pages 2650–2663.

[163] Olle Sundstrom and Carl Binding. “Flexible charging optimization for electric vehicles considering distribution grid constraints”. In: *IEEE Transactions on Smart Grid* 3.1 (Mar. 2012), pages 26–37.

[164] Theodoros Theodoropoulos et al. “Proportionally fair and scalable EV charging under distribution line voltage constraints”. In: *Electric Power Systems Research* 208 (July 2022). Article 107797.

[165] Mahyar Tofighi-Milani, Sajjad Fattaheian-Dehkordi, Mahmud Fotuhi-Firuzabad, and Matti Lehtonen. “Decentralized active power management in multi-agent distribution systems considering congestion issue”. In: *IEEE Transactions on Smart Grid* 13.5 (Sept. 2022), pages 3582–3593.

[166] Georgios Tsaousoglou et al. “Managing distributed flexibility under uncertainty by combining deep learning with duality”. In: *IEEE Transactions on Sustainable Energy* 12.4 (Oct. 2021), pages 2195–2204.

[167] Emin Ucer, Mithat C. Kisacikoglu, Murat Yuksel, and Ali Cafer Gurbuz. “An Internet-inspired proportional fair EV charging control method”. In: *IEEE Systems Journal* 13.4 (Dec. 2019), pages 4292–4302.

[168] Andreas Ulbig and Göran Andersson. “Analyzing operational flexibility of electric power systems”. In: *International Journal of Electrical Power & Energy Systems* 72 (Nov. 2015). The Special Issue for 18th Power Systems Computation Conference, pages 155–164.

[169] Mark Voorneveld. “Characterization of Pareto dominance”. In: *Operations Research Letters* 31.1 (Jan. 2003), pages 7–11.

[170] Yanni Wan, Jiahu Qin, Fangyuan Li, Xinghuo Yu, and Yu Kang. “Game theoretic-based distributed charging strategy for PEVs in a smart charging station”. In: *IEEE Transactions on Smart Grid* 12.1 (Jan. 2021), pages 538–547.

[171] Jun Wang et al. “Queueing theory-based optimal decision-making model of battery energy storage-assisted fast charging station participating in emergency demand response”. In: *2020 IEEE Sustainable Power and Energy Conference (iSPEC)*. Chengdu, China: IEEE, Nov. 2020, pages 2110–2115.

[172] Minhan Yoon et al. “Utilization of energy storage system for frequency regulation in large-scale transmission system”. In: *Energies* 12.20 (Oct. 2019). Article 3898.

[173] Yung-Terng Wang and R. J. T. Morris. “Load sharing in distributed systems”. In: *IEEE Transactions on Computers* 34.3 (Mar. 1985), pages 204–217.

[174] Yunkai Zhou and Harish Sethu. “On the relationship between absolute and relative fairness bounds”. In: *IEEE Communications Letters* 6.1 (Jan. 2002), pages 37–39.

[175] Jan Zacharias. *Konkretisierung der Ampelkonzepts im Verteilungsnetz*. Technical report. Bundesverband der Energie- und Wasserwirtschaft e.V., Feb. 2017. URL: `https://www.bdew.de/media/documents/20170210_Konkretisierung-Ampelkonzept-Smart-Grids.pdf` (visited on Mar. 11, 2021).

[176] Ioannis Zenginis, John S. Vardakas, Nizar Zorba, and Christos V. Verikoukis. “Analysis and quality of service evaluation of a fast charging station for electric vehicles”. In: *Energy* 112.1 (Oct. 2016), pages 669–678.

[177] Jie Zhang et al. “Deadline-based V2V charging under spatial resource constraints”. In: *2021 International Conference on Control, Automation and Information Sciences (ICCAIS)*. Xi’an, China: IEEE, Oct. 2021, pages 240–245.

[178] Jinye Zhao, Tongxin Zheng, and Eugene Litvinov. “A unified framework for defining and measuring flexibility in power system”. In: *IEEE Transactions on Power Systems* 31.1 (Jan. 2016), pages 339–347.

[179] Yingjie Zhou, Nicholas Maxemchuk, Xiangying Qian, and Yasser Mohammed. “A weighted fair queuing algorithm for charging electric vehicles on a smart grid”. In: *2013 IEEE Online Conference on Green Communications (OnlineGreenComm)*. Piscataway, NJ, USA: IEEE, Oct. 2013, pages 132–136.

[180] Yingjie Zhou, Nicholas Maxemchuk, Xiangying Qian, and Chen Wang. “The fair distribution of power to electric vehicles: An alternative to pricing”. In: *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. Venice, Italy: IEEE, Nov. 2014, pages 686–691.