Learning-based Distributed Resource Flexibility Estimation for Enhanced TSO-DSO Coordination

Chrysostomou, D.

DOI
10.4233/uuid:6543ab05-deed-4b9e-bd31-fba18c0e4d82

Publication date
2026

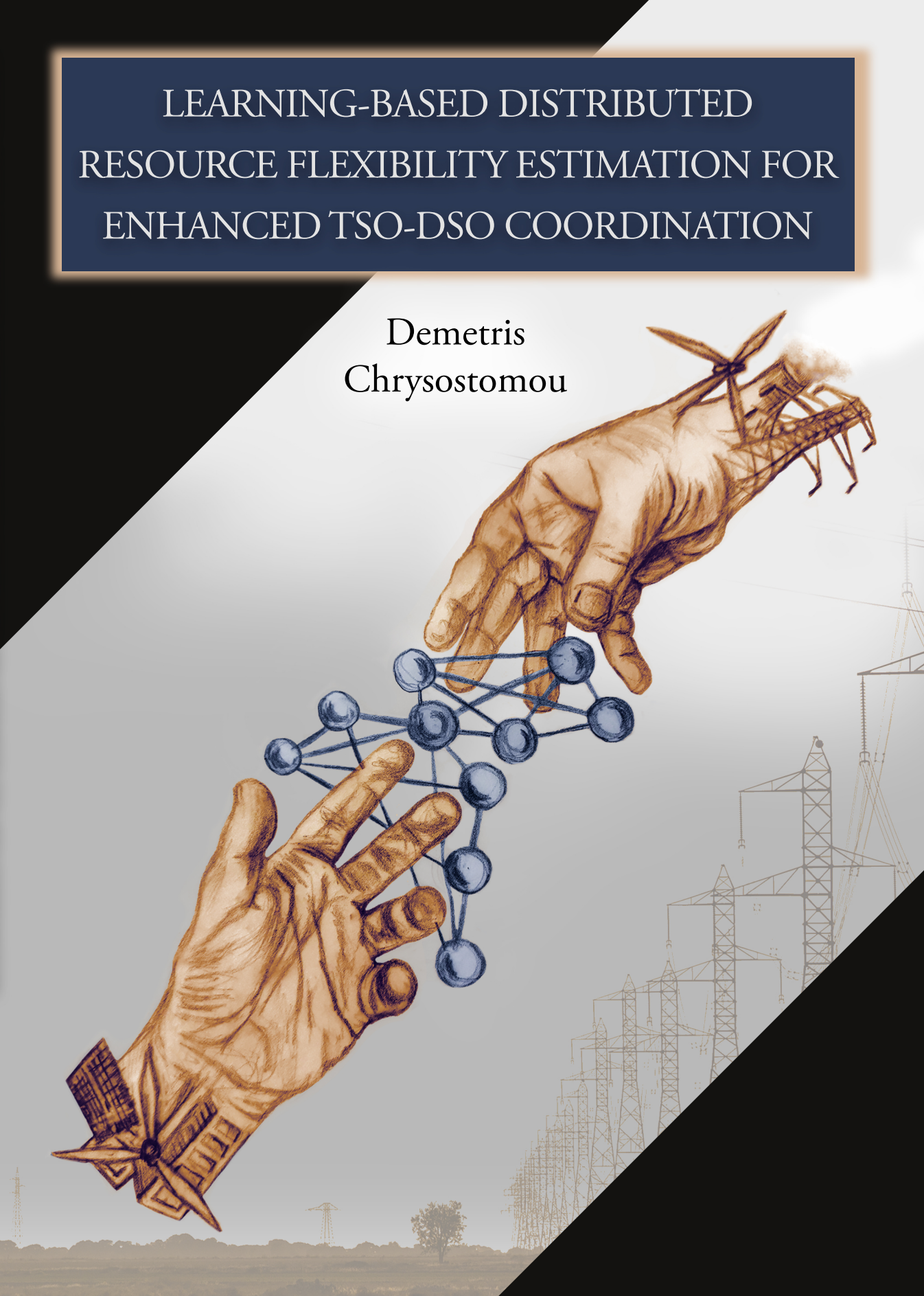
Document Version
Final published version

Citation (APA)
Chrysostomou, D. (2026). *Learning-based Distributed Resource Flexibility Estimation for Enhanced TSO-DSO Coordination*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:6543ab05-deed-4b9e-bd31-fba18c0e4d82

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

LEARNING-BASED DISTRIBUTED RESOURCE FLEXIBILITY ESTIMATION FOR ENHANCED TSO-DSO COORDINATION

Demetris Chrysostomou



Learning-based Distributed Resource Flexibility Estimation For Enhanced TSO-DSO Coordination

Learning-based Distributed Resource Flexibility Estimation For Enhanced TSO-DSO Coordination

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl,
chair of the Board of Doctorates,
to be defended publicly on Friday, 27 February 2026 at 10:00 o'clock

by

Demetris Chrysostomou

Master of Science in Robotics, Systems and Control, ETH Zurich, Switzerland
born in Lefkosia, Cyprus.

This dissertation has been approved by the promotors:
Promotor: Dr. J.L. Rueda Torres
Promotor: Dr. J.L. Cremer
Promotor: Prof. ir. M. van der Meijden

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Dr. J.L. Rueda Torres	Delft University of Technology, promotor
Dr. J.L. Cremer	Delft University of Technology, promotor

Independent members:

Prof. dr. J.A. La Poutré	Delft University of Technology
Prof. dr. L.S. Reins	Erasmus University Rotterdam
Prof. dr. Z. Vale	Polytechnic of Porto
Prof. dr. S. Chatzivasileiadis	Technical University of Denmark
Prof. dr. J.K. Kok	Eindhoven University of Technology
Prof. dr. ir. A.H.M Smets	Delft University of Technology, reserve member

Keywords: TSO-DSO coordination, Distributed Controllable Flexibility, Convolution Operations, Recurrent Neural Networks, Long-Short Term Memory Networks, Bayesian Neural Networks
Printed by: Proefschrift Specialist
Cover Design: Demetris Chrysostomou

ISBN 978-94-6518-225-4

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Ancora imparo
(I am still learning).

Michaelangelo, at age 87

NOMENCLATURE

LIST OF ABBREVIATIONS

ADRR	Aggregated dynamic response representation.
AI	Artificial intelligence.
ARMA	Autoregressive moving average.
BESS	Battery energy storage systems
BNN	Bayesian neural network.
DB	Dinstance based.
DE	Deep ensembles.
DFC	Density of feasible combinations.
DS	Distribution system.
DSO	Distribution system operator.
DSSE	DS state estimation.
FA	Flexibility area.
FB	Feature based.
FNN	Feedforward neural network.
FSP	Flexibility service provider.
GFL	Grid following.
GFM	Grid forming.
GNN	Graph neural network.
GRU	Gated recurrent unit.
HV	High voltage
IBR	Inverter-based resources.
KL	Kullback–Leibler
LSTM	Long-short-term memory.

LV	Low voltage
MCD	Monte Carlo dropout.
ML	Machine learning.
MOO	Multi-objective optimization.
MSE	Mean squared error.
MV	Medium voltage
NLL	Negative log likelihood.
$OB0$	Oberrhein network of substation 0.
$OB1$	Oberrhein network of substation 1.
OB	Oberrhein network.
OC	Operating condition of the DS.
$OLTC$	On-load tap changer
OP	Operating point as p, q of the PCC.
OPF	Optimal power flow.
PCC	Point of common coupling.
PF	Power flow.
$PICP$	Prediction interval coverage probability.
PV	Photovoltaic
$RMSE$	Root mean squared error.
RNN	Recurrent neural network.
RT	Research Track.
RTI	Real-time interface.
TS	Transmission system.
TSO	Transmission system operator.
VI	Variational inference.
$w.r.t.$	with respect to
WT	Wind turbine

CONTENTS

SUMMARY

This research focuses on a critical challenge in modern power systems: the effective coordination between transmission system operators (TSOs) and distribution system operators (DSOs) to harness distributed flexibility for grid stability and efficiency. As energy systems undergo a fundamental transformation toward decentralized, decarbonized, and digitalized operations, the ability to accurately estimate and activate aggregated flexibility in distribution systems is becoming a cornerstone of secure and sustainable grid management.

First, this research proposed and validated an algorithm for aggregated distribution system flexibility estimation. This algorithm enables DSOs to effectively aggregate distribution system flexibility to characterize the set of feasible active and reactive power adjustments that can be offered to TSOs under operational constraints. This algorithm evaluates the feasibility of all possible flexibility combinations, estimating the density of feasible combinations as a reliability metric. This contribution provides reliability considerations in flexibility areas and provides a fast estimation approach for meshed and radial topologies, disjoint, convex, and non-convex flexibility areas.

Second, this research developed a flexibility area estimation approach considering the real-time uncertainty in distribution systems due to limited observability. This method utilizes probabilistic machine learning to enable TSOs to select a safety probability given the task uncertainty. This method enables adopting the flexibility area estimation in distribution systems with limited measurements, which can support TSOs in utilizing distributed flexibility. Thus, the resulting flexibility areas are robust to measurement errors.

Third, this research developed and released an open-source Python package that encapsulates functionalities for flexibility area estimation. This tool provides a programmatic interface for system operators, researchers, and third-party service providers to compute and visualize flexibility areas. The package is modular, extensible, promoting collaborative innovation in this domain.

Fourth, this doctoral research developed a TSO-DSO coordination framework for flexibility activation aimed at improving the overall dynamic performance due to the action of primary controls. This work recognized that inverter-based resources affect the post-disturbance response in power systems, and controlling these resources can impact the post-disturbance power system stability. Therefore, this doctoral research proposed a dynamic system-aware coordination scheme that leverages flexibility at the distribution level to avoid post-disturbance instabilities in the interconnected distribution-transmission system. Case studies on the Dutch Zealand system showed how this framework can indicate and mitigate instability risks for TSOs while utilizing existing measurement platforms. This framework does not require sensitive information exchange.

The implications of this work extend beyond technical contributions. The growing penetration of distributed energy resources, such as solar and electric vehicles, introduced

new operational paradigms that challenge the conventional unidirectional flow of control and information. TSOs and DSOs must now interact more frequently and effectively, especially as power systems transition from centralized to a more decentralized dispatch. This doctoral research provides and proposes algorithms and methodologies to enable such interactions.

On a broader societal level, this doctoral research supports the integration of higher shares of renewable energy by developing algorithms that appraise the flexibility of decentralized actors to support power system operation. It contributes to the ongoing decarbonization of the energy sector by making system balancing and stability more distributed and cost-effective. The work also promotes digital transparency and open innovation through the public release of software tools, allowing academic and industrial stakeholders to co-develop future extensions.

SAMENVATTING

Dit onderzoek richt zich op een cruciale uitdaging in moderne energiesystemen: de effectieve coördinatie tussen transmissiesysteembeheerders (TSOs) en distributiesysteembeheerders (DSOs) om gedistribueerde flexibiliteit te benutten voor netstabiliteit en efficiëntie. Nu energiesystemen een fundamentele transformatie ondergaan naar gedecentraliseerde, CO2-arme en gedigitaliseerde bedrijfsvoering, wordt het vermogen om geaggregeerde flexibiliteit in distributiesystemen nauwkeurig te schatten en te activeren een hoeksteen van veilig en duurzaam netbeheer.

Ten eerste stelt dit onderzoek een algoritme voor en valideert het voor het schatten van geaggregeerde flexibiliteit in distributiesystemen. Dit algoritme stelt DSO's in staat en zo de flexibiliteit van distributiesystemen effectief te aggregeren om de set van haalbare aanpassingen van het actieve en reactieve vermogen te karakteriseren die aan TSOs kunnen worden aangeboden onder operationele beperkingen. Dit algoritme evalueert de haalbaarheid van alle mogelijke flexibiliteitscombinaties en schat de dichtheid van haalbare combinaties als betrouwbaarheidsmaatstaf. Deze bijdrage behandelt betrouwbaarheidsoverwegingen van flexibiliteitsgebieden en biedt een snelle schattingsmethode voor vermaasde en radiale topologieën, disjuncte, convexe en niet-convexe flexibiliteitsgebieden.

Ten tweede heeft dit onderzoek een benadering ontwikkeld voor het schatten van flexibiliteitsgebieden, rekening houdend met de realtime onzekerheid in distributiesystemen als gevolg van beperkte observeerbaarheid. Deze methode maakt gebruik van probabilistisch machine learning om TSOs in staat te stellen een veiligheidskans te selecteren, gegeven de taakonzekerheid. Deze methode maakt het mogelijk om de schatting van flexibiliteitsgebieden toe te passen in distributiesystemen met beperkte metingen, wat TSOs kan ondersteunen bij het benutten van gedistribueerde flexibiliteit. De resulterende flexibiliteitsgebieden zijn daardoor bestand tegen meetfouten.

Ten derde heeft dit onderzoek een open-source Python-pakket ontwikkeld en uitgebracht dat functionaliteiten voor het schatten van flexibiliteitsgebieden bevat. Deze tool biedt een programmatische interface voor systeembeheerders, onderzoekers en externe dienstverleners om flexibiliteitsgebieden te berekenen en te visualiseren. Het pakket is modulair en uitbreidbaar en bevordert collaboratieve innovatie in dit domein.

Ten vierde ontwikkelde dit promotieonderzoek een TSO-DSO-coördinatiekader voor flexibiliteitsactivering, gericht op het verbeteren van de algehele dynamische prestaties als gevolg van de werking van primaire regelaars. In dit onderzoek werd vastgesteld dat invertergebaseerde bronnen de respons van elektriciteitssystemen na een verstoring beïnvloeden, en dat de besturing van deze bronnen de stabiliteit van het elektriciteitssysteem na verstoring kan beïnvloeden. Daarom stelde dit promotieonderzoek een dynamisch, systeembewust coördinatieschema voor dat flexibiliteit op distributieniveau benut om instabiliteit na verstoring in het onderling verbonden distributie- en transmissiesysteem te voorkomen. Casestudies op het Nederlandse Zeeland-systeem lieten zien hoe dit ka-

der instabiliteitsrisico's voor TSO's kan aangeven en beperken, terwijl gebruik wordt gemaakt van bestaande meetplatformen. Dit kader vereist geen uitwisseling van gevoelige informatie.
De implicaties van dit werk reiken verder dan technische bijdragen. De groeiende penetratie van gedistribueerde energiebronnen, zoals zonne-energie en elektrische voertuigen, heeft nieuwe operationele paradigma's geïntroduceerd die een uitdaging vormen voor de conventionele unidirectionele stromen van besturing en informatie. TSO's en DSO's moeten nu frequenter en effectiever met elkaar samenwerken, vooral nu elektriciteitssystemen de overgang maken van een gecentraliseerde naar een meer gedecentraliseerde distributie. Dit promotieonderzoek biedt en stelt algoritmen en methodologieën voor om dergelijke interacties mogelijk te maken.
Op breder maatschappelijk niveau ondersteunt dit promotieonderzoek de integratie van een groter aandeel hernieuwbare energie door algoritmen te ontwikkelen die de flexibiliteit van decentrale actoren beoordelen ter ondersteuning van de werking van het elektriciteitsnet. Het draagt bij aan de voortgaande decarbonisatie van de energiesector door de balancering en stabiliteit van het systeem meer gedistribueerd en kostenefficiënt te maken. Het werk bevordert ook digitale transparantie en open innovatie door de openbare vrijgave van softwaretools, waardoor academische en industriële belanghebbenden samen toekomstige uitbreidingen kunnen ontwikkelen.

1

INTRODUCTION

Demetris CHRYSOSTOMOU

A system is not the sum of its parts, but the product of their interactions.

Russell L. Ackoff

This chapter introduces the need to estimate the flexibility available in distribution systems and to coordinate between TSO and DSO, which are key aspects of this dissertation. This chapter explains the scientific and social urgency for this coordination and flexibility estimation, reflections from the related literature, and a list of scientific gaps. This chapter includes the doctoral research scope and questions, the overall methodology of this dissertation, and concludes with the outline for the following sections.

Parts of this chapter have been published in: D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, "Exploring Operational Flexibility of Active Distribution Networks with Low Observability", 2023 IEEE Belgrade PowerTech, 2023. DOI: 10.1109/PowerTech55446.2023.10202841. [1].

Parts of this chapter have been published in: D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, "Exploring Operational Flexibility of Active Distribution Networks with Low Observability", ArXiv. DOI: https://doi.org/10.48550/arXiv.2304.04192

1

Power system operation includes transmission systems (TSs), operated by the transmission system operators (TSOs), and distribution systems (DSs), operated by the distribution system operators (DSOs). One key difference between the transmission and distribution systems is that the TS operates at higher voltage levels than the DS. Before the rise of decentralized or distributed generation (DG), the TSs were responsible for transmitting energy from the large conventional power generation sites to the distribution system substations. The distribution systems would then distribute the energy from the substation to the consumers. Hence, the power flow between the transmission and distribution systems was unidirectional [2]. In the absence of DG, a TSO would utilize statistical or persistence models of the expected load consumption to operate and plan its generation, to ensure the demand-generation balance. In case of forecast errors or unexpected events such as the loss of a generator, the TSO would resolve the issues by applying control actions and utilizing its generation capabilities. However, the renewable energy sources (RES) and DG growth altered the dynamics between the DS and the TS and invoked the possibility of a bidirectional power flow between them [2].

The most prominent renewable sources of energy depend on uncontrollable weather conditions, such as wind speed or solar radiation. Thus, researchers pursued identifying the weather parameters affecting the efficiency of installations [3],[4]. Simultaneously, forecasting the operating points of renewable sources of energy became a great challenge for power engineering researchers [5], [6]. Prolific methods for wind generation forecasting, have been based on artificial neural networks (ANNs) [6]. Fig.1.1(a) shows the total capacity of different energy resources in the Netherlands each year using data from the Entso-E transparency platform [7]. The total photovoltaic and wind capacity has increased greatly over the last 10 years, whereas conventional fossil-based generation capacity has not expanded but rather decreased slightly. These data highlight the shift from conventional to renewable, mostly-decentralized generation. In addition, Fig.1.1(b), Fig.1.1(c), Fig.1.1(d) show the total fossil, PV and wind generation in the Netherlands, for each quarter of an average day for each season in 2019 and 2024. The data obtained by Entso-E [7] show a large variability between different days of the same season, highlighting the uncertainty and variability that TSOs and DSOs must endure for a reliable operation. The total PV and wind generation increased in the 5 year margin, which also increased the variability in the daily contribution by PV, wind, and fossil generation. The total values for PV-generation in Fig.1.1(c) only include the PV generation identified by the Dutch TSO, TenneT, with the majority of PV generation data being unavailable.

The rise of distributed generation has reduced the controllability and capabilities that the TSO would have to resolve forecast issues and unexpected events. Simultaneously, renewable sources of energy increased the fluctuations within the power grid. Therefore, the need for fast response capabilities from the TSO and the DSOs has increased. The increased uncontrollability, unpredictability, unobservability, and variability of renewable energy sources have made cooperation and coordination between distribution and transmission system operators a necessity [8]. This doctoral research employs data-driven approaches and machine learning models to help coordinate the TSOs and DSOs to utilize the distribution system flexibility for stable operation.

This doctoral research considers the DER support for steady-state operation and time-

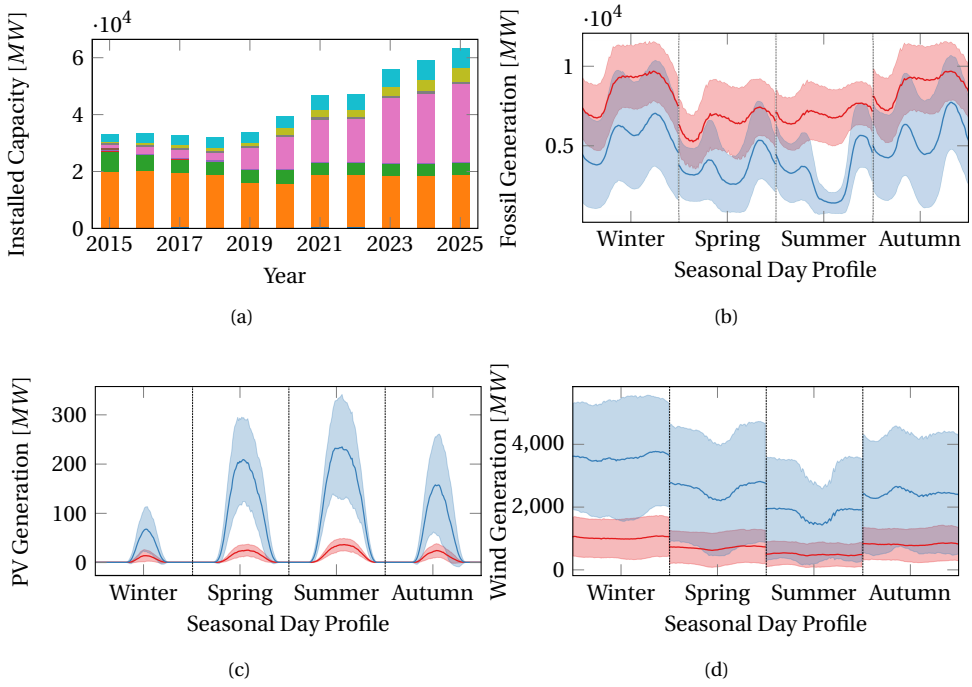


Figure 1.1: In (a) the installed production capacity in the Netherlands per year for biomass (), fossil gas (), fossil coal (), hydro (), nuclear (), other (), solar (), waste (), offshore wind (), and onshore wind () energy. The 2024 () and 2019 () Netherlands total (b) fossil, (c) PV, (d) wind generation per quarter of a day for each season.

varying disturbances. DER support for congestion management, or voltage control, would rather require steady-state simulations to consider whether the support respects the DS voltage and loading constraints. Time-varying disturbances, such as faults, would require considering the network dynamics to evaluate the post-disturbance system response for the available DER support.

1.1. MEGAMIND

The MEGAMIND program, short for MEasuring, GAthering, and MINing Data for Self-management in the Edge of the Electricity System, is a cross-disciplinary initiative funded by the Dutch Research Council (NWO) under its Perspectief scheme. MegaMind is led by Prof. Koen Kok at TU Eindhoven and supported by TU Twente, TU Delft, Tilburg, TNO, and nine industry partners: TenneT, Liander, Stedin, Enexis, PWC, IBM, Transdev, Smart State Technology, and Equans.

MegaMind aims to enhance local electricity systems by integrating AI and addressing regulatory barriers. MEGAMIND focuses on the edge of the power system, the distribution systems, and devices that connect to them. The project leverages distributed AI to monitor the network efficiently, predict potential overloads or imbalances, and enable

smart, autonomous coordination of devices to maintain grid stability.
A core pillar of MEGAMIND is ensuring that innovation in legislation and regulation keeps pace. Researchers in energy law are crafting regulatory frameworks for transparent, equitable data and energy sharing, preserving privacy, competition, and public values while enabling digital innovations.
The program is structured around 10 research tracks, one for each researcher. Furthermore, the program includes three work packages:

1. Grid-edge decision-making mechanisms and architectures.

2. Distributed multi-stakeholder energy operation technologies.

3. Data sharing and data governance methodologies for sustainability.

TU Delft participated in the MegaMind project through the research track (RT) 10: Stability of DSO-TSO operations. The researcher for RT10 is Demetris Chrysostomou, with supervision team from Prof. Ir. M.A.M.M. van der Meijden, (TenneT and TU Delft), Dr. ir. J.L.Rueda Torres (TU Delft), and Dr. ir. Jochen. L. Cremer. User partners mainly involved with RT10 include TenneT, Enexis, IBM, Stedin, and Liander. RT10 spans all work packages.
Throughout the MegaMind project, the RT10 team maintained frequent discussions and updates with TenneT to adapt the research and consider directions impactful in real-world applications. Furthermore, discussions were also regularly performed with Liander and Stedin to consider realistic scenarios and guide the developed algorithms to real-world issues.
RT10 research activities include this Ph.D. thesis, participation in knowledge and awareness sessions, panel participation, posters and presentations, industry workshops, and 3 deliverables:

1. Report on “Future scenarios and research boundaries”. Delivery month: October 2023.

2. Generic power system model in Digsilent PowerFactory, for simulation of interconnected future-like DSO-TSO power networks. Delivery month: October 2023.

3. Report on “Frequency stability performance of future integrated DSO-TSO systems”. Delivery month: October 2025.

1.2. TSO-DSO COORDINATION

TSO-DSO coordination becomes significant to allow DSOs and TSOs to facilitate and utilize flexibility from DS to ensure the resilient operation of the overall system. A recent report on TSO-DSO challenges and opportunities [9] clusters these challenges in (i) customers, business, market, data and information exchange, (ii) system planning, future flexibility, and asset lifecycle, (iii) system operations, dynamics, and control rooms of the future. For the first cluster [9] recognized the observability and effective control over new interconnected devices at different voltage levels as a key mid-term challenge. For short-term challenges, one focus is on increasing the TSO-DSO coordination and

cooperation to reduce the costs of administrative burdens and aligning processes. For the second cluster, [9] short-term challenges include extensive grid modeling to ensure effective analysis of impacts and integrating automatic RES control and intelligent load shedding. For cluster (ii), a main long-term challenge is improving resilience in extreme conditions, with mitigating directions including increasing monitoring for early identification and redundancy planning. For the third cluster, [9] key short-term challenges include monitoring ancillary services, enhancing active system management, establishing control governance, and the lack of proper coordination mechanisms, resulting in harmful effects. Key long-term challenge includes the system's resilience through shared risk assessment, which requires risk assessment models, data-sharing, and coordination between stakeholders.

The review in [10] classifies TSO-DSO coordination techniques in TSO-managed, DSO-managed, and hybrid-managed TSO-DSO models. These techniques focus on supporting the TSO to utilize flexibility from DS DER and consumers, but also to avoid potential conflicts with the DSO. In the TSO-managed model, TSOs are responsible for the DER dispatch, but TSOs also need to consider the DS constraints. TSO-managed approaches can have high computational and modelling challenges for TSO, and require a large data transfer of DS operating conditions from DSOs to the TSOs [10]. In the hybrid approach, TSOs are responsible for the DER dispatch while DSOs validate the TSO bids to consider DS constraints. Hybrid approaches can have less computational requirements and data transfer requirements, but complicate the coordination process. In the DSO-managed approach, TSOs consider the system dispatch while DSOs validate and consider the DER dispatch, facilitating distribution-level markets [10]. DSO-managed approaches do not need operational data transfer and limit the modelling requirements, but can still have computational challenges for DSOs. Between the coordination approaches [10] found the DSO-managed as the most commonly implemented. For DSO-managed coordination, flexibility areas (FA) enable DSOs to aggregate and visualize the total flexibility available in DS at the TSO-DSO interconnection, considering the DS constraints.

1.3. DISTRIBUTION SYSTEM OBSERVABILITY

The expected high impact of DG intrigued researchers to analyze their modeling within the DS and their effects on system power quality, safety, and reliability [11]. Therefore, the complexity of accurately modeling a DS with high DG penetration posed a significant challenge to researchers and system operators [12]–[14].

To deal with uncertainties in DS, researchers have focused on areas such as distribution network state estimation (DSSE), topology estimation, and parameter estimation. DSSE methodologies typically differ from the conventional transmission system state estimation methods due to some inherent differences between transmission and DS [14]. Those differences include the lower line reactance to resistance (X/R) ratio in DS, the larger number of nodes present in DS, and the higher uncertainty of system parameters due to inaccurate data or measurement devices. The non-observability of DGs provides further uncertainties to DSOs and TSOs, deteriorating their capability of accurately modeling the DSs. Therefore, as illustrated in [15], many researchers and EU projects on the DSO-TSO cooperation emphasize in increasing the grid observability and controllability.

1.4. AGGREGATED DISTRIBUTION SYSTEM FLEXIBILITY

Flexibility in power systems is a term used from various directions. These directions include market design [16], a specific resource type [17], or the aggregated DS flexibility offered at the TSO-DSO interconnection [18]–[20]. This doctoral research focuses on the DS flexibility at the TSO-DSO interconnection, in the context of TSO-DSO coordination. The objective of existing algorithms that estimate DS flexibility is to explore the limits of active and reactive power on the TSO-DSO boundary nodes. Flexible and distributed energy sources providing this flexibility are devices that can alter their operation to help the network avoid technical issues [19]. The flexibility of moving the current operating point to a new one is typically illustrated in a two-dimensional plot, the FA. This plot shows the active and reactive power on the TSO-DSO boundary nodes, as shown in Fig. 1.2. The red circle illustrates the initial operating point. The orange area shows the feasible apparent power values in which the operating point can shift using the flexibility of DS sources. The blue area illustrates the values in which the operating point shift through flexibility service providers (FSPs) could breach the DS constraints.

SO should know the available range of flexibility to timely detect in which directions and how much they can shift the TSO-DSO interconnection's operating point to ensure the system's stability and operation. Therefore, a higher range of flexibility can instigate a larger space for the SO to alter the power change between the TS and DS. In addition to the flexibility range, SO should know the population of each flexible point. More populated FA points indicate more feasible options for flexibility shift combinations. Therefore, SOs could have more options for FSP setpoints to reach each point, optimize costs, and maximize reliability considerations.

FA estimation methods primarily rely on the DS operating conditions (OC) as input data [19], [21]–[26]. However, DSs often suffer from limited real-time observability, as measurement units are only deployed on a subset of network components [1], [27]–[29]. Some methods have incorporated forecasting uncertainty into day-ahead FA estimations [18], [20], [30]. However, the issue of limited observability remains unaddressed in real-time FA estimation.

1.5. REPRESENTING DISTRIBUTION SYSTEM AGGREGATED DYNAMIC RESPONSE

The increase of DER in DS, and inverter based resources alters the impact of DS in the overall power system dynamics [31], [32]. Therefore, TSOs should not further rely on static load models to represent DSs [33], [34] but rather consider the DS response to dynamic events when evaluating the overall system stability. However, TSOs do not directly incorporate dynamic DS models in their dynamic studies, as:

1. The modeling complexity and maintenance for a system including DSs and TS can challenge TSOs [35].

2. DSs can have limited real-time observability and high DER uncertainty [14], [15].

3. Co-simulating transmission and DS dynamics can be computationally challenging [33], [34].

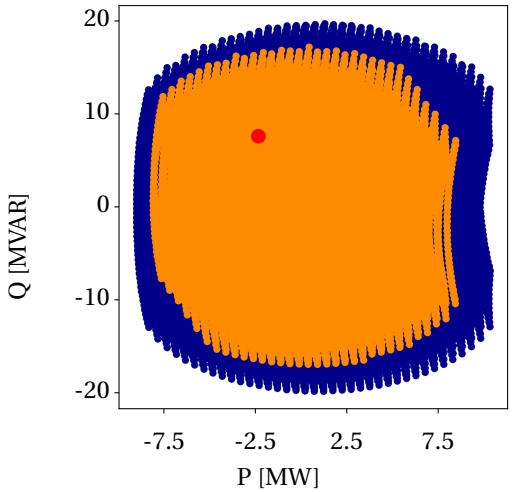


Figure 1.2: Example flexibility area of a load and a generator connected to the TSO-DSO interconnection in parallel. Feasible flexibility samples (■), infeasible flexibility samples (■), and initial operating point (■).

Aggregated dynamic response representation (ADRR) methods, also referenced as dynamic equivalent models, for DS aim to support dynamic studies for TSOs and DSOs by reducing the complexity of DS modeling and simulation. However, prior ADRR approaches generally assume fixed TS dynamic properties and do not enable control of grid-forming (GFM) inverter-based resource (IBR) outputs within the DS. This assumption of fixed, high TS inertia limits the ability to accurately capture DS responses, particularly when assessing the impact of controlling GFM IBR setpoints. Therefore, ADRR models should incorporate variable TS dynamic characteristics and support GFM IBR control within the DS.

1.6. SCIENTIFIC GAPS

The increasing number of DG in DS, and the low DS observability challenge the system operators' capability to maintain stable operation, utilize DS flexibility, and coordinate to consider the network's integrity. The low inertia in power systems due to the decrease in synchronous generation also challenges the dynamic stability in power systems. Thus, TSOs should consider the impact of DER setpoints on the system stability and control these DER setpoints to minimize instability risks.

Fig.1.3 visualizes the prior research on TSO-DSO coordination relating to the identified scientific gaps that this thesis aims to address. Prior research for DS flexibility in TSO-DSO coordination focused on aggregating the steady-state flexibility through flexibility areas. These flexibility areas can inform the TSOs on the feasible operating condition shifts. These feasible shifts can include static services such as congestion management or voltage support, and dynamic services such as automatic frequency restoration reserve (aFRR) or manual frequency restoration reserve (mFRR). However, prior research did not provide a TSO-DSO coordination method to visualize and control the DS dy-

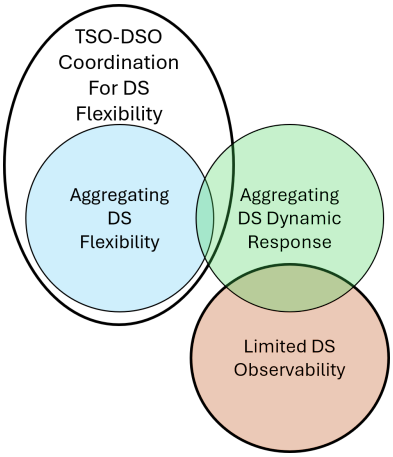


Figure 1.3: Prior research on TSO-DSO coordination relating to the identified scientific gaps.

namic response to large disturbances such as faults. Prior research on aggregated DS flexibility did not provide tractable aggregation algorithms to consider the FA reliability or feasibility density. Prior research in aggregated DS flexibility neglected the limited DS observability that can challenge the confidence and accuracy of the estimated FAs.
For steady-state consideration of DS flexibility, the main scientific gaps of related literature include:

1. While robustness and reliability are significant for system operators, considering reliability metrics for FAs can be intractable with prior FA estimation approaches.

2. FA estimation methods can be computationally expensive, whereas simplified, faster approaches have generalization issues for meshed DSs and cannot estimate cases where distinct, non-convex FAs exist.

3. Existing approaches for aggregated DS flexibility estimation in real-time assume full observability, challenging their adoption by system operators.

For ADRR, the main scientific gaps of related literature include:

1. ADRR approaches do not provide a coordination process enabling TSOs to control DS conditions, nor to evaluate the control actions' impact on the DS's post-disturbance response.

2. ADRR approaches do not consider grid-forming inverters, and their impact on post-disturbance stability.

1.7. RESEARCH OBJECTIVE AND QUESTIONS

Considering the scientific gaps of Sec.1.6, the objective of this doctoral research is "to develop approaches to estimate the steady-state and dynamic flexibility in distribution sys-

tems under limited observability to support secure TSO-DSO operation". This objective aims to address the scientific gaps of Fig.1.3, to result in the research directions of Fig.1.4; recognizing the connection between aggregating DS flexibility, limited DS observability, and aggregating the DS dynamic response in the context of TSO-DSO coordination to utilize DS flexibility.

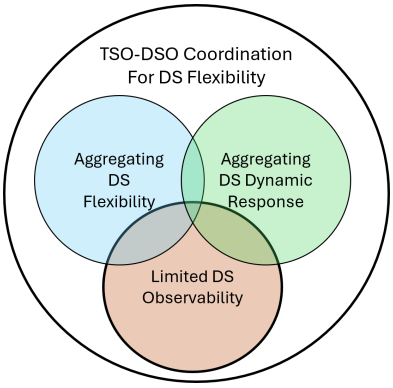


Figure 1.4: Proposed research on TSO-DSO coordination to address the identified scientific gaps.

1.7.1. Q1: HOW CAN FLEXIBILITY AREA ESTIMATION ALGORITHMS INCLUDE RELIABILITY METRICS WITH COMPUTATIONAL EFFICIENCY?

By answering this research question, the impact of FA algorithms for system operators can be improved, as robustness and resilience are important for TSOs and DSOs to ensure grid operational stability. The dependency on specific flexibility combinations and operating conditions close to the constraint thresholds can challenge DSOs and TSOs. Dependency on limited flexibility combinations could mean that the TSOs are vulnerable to any unreliability of FSPs (e.g., failure to achieve the requested setpoints). In addition, such dependency could also mean that TSOs need to suffer any flexibility costs requested by the FSPs in the limited feasible combinations. FA setpoints close to non-feasible conditions can also challenge DSOs, as any issue of unexpected change or over-/under-shoot for the FSP setpoints could lead the DS to non-feasible operation. Therefore, considering reliability metrics in the FA can improve the impact and adoption of FA estimation algorithms for TSOs and DSOs.

This research question represents reconsidering the assumptions and simplifications of prior approaches. Prior approaches, in the process of relieving computational burden, focus on exploring the extreme limits of FAs. However, FA operating points that can be reached with more feasible flexibility options are more reliable, e.g., system operators can select among any of the feasible combinations, without relying on limited options. However, evaluating the feasibility of all possible flexibility combinations rather than only the extreme feasible limits requires developing a more fundamental FA estimation approach rather than incrementally modifying prior approaches. To answer Q1, a key question derived is "how can the feasibility of each flexibility combination be evaluated with computational tractability?"

1.7.2. Q2: HOW CAN FLEXIBILITY AREA ESTIMATION ALGORITHMS GENERALIZE TO MESHED AND RADIAL NETWORK TOPOLOGIES, INCLUDING DISJOINT AND NON-CONVEX FLEXIBILITY AREAS?

By answering this research question, the applicability of FAs can be improved, which would enhance the adoption of FAs in power system operation. Limiting FA estimation approaches to specific network structures or types of flexibility resource can mitigate the appeal for their adoption, as system operators urge more robust and generally applicable algorithms in their operation.
This research question also represents reconsidering assumptions and simplifications in prior FA estimation approaches. Typical prior FA estimation algorithms focus on radial networks. However, non-radial DS topologies are present and emerging. Further, in the process of improving the computational speed or convergence properties of prior FA approaches, prominent FA estimation approaches fail to estimate disjoint and non-convex FAs. In answering Q2, more foundational questions arise:

1. How can the flexibility from multiple resources be aggregated, including non convexities?
2. How can flexibility estimation algorithms include flexibility resources with a non-continuous set of flexibility setpoints?

These foundational questions were identified after the relevant literature review, explored in later sections.

1.7.3. Q3: HOW CAN FLEXIBILITY AREA ESTIMATION INCORPORATE THE LEVEL OF OBSERVABILITY ON DISTRIBUTION SYSTEMS?

Answering this question can alleviate an important gap in realizing flexibility areas in power system operation. DSOs have limited measurements in DSs, which can limit the application of FA estimation algorithms that require the initial network state as input. In answering this research question, a more fundamental question arising is “What is the impact of limited DS observability on estimated FAs?”

1.7.4. Q4: WHAT IS THE IMPACT OF CONTROLLING INVERTER-BASED RESOURCES ON THE OVERALL SYSTEM’S DYNAMIC STABILITY?

The rise of IBRs alters the DS dynamics, and TSOs should consider the more complex DS responses when evaluating the overall system stability. However, examining the impact of control actions on IBRs on the overal system stability could potentially support TSOs in ensuring dynamic system stability. Therefore, this research question aims to examine if controlling DS IBRs can indeed be impactful for instability prevention.

1.7.5. Q5: HOW CAN TRANSMISSION AND DISTRIBUTION SYSTEM OPERATORS COORDINATE TO CONTROL DISTRIBUTION SYSTEM RESOURCES IN REAL-TIME TO ENSURE STABILITY?

By answering this research question, a method for TSO-DSO coordination can be developed to enable TSOs consider DS dynamic response, and directly control the DS resources. Such method could improve the dynamic stability of the overal power system,

considering the proliferation of DER and reduction of conventional energy resources. However, in answering Q5, further questions arise:

1. How can TSOs consider and evaluate the impact of inverter-based resources in DSs for dynamic simulations?
2. How can DSOs classify the types of inverter-based resource operations?
3. What existing real-time measurements can help the DS FA estimation and ADRR?

1.8. OVERALL METHODOLOGY

This doctoral research investigates interrelated approaches to improve coordination between TSOs and DSOs to enable the utilization of DS flexibility. These approaches address critical aspects of modern DS operation in the context of increasing IBR penetration, grid decentralization, and limited DS observability.

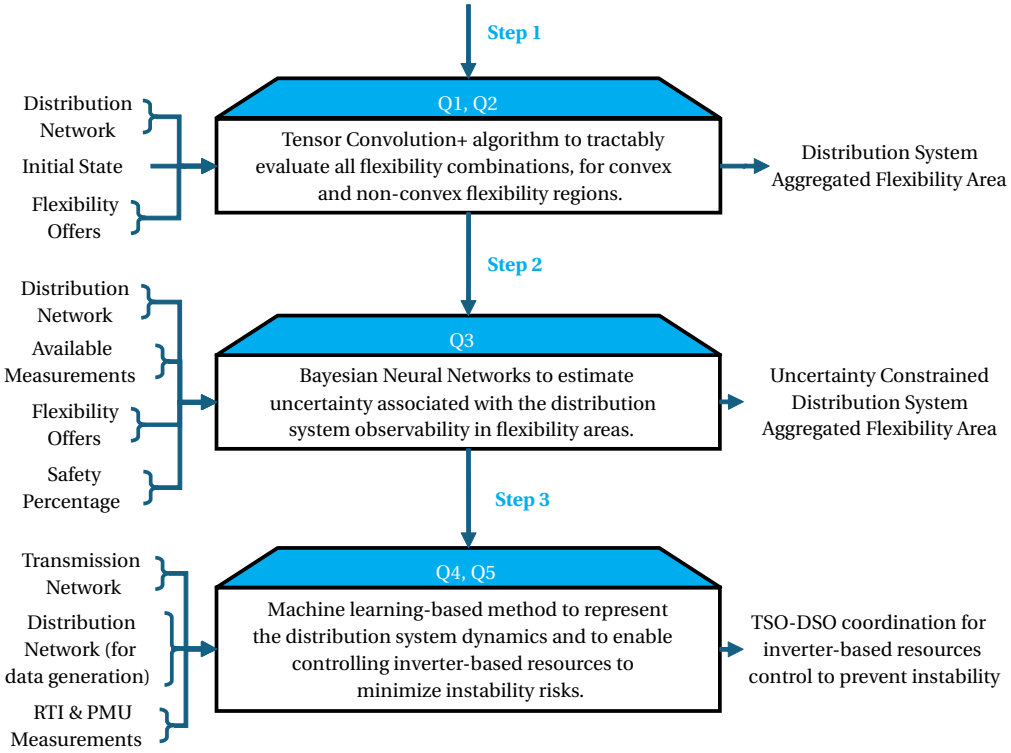


Figure 1.5: Methodology steps in addressing research questions.

Fig. 1.5 illustrates the methodology steps and developed approaches to address the research questions. The first approach develops an algorithm for the aggregation of DS flexibility for TSO-DSO coordination, with explicit consideration of computational efficiency and reliability metrics. The goal is to provide the TSO with tractable yet accu-

rate flexibility representations derived from the underlying distribution network constraints. The proposed algorithm utilizes convolution operations to aggregate flexibility areas, and tensor operations to evaluate the feasibility of each flexibility combination. The output of this algorithm returns not only the range of FA, but also the density of feasible combinations for each FA operating point. Hence, this algorithm enables the DSO to communicate actionable flexibility offers that the TSO can confidently use in system-wide decision-making. To increase the outreach of the developed algorithm and further enhance the adoption of FA estimation apporaches by researchers and industry professionals, a Python package for FA estimation was also developed and mad available through the Python package index (PyPi). The details and examples for the package are included in Appendix 1.
The second approach utilizes Bayesian neural networks to estimate the initial conditions of the DS constrained variables and approximate the aleatoric and epistemic uncertainties associated with the DS measurements. This approach develops an FA estimation method that enables system operators to select the safety margins in the estimated FA, considering uncertainties due to the limited DS observability.
The third approach explores the DS flexibility for dynamic stability, in the context of TSO-DSO coordination. The proposed coordination scheme focuses on enabling TSOs to control DS IBRs to minimize instability risks after potential disturbances. This approach embeds a DS ADRR model that TSOs can use to evaluate the DS response to dynamic events, and identify the impact of IBR setpoint control actions. This approach considers measurements available to DSOs that can further support the ADRR performance under varying DS operating conditions.

1.9. LIST OF SCIENTIFIC CONTRIBUTIONS

The scientific contributions pertaining to the objective and research questions are:

- Developed the TensorConvolution+ algorithm [36]. This algorithm addresses Q1, and Q2. It efficiently evaluates the feasibility of all possible FSP flexibility combinations and quantifies the density of feasible combinations as a reliability metric. This algorithm generalizes to meshed and radial DS topologies, and includes disjoint and non-convex FAs. Thus, Chapter 3 contributions include:

    1. Proposed the first method capable of tractably evaluating all discrete combinations of flexibility, thus enriching the characterization of FAs.
    2. Introduction of the convolution operation in flexibility estimation, leveraging convolution properties to reduce the complexity of the FA estimation problem.
    3. Using tensor structures in flexibility estimation to efficiently store and assess the feasibility of discrete flexibility combinations.

- Proposed an approach for FA estimation with BNNs to consider the estimation and DS observability uncertainties, addressing Q3. Chapter 4 analyses the performance of BNN approaches considering the impact of network sizes, assumptions about noise levels, and data distributions for the power system task. The proposed

FA approach also includes an algorithm for networks with 2 PCC. Therefore, Chapter 4 contributions include:

1. Developing an FA estimation approach considering the uncertainty for DS constraint margins due to limited DS observability and estimator uncertainty.
2. Developing BNN model structures to estimate FAs considering real-time uncertainties and data distributional changes.
3. Creating an approach approximating and representing FAs in networks with 2-PCCs.

- Proposed a TSO-DSO coordination method using an ADRR model to enable TSOs to select and evaluate flexibility setpoints for the DS IBR. This method addresses Q4 and Q5. Hence, Chapter 5 contributions include:

  1. Developing a TSO-DSO coordination method for dynamic stability, considering the aggregated dynamic response of DS estimated with measurements available to system operators.
  2. Creating an algorithm for risk-based evaluation and selection of TSO actions to prevent instability, within the proposed TSO-DSO coordination method.
  3. Developing a ML-based approach to aggregate and represent the DS dynamic response to dynamic events, considering for variable dynamic system properties, the aggregated DS power output, and the IBR type of primary control, combining sequential and non-sequential features.

- Released an open-source Python package for aggregated FA estimation in Appendix A. This package aims to support the adaptability of FA estimation algorithms and the reachability of TensorConvolution+.

1.10. OUTLINE OF THESIS

The thesis outline is:

- Chapter 2: Background on existing FA literature, ADRR models, and ML applications.
- Chapter 3: Proposes a new approach for FA estimation, evaluating the feasibility of all flexibility combinations, and including the density of feasible combinations.
- Chapter 4: Proposes using Bayesian neural networks to estimate and consider the uncertainty of the DS operating conditions when estimating FAs.
- Chapter 5: Proposes a new TSO-DSO coordination to control DS IBRs for dynamic stability using LSTM-based models.
- Chapter 6: Concludes on the overall research findings, addressing the research questions, and introducing future research directions.

1

Each chapter is associated with a famous quotation reflecting the chapter's topic or intention. The introduction chapter included the quotation "A system is not the sum of its parts, but the product of their interactions" from Russel L. Ackoff, which reflects the overall impact of TSO-DSO coordination, and the complexity in aggregating or evaluating DS flexibility to consider the feasibility and network stability. In addition, this quotation reflects the intent of the introductory section to explain the connection between the research directions as also shown in Fig.1.4. The background section is introduced with the quotation "The theory, hypothesis, framework, or background knowledge held by an investigator can strongly influence what is observed" from Norwood Russell Hanson. This quotation emphasizes the significance of background knowledge in comprehending the impact of the subsequent chapters. The third chapter is introduced with the quotation "Act always so as to increase the total number of choices" from Heinz von Foerster. This quotation highlights one of the main appeals of flexibility areas and the proposed TensorConvolution+ algorithm, i.e., to provide TSOs with information on the feasible flexibility options to optimize their system operation. The fourth chapter includes the quotation "What we observe is not nature itself, but nature exposed to our method of questioning" from Werner Heisenberg, emphasizing the importance of acknowledging the impact of limited DS observability for the FA estimation problem, and how including prediction uncertainty through BNNs can further represent the task uncertainty. The fifth chapter includes the quotation "An ounce of prevention is worth a pound of cure" from Benjamin Franklin, highlighting that it is better to take steps to prevent a problem from happening than to deal with the consequences after it occurs. This is reflected by the fifth chapter's proposed approach for post-fault instability prevention. Finally, the sixth chapter includes the quotation "If you cannot measure it, you cannot improve it" from Lord Kelvin, reiterating the main theme of this doctoral research, where knowledge and information on flexibility and DS conditions are crucial for the power system operation, i.e., TSOs can utilize DS flexibility if they are informed on its availability, feasibility, and characteristics.

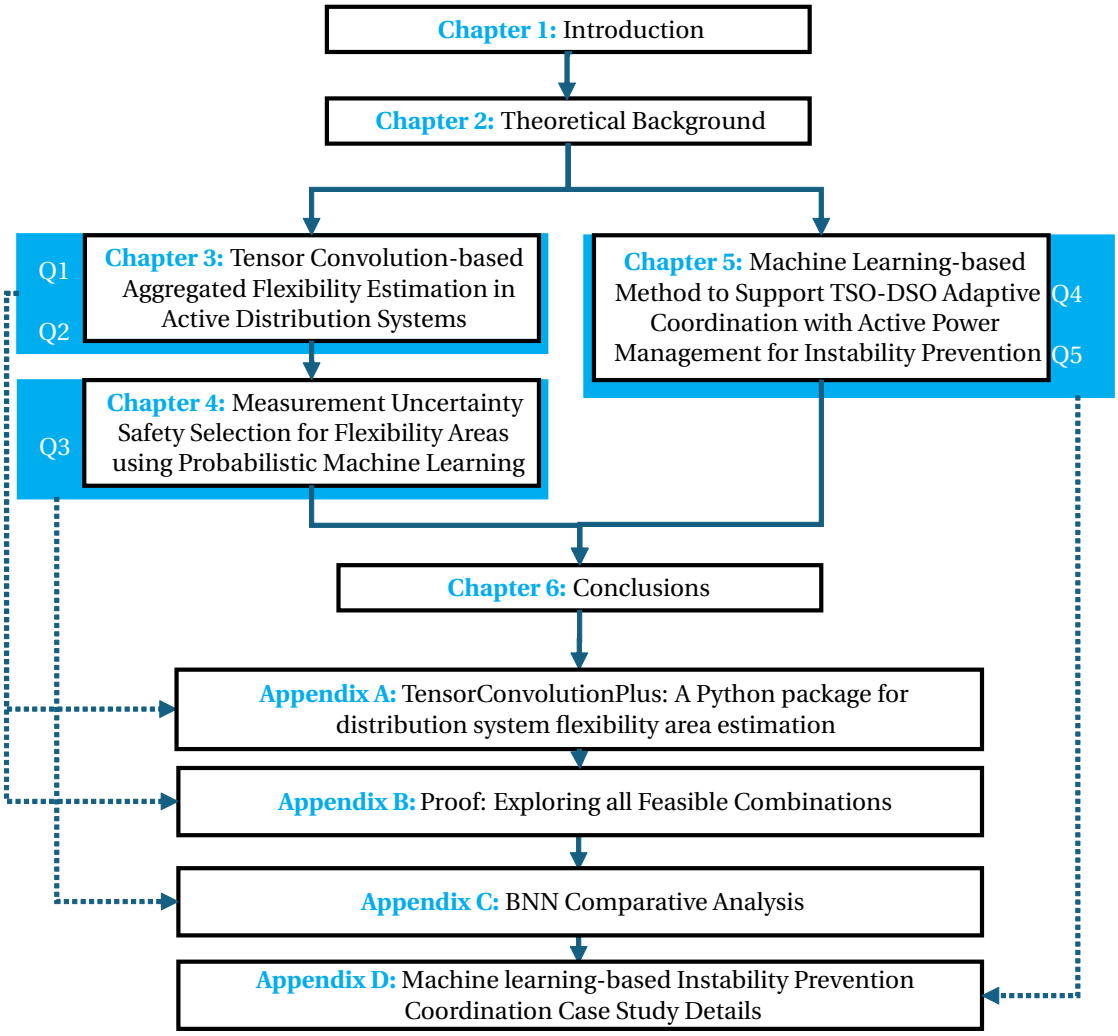


Figure 1.6: Outline of this thesis and connection with research questions.

2

BACKGROUND

The theory, hypothesis, framework, or background knowledge held by an investigator can strongly influence what is observed.

Norwood Russell Hanson

This section provides the necessary background on key topics relevant to this research. It begins by introducing the concept of distribution system flexibility areas, including definitions, prominent estimation methods, and the application of flexibility areas in TSO-DSO coordination. It then reviews the current practices in the power industry for utilizing DS flexibility, with a focus on Dutch frameworks. In addition, it examines distribution system aggregated dynamic response representation models, prominent approaches, and limitations. Finally, the section examines machine learning applications, with a focus on power systems tasks, including supervised and unsupervised methods, emphasizing both the opportunities and limitations of data-driven approaches.

Parts of this chapter have been published in: D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, "Exploring Operational Flexibility of Active Distribution Networks with Low Observability", 2023 IEEE Belgrade PowerTech, 2023. DOI: 10.1109/PowerTech55446.2023.10202841. [1].
Parts of this chapter have been published in: D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, "Exploring Operational Flexibility of Active Distribution Networks with Low Observability", ArXiv. DOI: https://doi.org/10.48550/arXiv.2304.04192

This chapter includes background on distribution system flexibility area estimation, industry applications to utilize distribution system flexibility, ADRR models and applications, and machine learning applications in power systems.

2.1. AGGREGATED DISTRIBUTION SYSTEM FLEXIBILITY AREAS

Flexibility can be defined in a variety of ways depending on the point of view, or purpose of usage. Table 2.1, includes the flexibility definitions within the relevant literature, which are typically non-analytical. This inconsistency of a definition for a common goal, and the consistency of non analytical definitions, provoked this research to create an alternative generalised definition for flexibility.

Existing algorithms approach the flexibility estimation problem as range exploration. The objective of existing algorithms that estimate this flexibility is to explore the limits of active and reactive power on the TSO-DSO boundary nodes. Flexible and distributed energy sources providing this flexibility are devices that can alter their operation to help the network avoid technical issues [19]. Prior FA estimation algorithms can mainly be categorized as power flow (PF)-based and optimal power flow (OPF)-based approaches. This section first introduces the proposed generalized flexibility area definition, and later describes the prominent PF-based and OPF-based approaches through that definition.

For the proposed definition, the following non-linear equation system is adopted:

$$\dot{x}(t) = f(x(t), u(t), v(t), w_x), \tag{2.1}$$

$$\tilde{x}(t) = r(x(t), \omega_o), \tag{2.2}$$

$$y(t) = h(\tilde{x}(t), u(t), \tilde{v}(t), w_y). \tag{2.3}$$

$t \in \mathbb{R}_+$ is the time component to include the time-variance of the system, $x(t) \in X \in \mathbb{R}^{2\times n}$ is the time-dependent state matrix representing the active and reactive power at each of the network's n nodes. Furthermore, $u(t) \in U$ is the action matrix representing the active, reactive power shifts for each controllable device. The output $y \in Y \subset \mathbb{R}^2$ is the TSO-DSO interconnection's observable active and reactive power flow. The matrices w_x, w_o, w_y are the system noise, the matrix $v(t) \in \mathbb{R}^{2\times n}$ includes the real and imaginary voltage values per network node, and $\tilde{v}(t)$ are the observable nodal voltages. Furthermore, $r(\cdot)$ is the function limiting the states observable by the SO (i.e., for a completely observable system $r(\cdot) = id(\cdot)$, the identity function). $h(\cdot)$ is the function for flexibility estimation, which uses the actions and observable states to estimate which y are reachable. The system is subjected to inequality technical constraints, e.g., due to allowable min-max limits for voltage magnitudes $c_v^{min} \leq g_v(x(t), v(t), u(t)) \leq c_v^{max}$, branch current flow $g_l(x(t), v(t), u(t)) \leq c_l^{max}$, and system stability $c_{DVR}^{min} \leq g_{DVR}(x(t), v(t), u(t)) \leq c_{DVR}^{max}$, where DVR is the dynamic variable time response. DVR represents the stability of the network when flexibility is activated, such as dynamic frequency excursions within the time frame of primary frequency control.

Through this representation, the set of equilibrium points X_e, where the power network is steady in the absence of an input and a disturbance, is defined as:

$$X_e = \{x_e | f(x_e(t), 0, v(t), 0) = 0\}. \tag{2.4}$$

Table 2.1: Flexibility definitions identified within relevant literature.

Flexibility definitions within literature
"The capability of a distribution network to adjust its active and reactive power flow at the TSO-DSO connection node by using optimization techniques" [21].
"The modification of generation injection and/or consumption patterns in reaction to an external signal (price signal or activation) in order to provide a service to the system."[19].
"Capability of active distribution networks to provide ancillary services (e.g. frequency control/power balancing, congestion management, or voltage support/security) at the TSO–DSO interface so as to aid the secure operation of the transmission network. .. TSO can activate this flexibility & notify the DSOs to change the setpoints accordingly."[22].
"A service, like active/reactive power reserves, that a resource provides to the grid by adjusting its operating point"[18].
"The ability of the system components to adjust their operating point, in timely and harmonized manner, to accommodate expected, as well as unexpected, changes in system operating conditions"[37].
"The response of a specific resource to an external signal (e.g., price signal or activation) through the modification of its injection and/or consumption pattern, thus providing a service to the system"[38].
"the ability of a system to deploy its resources to respond to changes in netload, where netload is defined as the remaining system load not served by variable generation"[39].
"Encompass controllability and observability information about the underlying power system which is constrained by the dynamics of its resources"[40]. "The ability of a power system to cope with variability and uncertainty in both generation and demand, while maintaining a satisfactory level of reliability at a reasonable cost, over different time horizons"[40]. "The system's capability to respond to a set of deviations that are identified by risk management criteria through deploying available control actions within predefined time-frame and cost thresholds"[40].
"flexibility refers to the capability of modifying energy usage schedules without violating operational constraints or compromising occupants' comfort"[17].

The definition of the flexibility range for the time window $[\tau_0, \tau_1]$, and the initial state $x(\tau_0)$, can be the set of all $y(\tau_1) \in Y$ in which the system can converge after a disturbance or an action, validating the network constraints. Disturbance can be an event (e.g. a fault) that will cause the FSPs to automatically respond through their controllers, and shift the state from the initial equilibrium point. This range is the orange area of the flexibility example in Fig. 1.2.
If assumed that there are either a finite amount of equilibrium points reachable from $x(\tau_0)$, or $y(t)$ is rounded to k decimals when represented or illustrated, then there is a possibility of multiple $y_1(\tau_1) = y_2(\tau_1)$ reachable from different actions $u_1(\cdot) \neq u_2(\cdot)$. Therefore, a set cannot include all reachable $y(t)$. Nevertheless, a multi-set can be theoretically defined to include the number of actions leading to each $y(t)$. A multiset is an element collection in which some elements may occur more than once [41], where Y is the set of distinct elements, and $m : Y \rightarrow \mathbb{Z}_+$ is the multiplicity, which defines the number of times each element of set Y exists.

2.1.1. PROPOSED DEFINITION

The flexibility of the system described by (2.1)–(2.3) on the time interval $[\tau_0, \tau_1]$ is the multiset of all $y(\tau_1) \in Y$, and $m_{y(\tau_1)} : Y \rightarrow \mathbb{Z}_+$ whose corresponding $x(\tau_1)$ is an equilibrium point, in which the system can converge after a disturbance or an action, validating the network constraints.
This definition generalizes to the dynamic response of power networks through (2.1), the time variability through the continuous and time-dependent (2.1)–(2.3), and low observability through the inclusion of $r(\cdot), \tilde{x}(t)$ in (2.2)–(2.3). Considering the disturbance term allows the response of active components, such as grid-forming inverters, to be included as flexibility resources, even without explicitly being shifted through a real-time activation action $u(t)$. The term $u(t)$ is the time-variable shifts of FSPs' power output, i.e., flexibility activation. Finally, this definition includes the range and multiplicity of flexibility by describing flexibility as a multiset instead of a set. Fig. 2.1 shows the flexibility when multiplicity m is also accounted for, i.e., the number of different actions $u(t)$ that reach the same feasible $y(t)$ as in (2.3). The projection into the PQ plane results in the range of flexibility as illustrated in Fig. 1.2. The multiplicity of flexible points at the boundaries of the PQ-plane projection is significantly lower than the centers. Hence, most actions leading to those points are non-feasible due to the network's technical constraints, or reaching those points is possible only through limited action combinations. On the other hand, the highly populated flexible points (with large m) near the center suggest multiple options for the SO.

FLEXIBILITY ESTIMATION OBJECTIVES AND CONSTRAINTS

The objectives for the estimation of the proposed multiset described flexibility are

$$\max\left(range(h \circ r)_{[\tau_0,\tau_1]}\right), \tag{2.5}$$

$$\max\left(m(y_i)\right) \quad \forall y_i \in range(h \circ r)_{[\tau_0,\tau_1]}, \tag{2.6}$$

here $range(\cdot)$ is the flexibility range. $h \circ r$ is the composition of flexibility estimation function h and observable state limitation function r [42]. The first objective (2.5) explores all possible active and reactive power shifts y reachable from an initial starting

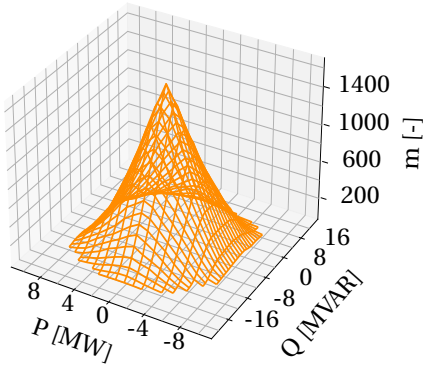


Figure 2.1: The flexibility area of Fig. 1.2 when considering m the multiplicity.

time τ_0 until the time τ_1. The second objective (2.6) explores the multiplicity m of each reachable shift. Thus, the number of action combinations that exist for each reachable shift. Considering $h(\cdot)$ and $r(\cdot)$ in the objectives highlights the influence of the flexibility estimation algorithm and the observable states on the result. Accordingly, improving the flexibility estimation could be through upgrading the estimation algorithms $h(\cdot)$, upgrading the network observability levels $r(\cdot)$, or both.
As shown in the generalized flexibility definition, the constraints accompanying the network are nodal under and over-voltage c_v^{min}, c_v^{max}, the branch loading limitations c_l^{max}, and the minimum and maximum system dynamic variable time response $c_{DVR}^{min}, c_{DVR}^{max}$. However, including additional constraints based on the needs of the flexibility estimating authority is feasible. As an example, a system operator who has a limitation on funds for flexibility activation can add a constraint $g_e(u) \leq c_e$. This additional constraint depends on the activation actions, which form costs depending on the shift amount, duration, and type of FSP.

2.1.2. POWER FLOW-BASED APPROACHES

Power flow-based approaches, starting from observable networks sample various shifts in the controllable flexibility sources. Using these shifts, they solve power flows to detect whether the new operating points would meet the constraints or not [20], [24]. Termination conditions for such algorithms can be the number of power flows run or time spent. Hence, the overall methodology followed by such algorithms has the following four steps:

1. Identify the shift capabilities U of each FSP and the initial operating state x of all network nodes.
2. Sample an operation shift for each FSP within the identified shift capabilities $u \in U$.
3. Run power flow to obtain the new TSO-DSO node's operating state y and check if it is feasible regarding network constraints.
4. Termination condition reached?

- If yes, plot flexibility area Y.
- If not, starting from the initial operating state, sample a different shift per FSP and go to step 3.

This methodology does not include the (2.2) since step 3 runs the power flow in the complete network, thus taking x as an input. The dynamics of (2.1) are ignored since these approaches are for steady state, and output Y does not include multiplicity.
The advantages of power flow-based algorithms include no need for linearization nor convexification of the objectives (2.5)–(2.6). Therefore, non-linear variables (e.g., on-load tap changers (OLTC)) do not affect their performance and can result in non-convex flexibility areas. Limitations of these algorithms include being significantly slower than OPF-based approaches [19], and their performance's dependence on the sampling distributions used for the FSP actions, i.e., which actions u are sampled and passed to (2.3) to find feasible points y.

2.1.3. OPTIMAL POWER FLOW-BASED APPROACHES

OPF-based approaches use multi-objective optimization (MOO) to compute the flexibility areas. In general terms, the four objectives that can theoretically be pursued are:

$$\min\left(P_{TSO-DSO}\right), \tag{2.7}$$
$$\max\left(P_{TSO-DSO}\right), \tag{2.8}$$
$$\min\left(Q_{TSO-DSO}\right), \tag{2.9}$$
$$\max\left(Q_{TSO-DSO}\right), \tag{2.10}$$

where $P_{TSO-DSO}, Q_{TSO-DSO}$ are the active and reactive power flowing from the TSO to the DSO, respectively. The approaches adapt the OPF constraints to include flexibility source operating point shifts.
Objectives (2.7)–(2.8) and (2.9)–(2.10) conflict with each other. Thus, these multi-objective optimization problems cannot combine all objectives in one. Therefore, OPF-based algorithms solve four MOOs for each FA $\min\left(P_{TSO-DSO}\right)$ and $\min\left(Q_{TSO-DSO}\right)$, $\min\left(P_{TSO-DSO}\right)$ and $\max\left(Q_{TSO-DSO}\right)$, $\max\left(P_{TSO-DSO}\right)$ and $\min\left(Q_{TSO-DSO}\right)$, and $\max\left(P_{TSO-DSO}\right)$ and $\max\left(Q_{TSO-DSO}\right)$. These MOOs are solved with approaches such as the ϵ constraint method [22], [43], weighted sum method [19], and radial reconstruction-based method [18], [23].
The difference between ε-constrained and weighted sum-based approaches is how they combine objectives. The ε-constrained considers only one of the objectives to be optimized and considers the other objectives as inequality constraints greater than ϵ. The ϵ is adjusted to approach different limits of the flexibility area. The weighted sum approach optimizes a single objective, the weighted summation of each objective of the initial MOO problem. The weights are adjusted to approach different flexibility areas. These two optimization-based approaches have the limitation that their resulting flexibility areas are convex hulls, which might not represent the actual case as explained in [19]. These approaches either perform linearization of non-linear control variables (OLTC) [19], [22] or do not include non-linear control variables [43]. Radial reconstruction iteratively considers an angle θ of active and reactive power proportions, and it finds

the active and reactive power limits along the line of θ [18]. Radial reconstruction-based approaches include both convex optimization [18] and non-convex optimization [23].
The advantages of OPF-based approaches are their speed and independency with FSP shift sampling distributions. A limitation of the above algorithms is not dealing with disjoint flexibility areas, which can be the case when FSPs include discrete variables [19]. OPF-based approaches ignore the dynamics of (2.1) and assess the steady-state performance. The OPF considers the complete network, thus ignoring (2.2). The output of these approaches aims only at objective (2.5).

2.1.4. ASSESSING THE EXISTING APPROACHES WITH THE PROPOSED DEFINITION

The approaches analyzed in the previous two subsections can be characterized based on the proposed definition in terms of low observability, multiplicity, and time continuity. The two approaches based on power flow and OPF need the initial network state $x(t)$ to explore the boundaries of the area. Therefore, the $h(\cdot)$ of (2.3) takes $x(t)$ as an input, and the low observability of (2.2) is neglected. Thus, the objectives (2.5) and (2.6) are altered into:

$$\max\left(range(h)_{[\tau_0,\tau_1]}\right), \tag{2.11}$$

$$\max\left(m(y_i)\right) \quad \forall y_i \in range(h)_{[\tau_0,\tau_1]}. \tag{2.12}$$

Regarding the objectives, OPF-based approaches' objectives (2.7)–(2.10) are aligned with the proposed objective of (2.5) since pushing the limits of feasible $P_{TSO-DSO}, Q_{TSO-DSO}$ maximizes the set of feasible y, and the range of h. However, OPF-based approaches do not deal with (2.6) since (2.7)–(2.10) do not include multiplicity. The power flow-based approaches can deal with the objectives (2.11) and (2.12) as they can sample actions leading to a similar flexibility point. However, power flow-based approaches sample actions using predefined distributions. Hence, the multiplicity reported by power-flow-based approaches can be faulty and biased toward these distributions. In addition, multiplicity results are not visualized nor referenced within the found literature.
The two approaches do not perform dynamic simulations in terms of state continuity and transients. The approaches also do not account for the response that inverter-based generators can have to external disturbances. Thus, the continuous state dynamics are discretized, and the system of (2.1)–(2.3) becomes:

$$x(\tau + \Delta\tau) = \hat{f}(x(\tau), u(\tau), v(\tau), \omega_x), \tag{2.13}$$

$$y(\tau) = h(x(\tau), u(\tau), v(\tau), \omega_y), \tag{2.14}$$

Furthermore, the action or shift capabilities of flexibility service providers are typically assumed not to be constrained by resource time characteristics. Hence, $\Delta\tau$ is assumed large enough that all machines can change their outputs to their limits. Thus, $u(\tau)$ is sometimes replaced by u.

2.1.5. FLEXIBILITY SERVICES

Evaluating the operating point shifting capabilities (without violating system constraints) of controllable devices within the distribution grid, can inform the transmission system

Table 2.2: Type of flexibility services referenced in relevant literature.

Ancillary services	
Type of service	Referenced within paper
Voltage regulation	[18], [20]–[24], [43], [44]
Frequency stability	[18], [20], [22], [23]
Power quality/support	[18], [20], [22], [43]
Congestion management	[18], [20], [22], [23], [44]
Generic terms	
Quality and security of supply	[18], [20], [22]
Balancing, stability and efficiency	[20], [21]
Deal with forecast uncertainty, react to sudden changes	[19]

operator which nodes can be used and how much active and reactive power can be requested to solve technical problems. These technical problems are frequency stability, voltage regulation, power quality, and congestion management.
The time horizon of the flexibility area utilization can be determined by its computational time and the inclusion of uncertainty and forecast data. Algorithms that take tens of minutes to compute the flexibility area, should not be useful for real-time operation [44]. On the other hand, not including the uncertainty when defining the load, or RES generation values, can be inaccurate on the planning time-frame [18]. The services provided by the flexibility area, can range through all time frames, from operational planning to real time operations [19].
The aforementioned literature typically specifies the types of ancillary services that can be offered through flexibility areas, or use some generic terms to explain the contribution that flexibility areas can have. Accordingly, Tab.2.2 includes services referenced within the literature. The research in this doctoral thesis considers the generic terms for FAs' contributing factors in power systems. This selection highlights that FA estimation algorithms primarily do not depend on the type of service for which the TSO would use the FSP flexibility.

2.2. INDUSTRY APPLICATIONS FOR DISTRIBUTION SYSTEM FLEXIBILITY

Report [8] outlines the operational challenges presented, or expected to be present by 2030, to the TSOs and DSOs. Those challenges include an electric system based on renewable sources, the development of real-time markets, decentralized generation, and new actors in the system. Furthermore, fast response to events or errors to ensure demand-generation balancing by the system operators is a significant challenge given the expected decrease of conventional generation sites on the transmission level.
In response to these challenges, system operators created flexibility platform to allow DERs to trade energy with the TSO and DSOs for economic returns. These flexibility platforms primarily focus on trade coordination, dispatching, and settlement of energy

or system services between system operators and the DERs [7]. Additionally, these flexibility platforms focus on flexibility that can help balance the electricity system and resolve constraints. Functions performed by such platforms include [7] asset registration and pre-qualification, notification of flexibility requirements and submission of offers, the priority of access, matching, price formation, issuing dispatching instructions and activation, as well as validation and settlement. [7] categorizes the operational models of flexibility platforms as administrative flexibility scheme coordinators, market intermediaries, and marketplaces. TSOs mainly use those platforms for short-term balance between active power supply and power demand maintenance, for balancing and congestion management. DSOs use those platforms mainly for congestion management. The challenges of these platforms are the DER integration, TSO-DSO coordination, and market design [7]. DER integration challenges include the prequalification requirements for DERs, such as minimum capacity thresholds. Even when DERs pass the requirements, the economic costs for them in order to participate in the markets might be high. Other revenue possibilities for DERs, and punishments when unable to deliver the flexibility they were asked to reserve, can disincentivize the DER operators from participating in the markets. Challenges for the TSO and DSO coordination include information exchange between the system operators. This information exchange becomes more exigent as the complexity, uncertainties, and unobservable or uncontrollable areas of distribution systems increase. Hence, obtaining information on distribution system constraints, and flexible devices is important. Notable platforms in industry for flexibility consideration include GOPACS, DA/RE, and RTI. These platforms support system operators to utilize distributed flexibility as response to events, but do not inform TSOs on the available DS feasible flexibility to support their operating decisions, i.e., are event-driven. In addition, these platforms currently only support grid congestion services. The FA algorithms in this thesis aim to inform TSOs on the available flexibility to support their operational decisions and do not depend on specific services.

GOPACS Platform Within the Netherlands, GOPACS (Grid Operators Platform for Congestion Solutions) acts as a market intermediary platform. This platform is owned by the Dutch grid operators. GOPACS supports the coordination of market-based procurement of congestion management services from energy markets within the Netherlands. Currently, flexibility providers offer their services through the ETPA market platform. ETPA allows short-term electricity trading in 15-minute blocks. The minimum flexibility service provider capacity for participation is 0.1 MW. Furthermore, the economic compensation provided by the operators to the flexibility providers is with regards to the activation costs in $€/MWh$. The operating time window of the GOPACS platform is within hours, i.e., a system operator can request flexibility until hours prior to the activation.
GOPACS allows the TSO and DSOs to issue a congestion notification to the markets. For each notification of a congestion, the system operator acquires at least two orders (eg, one to buy and one to sell) from the intraday markets. Flexibility providers (such as DER participating in the markets) can submit an order based on the congestion notification. Flexibility providers who are within the congested area can submit a buy or sell order. Providers outside the congested area can submit the opposite order than the order from the providers within the congested area, to balance the changes. The GOPACS

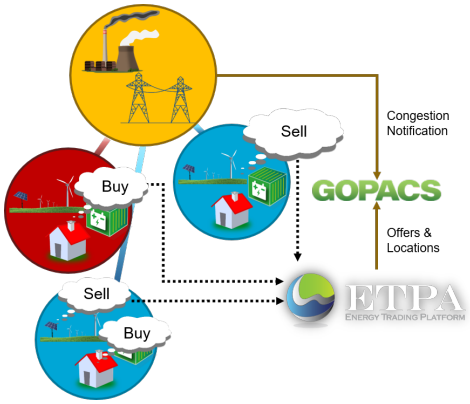


Figure 2.2: GOPACS operation example.

platform is informed about the bids and the service providers' location and runs an optimization algorithm. The optimization algorithm's objective includes the minimization of mismatch between buy order prices and sell order prices while solving the congestion. This mismatch is named IDCONS (intraday congestion spread). For each order and its location, the GOPACS algorithm checks whether the order can solve the congestion and if it does not provoke a congestion somewhere else in the grid. The final IDCONS is paid by the network operator who submitted the initial request.

Fig 2.2 illustrates an example of the GOPACS operation. The TSO (yellow circle), recognizes a congestion on a connection with the distribution network (red area), and sends a congestion notification on the GOPACS platform. The flexibility service providers that agree to submit their bids for congestion management purposes are then available to submit offers to ETPA. ETPA also checks whether there are offers submitted prior to the congestion notification, which agrees to be used for congestion management if needed. Within the example Fig.2.2, a buy offer is submitted from a flexibility service provider from the congestion region, and other offers are submitted by the flexibility service providers from other regions (blue circles). Offers submitted on the ETPA that did not agree to participate to IDCONS, are not considered by the GOPACS algorithm. The GOPACS algorithm will then match the offers from the congestion region with the opposite orders from outside that region to ensure balanced power flow and cover the cost differences. Accordingly, from Fig.2.2, the GOPACS algorithm will match the buy offer from the congestion area with one or more sell offers from the other regions.

DA/RE Platform DA/RE platform was developed in Germany and is focused on enabling and smoothing the participation of flexibility providers to the Redispatch 2.0 congestion management scheme.

This platform does not participate in markets (which have a voluntary nature) since Redispatch 2.0 mandates participation from all flexibility providers (with capacity greater than $100kW$), and economically neutral operators. The economically neutral providers principle ensures that the economic compensation provided to the flexibility providers

by the operators only covers expenses related to the adjustments required for the flexibility activation and loss of other revenue opportunities. DA/RE can be operated within minutes prior to the deployment of flexibility services, and it compensates in $€/MWh$. The process followed by the DA/RE platform can be summarized in 6 steps:

1. The TSO or DSO submits redispatch requirements.
2. The redispatch requirements are aggregated along with the planning data from the other system operators.
3. A central optimization algorithm selects which flexibility providers will activate their flexibilities based on:
   - Network restriction from all operators.
   - Simultaneously optimizing for all network levels.
   - Operating costs.
4. The selected flexibility providers inform the platform about their expected operating costs, to estimate the economic compensation.
5. An automatic or manual activation of the chosen flexibility services is performed.
6. The data measured from the flexibility activation are used to compute the actual operating costs.

Real-Time Interface The Real-Time Interface (RTI) in the Netherlands is a standardized technical solution designed to facilitate real-time communication between DSOs, the TSO, and large DERs, such as solar PV systems, wind farms, and battery storage units. RTI became mandatory in 2024 for newly connected generating installations above $1MW$. The primary objective of RTI is to enable real-time monitoring and control of DER output, thereby enhancing grid security while allowing more renewable energy to be connected without extensive infrastructure upgrades. Presently, RTI considers congestion management. However, RTI expansions for additional services are planned in the future.

Technically, the RTI consists of two key components: a system operator endpoint and a customer endpoint. The customer endpoint is installed behind the meter and communicates with the system operator through secure communication channels. The interface allows system operators to send active or reactive power setpoints, request curtailment actions, and receive real-time measurements of voltage, current, active, and reactive power. To ensure reliability, devices installed at the customer endpoint must be certified according to the national technical specification maintained by Netbeheer Nederland.

This doctoral research considered the measurement units required in the context of RTI to provide realistic measurement assumptions in the developed data-driven algorithms.

2.3. DISTRIBUTION SYSTEM AGGREGATED DYNAMIC RESPONSE REPRESENTATION

The expected high impact of distributed generation (DG) intrigued researchers to analyze their modeling within the distribution systems, and their effects on distribution system power quality, safety, and reliability [11]. However, uncertainties associated with the DG placement and operation challenge the assessment of DG-equipped feeder performance [12]. Therefore, the complexity of accurately modeling DSs with high DER penetration posed a significant challenge to system operators [13], [14]. When considering dynamic simulations, TSOs should consider the dynamic response of interconnected DS, but fully modeling DS feeders could be highly complex and computationally challenging.

The primary purpose of DS equivalent modeling, or ADRR, is to provide simple DS models for the system operators to perform dynamic studies. The main approaches for ADRR are white box, black box, and grey box [45].

The white box approach is applicable for cases where the complete DS structure is known. Thus, white box models use prior knowledge and physical insights [45]. However, as DS have limited observability, and DER details can be unknown to DSOs, white box approaches are not common. Black box approaches do not utilize any physical network information but rather fit models to approximate the dynamic responses of collected data [45]. Grey box approaches can be considered a combination of black box and white box since they estimate parameters similarly to black box models but also use a known structure of the equivalent model [45], [46]. A grey box approach will model the equivalent network components but not specify the parameters based on system knowledge [46]. Hence, the diversification of black box, grey box, and white box approaches can be based on the physical insight into the DS.

2.3.1. WHITE BOX

As white box approaches expect complete network structure knowledge, and distribution networks suffer from reduced observability, there are not many recent equivalent modeling studies focused on these approaches. One study that could be considered as white box is [47]. This study, unlike the rest of the referenced studies, is not focused on system dynamic studies, but rather for PV impact studies. Accordingly, the approach in [47] simplifies and represents the feeders to the specified buses of interest.

2.3.2. GREY BOX

Grey box methodologies can further be categorized in coherency based, system identification based, and dynamic model reduction based. Coherency based methods try to identify groups of generators that are coherent. Dynamic model reduction methods, such as modal methods, are based on linearized models of external systems, while system identification methods are based on comparisons between measurement and simulation responses [48].

Dynamic model reduction methods There are multiple possible dynamic model reduction methods when linearizing the system, such as Hankel norm [46], Krylov, or

modal reduction [49]. Furthermore, modal reduction can be split into approaches such as selective modal analysis or matrix transformation technique [48]. The general approach when applying dynamic model reduction methods involves removing the eigenvalues of the linearized system located further from the origin [48].

The equivalent model developed in [48] is represented as a single generator or load. However, in the process of linearization, the system can become less accurate, or sensitive to the single operating condition in which it is linearized [46]

System identification methods These methods appear to be the most prominent out of the found literature on grey box models. One strength of these methods, is the avoidance of linearization of the network representation around an operating point.

One of the main issues of system identification methods, is detecting which parameters will be fine-tuned for the equivalent model. Therefore, [50] also performed trajectory sensitivity analysis. This approach, allowed [50] to investigate how sensitive the measured output variables (measured signals on the PCC) are to each parameter to be used in the equivalent microgrid. After the set of parameters to be tuned was decided, the ADRR evolutionary algorithm was applied to identify their values. [50] argues that a voltage source converter, a synchronous generator, and a composite load are suitable for ADRR. [51] also applied this representation, where an evolutionary particle swarm optimization algorithm detected the values of the parameters of this representation. Other representations, such as a converter connected generator, and a composite load in parallel, were used [46].

Generally, grey box approaches are prominent, but are developed and valid for limited ranges of DS operating conditions. Thus, grey box approaches might have issues with generalizing to alternate operating conditions or topological shifts. Considering the highly variable state in DSs, the research in this thesis developed a black box-based approach, considering a wide range of DS operating conditions.

2.3.3. BLACK BOX

Black box approaches do not rely on physical knowledge of the system, but use measurements to identify the response that the ADRR model should give to different disturbances. Accordingly, machine learning-based approaches are the most prominent for black box ADRR models [52]–[54]. One advantage that these methods have compared to grey or white box methods, is that they do not require DS simplifications, do not apply linearization to the DS model, and can fit a large amount of data, including high variability of disturbance scenarios and DS conditions.

[52] recognized that the forward calculation of a long-short-term memory unit (LSTM) can represent the discretized differential equation of an aggregated load. [52] also identified that adding fully connected neural network hidden layers on top of the LSTM hidden layers can represent the discrete algebraic equations of an aggregated load. [53] utilized a recurrent artificial neural network to represent the active and reactive power flows from the ADN to the transmission grid.

2.4. MACHINE LEARNING

Machine learning (ML) is a branch of artificial intelligence (AI) that focuses on designing algorithms that enable computers to identify patterns and make decisions or predictions based on data. Instead of relying on a fixed set of rules, machine learning models build statistical representations of the underlying data to perform tasks such as classification, regression, clustering, or control. ML has become a foundational tool across various domains, including natural language processing, computer vision, power systems, healthcare, and finance.

At a high level, machine learning techniques can be categorized into supervised, unsupervised, and reinforcement learning, each addressing different types of problems and scenarios. This thesis deployed supervised ML models, as the applications are mainly predictive and require direct input-output mapping.

2.4.1. UNSUPERVISED LEARNING

Unsupervised learning deals with data not explicitly labeled. The unsupervised learning objective is discovering unknown patterns, structures, or relationships in the data. Techniques such as clustering [55] (e.g., grouping loads based on consumption behavior) and dimensionality reduction [56] (e.g., principal component analysis to reduce data complexity) are common applications. Unsupervised learning can be particularly helpful in preprocessing and data analysis tasks, where human-labeled data is scarce or unavailable.

2.4.2. REINFORCEMENT LEARNING

Reinforcement learning (RL) was inspired by trial-and-error in psychology, optimal control, and engineering [57], where an agent learns to make sequential decisions through environment interactions. The agent obtains rewards or penalties in response to its actions and aims to learn a policy that maximizes the total reward accumulated over time. RL is well-suited for problems where feedback is delayed and actions influence future outcomes, such as in robotics, game playing, or energy management in power systems (e.g., electric vehicle charging scheduling) [58]. Key components of an RL framework include the agent, environment, states, actions, and the reward signal.

2.4.3. SUPERVISED LEARNING

In supervised learning, model training involves data with known target labels or outputs. The objective of supervised learning is to learn a function that maps inputs to outputs while generalizing effectively to unseen data. Typical supervised learning tasks are classification (e.g., identifying whether a DS is overloaded or not) and regression (e.g., predicting DER output active power). This thesis mainly deployed supervised learning algorithms.

Supervised learning algorithms can be diversified into numerous categories depending on the task characteristics. Fig.2.3 diversifies major supervised learning categories and subcategories. Nevertheless, due to the rapidly evolving nature of ML, these categories are consistently expanding. Supervised linear models are a foundational class of ML algorithms that learn a linear relationship linking input features and target outputs from

labeled data. Ensemble models consider the predictions of multiple base learners to enrich overall performance and robustness. Neural networks (NNs) are machine learning models inspired by the human brain structure, formed from interconnected layers of nodes (neurons). The uncertainty in DS conditions guided this research to BNNs. In addition, the sequential and non-sequential data impacting ADRRs guided this research to hybrid architectures, consisting of RNNs and FNNs.

NEURAL NETWORKS

In supervised learning, NNs learn complex, non-linear mappings from inputs to outputs through backpropagation and gradient descent. For a simple feedforward NN with L layers, the forward pass is defined as:

$$\begin{aligned} h^0 &= x, \\ h^1 &= \phi(W^1 h^0 + b^1), \\ &.. \\ \hat{y} &= \phi(W^L h^{L-1} + b^L), \end{aligned} \tag{2.15}$$

where x is the input, W^l, b^l are the l^{th} layer's weights and biases, and $\phi(\cdot)$ is a non-linear activation function. $\hat{y}$ is the model prediction.

Activation function An activation function introduces non-linearity into a NN, enabling it to model complex patterns beyond simple linear relationships. Common activation functions encompass the rectified linear unit (ReLU), sigmoid, and tanh. Without activation functions, a multi-layer neural network could be represented as a single-layer linear model.

Loss function A supervised NN is trained with a loss function that quantifies the difference between predicted $\hat{y}$ and the target values y. Typical NN loss functions for regression include the mean squared error (MSE), the root mean squared error (RMSE), and the mean absolute error (MAE) functions. Typical classification losses include hinge loss and sigmoid cross entropy loss for binary classification, and softmax cross entropy loss for multi-class classification.

Backpropagation Backpropagation is the fundamental algorithm used to compute gradients in NNs. Backpropagation calculates the loss gradient concerning the NN weights by recursively applying the chain rule from the output layer back to the input. These gradients inform how each parameter will be adjusted to reduce the loss. Backpropagation makes training deep NNs computationally feasible and efficient by reusing intermediate computations.

Parameter Update A parameter update is the process of adjusting the weights and biases in a model to reduce the missmatch between predictions and true outputs. Guided by the gradients computed through backpropagation, an optimizer like gradient descent or Adam modifies the parameters in small steps. This iterative adjustment process enables models to learn from data over time and to enhance their performance.

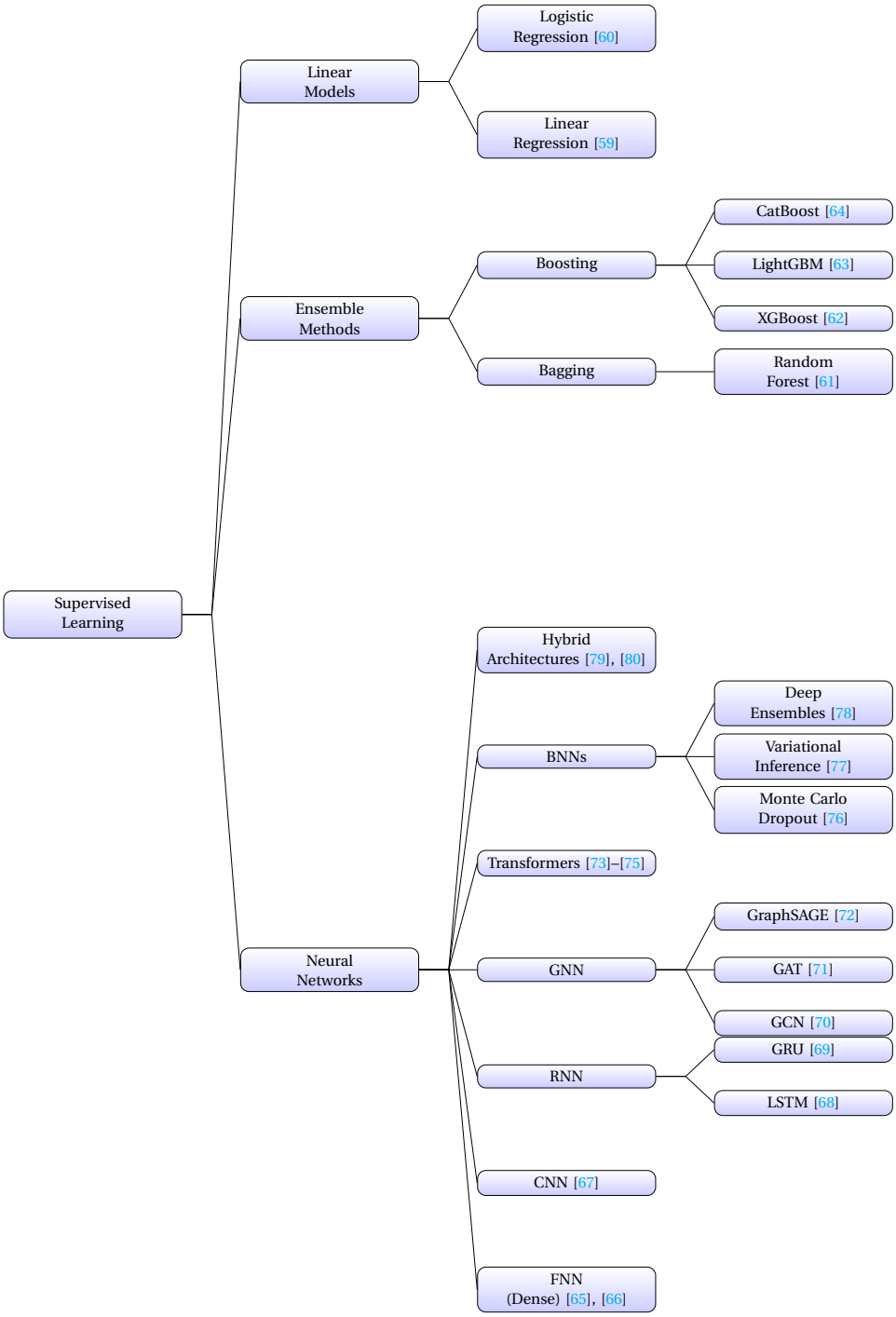


Figure 2.3: Supervised learning diversification into widely used categories.

BAYESIAN NEURAL NETWORKS

Bayesian Neural Networks are a probabilistic extension of conventional neural networks where weights are treated as distributions rather than fixed values. This allows the model to capture epistemic uncertainty, reflecting confidence in its predictions. Training a BNN involves estimating the posterior distribution over weights given the data, often using variational inference or Monte Carlo sampling. The output of a BNN is typically a predictive distribution rather than a single point estimate, making it useful in safety-critical or uncertainty-aware applications.

KL Divergence The Kullback–Leibler (KL) divergence is a measure of how one probability distribution diverges from another, often interpreted as the information lost when approximating a true distribution with a simpler one. Minimizing the KL divergence can be similar to minimizing the negative log likelihood (NLL). NLL is a common objective in BNNs and supervised learning.

Bayes Theorem The Bayes theorem is a key rule in probability theory that describes how new evidence can update the probability of a hypothesis. Bayes theorem is expressed as:

$$p(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}, \tag{2.16}$$

where $P(B)$ is the marginal likelihood. The Bayes theorem underlies Bayesian inference approaches.

LONG SHORT-TERM MEMORY

LSTM networks are a specialized type of recurrent neural network (RNN) designed to model sequences and retain long-term dependencies. Unlike standard RNNs, LSTMs use gates (input, forget, and output gates) to manage the information flow through memory cells, effectively mitigating the vanishing gradient issue in training. This makes LSTMs highly effective in tasks where both recent and distant past information can be important.

3

Tensor Convolution-based Aggregated Flexibility Estimation in Active Distribution Systems

Act always so as to increase the total number of choices.

Heinz von Foerster

Power system operators require advanced applications in the control centers to tackle increasingly variable power transfers effectively. One urgently needed application concerns estimating the feasible available aggregated flexibility from a power system network, which can be effectively deployed to mitigate issues in interconnected networks. This chapter proposes the TensorConvolution+ algorithm to address the above application. Unlike related literature approaches, TensorConvolution+ estimates the density of feasible flexibility combinations to reach a new operating point within the p-q flexibility area. This density can improve the decision-making of system operators for efficient and safe flexibility deployment. The proposed algorithm applies to radial and meshed networks, is adaptable to new operational conditions, and can consider scenarios with disconnected flexibility areas. Using convolutions and tensors, the algorithm efficiently aggregates the combinations of flexibility providers' adjustable power output that can occur for each flexibility area set point. Simulations on the meshed Oberrhein and radial CIGRE test networks illustrate the effectiveness of TensorConvolution+ for flexibility estimation with high numerical confidence and a minor computing effort. Additional simulations highlight how system operators can interpret the estimated density of feasible flexibility combinations for decision-making purposes, the algorithm's capability to estimate disconnected flexibility areas, and adapt to new operating conditions.

3.1. INTRODUCTION

The coordination between transmission system operators (TSOs) and distribution system operators (DSOs) faces challenges in data exchange [81] as the flexibility services from providers connected to the distribution networks (DNs) become important for services such as balancing and congestion management. TSOs need to anticipate the available DN flexibility, to effectively use this flexibility in their operation. DSOs must ensure that said flexibility services respect the DN operational constraints and use these services in distribution-level markets [10]. Between existing approaches for TSO-DSO coordination, the DSO-managed approach requires exchanging non-sensitive information from the DSOs to the TSOs. This information is flexibility areas (FAs); areas in the active and reactive power plane designating the extreme values of flexibility that the DN can "offer" to the transmission network at their point of common coupling (PCC) [18]. The DSO-managed coordination process involves two tasks for DSOs. In the first task, the DSO obtains offers from flexibility service providers (FSPs), estimates the theoretically feasible FA as in [18]–[26], [30], [38], [82], and informs the TSO of that FA. The TSO can request an operating point (OP) within the FA from the DSO. The DSO performs the second task, optimizing the individual FSP shifts achieving the TSO request while respecting the DN constraints and minimizing the costs as in [83], [84].

In existing FA estimation algorithms, the area surrounded by the FA curve's extreme values is mainly assumed as equally reachable and feasible. However, a different set of flexibility combinations can reach each FA OP, and the algorithms cannot guarantee the FSP availability and actual desired response [84], [85]. Therefore, each FA OP should not be viewed as equal, but represent the density of feasible flexibility combinations through which the OP can be reached. This additional information on FA OP can influence the TSO to select OP based on expected delivery, reliability, and effectiveness. Currently, to the best of the authors' knowledge, no algorithm provides information on the density of feasible flexibility combinations (DFC). Hence, TSOs may be unable to select safe and efficient shifts for FA OPs. This chapter focuses on FA estimation, considering all combinations and providing the DFC for each FA OP efficiently.

Early distribution network FA estimation studies [20], [24] effectively proposed power flow (PF)-based algorithms. These algorithms provide simple and coherent methods for FA estimation but have limitations in exploring the flexibility area space [24] and high computational time [20]. Optimal power flow (OPF)-based approaches are faster and apply multi-objective optimization such as ϵ-constraint method [22], [85], weighted sum method [19], [26], and radial reconstruction-based method [18], [23]. OPF-based approaches provide straightforward and efficient algorithms to identify the FA limits, improving the potential of system operators to include flexibility in their decision-making. However, OPF-based approximations may have limited validity in meshed networks, as highlighted by [21], [26], [38]. The vast majority of studies use radial networks [18]–[25], [38], [82], [85].

The authors of [30] introduced an efficient chance-constrained FA on meshed networks, showcasing the benefits of power flow routers for FAs. The results of [30] showed rela-

Parts of this chapter have been published in: D. Chrysostomou, J. L. R. Torres and J. L. Cremer, "Tensor Convolution-Based Aggregated Flexibility Estimation in Active Distribution Systems," in IEEE Transactions on Smart Grid, vol. 16, no. 1, pp. 87-100, Jan. 2025, doi: 10.1109/TSG.2024.3453667 [36].

tively slower performance than the other OPF-based algorithms. A limitation of OPF-based algorithms, as identified by [19], [38], is the incapability to deal with FAs that are disconnected, which might be the case for FSPs with limited offered flexibility setpoints. This chapter's proposed algorithm performs in radial and meshed network topologies and can estimate disconnected FAs.

The objective of existing FA estimation approaches is to identify the limits of the aggregated DN flexibility [18]–[23], [25], [26], [30], [82]. Evaluating each possible FSP shift combination through existing approaches is intractable, as they are tailored to the above objective. The authors of [85] considered the reliability of each FSP in the algorithm as the confidence in delivering the offered flexibility. The approach of [85] does not evaluate all possible flexibility shift combinations to reach an OP between any set of FSPs. Nevertheless, the results of [85] showcased great insights on the inner area of FAs, but the task was computationally expensive with an average time of 970s. This chapter's proposed FA estimation approach considers all possible flexibility shift combinations and reports the DFC to reach each FA OP.

Existing FA estimation approaches mainly require the OC of the DN as input data to perform OPF and PF simulations as [19], [21]–[26]. However, DNs have typically limited real-time observability with measurement units on limited network components [1], [27]–[29]. To deal with uncertainties from renewable sources, [30] applied a chance-constrained method, and [18] applied robust optimization. [20] included the probability distributions of forecast errors to determine the probability of feasibility. The proposed algorithm can use the FA estimated under expected DN OC (bus voltages, power injections, and line loading), to adapt and approximate the FA in correlated real-time OC with observability limited to a subset of network components.

Deep learning-based algorithms have been recently explored for the FA estimation problem [86], and in tasks related to FA estimation such as OPF [87], [88] and PF [89], [90]. Deep learning models can improve efficiency in performing these tasks. However, some limitations concerning low generalization to unseen network topologies and requiring large training datasets can challenge the application of deep learning in algorithms for FA estimation. The proposed algorithm does not employ deep learning models and does not require training.

The proposed TensorConvolution+ algorithm deals with issues on the exploration of inner FA, evaluates all FSP combinations and informs on the DFC to reach each FA OP. Existing alternative algorithms explore the extreme limits that flexibility combinations can achieve. Evaluating all combinations informs the system operators on which PCC operating points have more feasible flexibility options. A higher DFC can correspond to safety regarding network constraints, less dependency on specific FSPs, and more options to optimize costs in algorithms such as [83], [84]. Alternative approaches cannot tractably evaluate all possible discrete combinations, whereas TensorConvolution+ is time efficient. In addition, TensorConvolution+ handles the limitation of existing FA estimation algorithms in dealing with disconnected FAs and discrete FSPs. This chapter's contributions are:

- Developing the first method to evaluate all possible discrete combinations of flexibility, improving the information encompassed in FAs.

Figure 3.1: DSO-managed approach for TSO-DSO Coordination. Steps order (#).

- Introducing the application of convolutions in flexibility estimation, and analyzing useful properties of convolution to simplify the FA problem's complexity.
- Introducing the application of tensors in flexibility estimation, to store and evaluate the feasibility of flexibility combinations.

Case studies on the 70-bus and 109-bus meshed medium voltage Oberrhein networks and the 15-bus CIGRE medium voltage radial network with photovoltaic and wind plant modules show the algorithm's capability to work for diverse network topologies, the need for DFC in FAs, the above contributions and the algorithm's adaptability to partially observable OCs.
The following sections are (II) flexibility estimation algorithm; introducing the problem formulation and objectives, (III) tensor convolutions; the application of tensors and convolutions in the algorithm, (IV) case studies; on the need for DFC and algorithm contributions, and (V) conclusion.

3.2. FLEXIBILITY ESTIMATION ALGORITHM

In DSO-managed coordination, as illustrated in Fig.3.1 the first DSO task concerns the FA estimation. This task informs the TSO about the available shifts $s^o = [\Delta p^o, \Delta q^o]^T$ from the initial PCC OP $[p^o, q^o]^T$, considering the DN network constraints. The FA estimation impacts the TSO selection of an aggregated shift, and subsequently the DSO second task of FSP shift optimization. Therefore, the FA estimation problem requires (i) exploring possible FSP shift combinations and (ii) evaluating whether each combination results in a feasible OC for the network constraints.
The requirement (i) on FA estimation approaches determines which area can be reached using FSP shifts. The set of FSPs is Ω^{FSP}. The generic infinite set of shifts is $\Omega^S = \{s | s = [\Delta p, \Delta q]^T \in \mathrm{R}^2\}$. Each FSP $i \in \Omega^{FSP}$ offers a set of shifts as $\bar{\Omega}_i^S \subset \Omega^s$. The function $\sigma : \Omega^S \rightarrow \Omega^S$ maps each shift of FSP i from its bus to the PCC. Thus, the set of shifts from each FSP, as observed at the PCC is $\Omega_i^S = \{s | s = \sigma(\bar{s}), \forall \bar{s} \in \bar{\Omega}_i^S\} \subset \Omega^S$. The set of all shift combinations at the PCC between all FSPs is:

$$\Omega^C = \Omega_1^S \times \Omega_2^S \times ... \Omega_{|\Omega^{FSP}|}^S, \in \mathrm{R}^{2|\Omega^{FSP}|}, \tag{3.1}$$

$$|\Omega^C| = \Pi_{i \in FSP} |\Omega_i^S|, \tag{3.2}$$

where $\times$ is a set cartesian product and $|\cdot|$ is a set cardinality. Each combination $\pi \in \Omega^C$ corresponds to one tuple of FSP-shifts, and $\pi \rightarrow \{s_1, s_2, .., s_{|\Omega^{FSP}|}\} := H(\pi)$. Thus, function $H(\pi)$ represents obtaining a set of individual FSP shifts participating in the combination π.

The issue fulfilling requirement (i) is that the number of flexibility combinations grows exponentially as FSPs and their offered shifts increase. Due to this issue, state-of-the-art approaches do not explore *all* possible FSP shift combinations. The proposed approach estimates *all* possible combinations, i.e. all $\pi \in \Omega^C$.

The requirement (ii) on FA estimation approaches determines whether each discrete FSP shift combination considers the network constraints. Let the reachable shifts at the PCC s^o expanded by the index π for each $\pi \in \Omega^C$ as:

$$s^o_\pi = \sum_{s \in H(\pi)} s, \tag{3.3}$$

as each combination π leads to one s^o. Using (3.3), the set of reachable shifts at the PCC from all possible combinations is:

$$\Omega^{S^o} = \{s^o_\pi, \forall \pi \in \Omega^C\} \subset \Omega^S. \tag{3.4}$$

As shown in Fig.3.2, multiple combinations π can reach the same PCC OP shift s^o, i.e. $\exists s^o_\pi = s^o_{\hat{\pi}}$ s.t. $\pi, \hat{\pi} \in \Omega^C$. Therefore, the authors expand Ω^C by the index s^o, indicating the set of combinations leading to each s^o as:

$$\Omega^C_{s^o} = \{\pi | s^o_\pi = s^o \forall \pi \in \Omega^C\} \subset \Omega^C. \tag{3.5}$$

The cardinality $|\Omega^C_{s^o}|$ represents the number of possible combinations reaching s^o. For example in Fig.3.2, for $s^o = [3,3]^T$, the $|\Omega^C_{s^o}| = 2$ (from π_1, π_2). However, out of the 2 combinations, only one leads to a feasible voltage (π_2). Therefore, an FA estimation approach should account for the network constraints. Feasible shift combinations require network constraints to be fulfilled at the resulting shifted OCs. Every shift combination π when applied, can impact the voltage magnitude $v_{b,\pi}$ of every network bus $b \in \Omega^B$, and the loading $l_{z,\pi}$ of every line or transformer $z \in \Omega^L$. The set of feasible combinations reaching each s^o is:

$$\begin{aligned} \Omega^{FC}_{s^o} = \{\pi | (&v_{\min} \leq v_{b,\pi} \leq v_{\max} \forall b \in \Omega^B, \\ &|l_{z,\pi}| \leq l_{\max} \forall z \in \Omega^L), \forall \pi \in \Omega^C_{s^o}\} \subset \Omega^C_{s^o}. \end{aligned} \tag{3.6}$$

$v_{\max}, v_{\min}$ are the maximum and minimum allowed voltage for network buses, $l_{\max}$ is the maximum allowed loading for network lines and transformers. The set of all shifts at the PCC that are feasible is $\Omega^{FS^o} = \{s^o | 1 \leq |\Omega^{FC}_{s^o}| \forall s^o \in \Omega^{S^o}\} \subset \Omega^{S^o}$. For each s^o its DFC $= |\Omega^{FC}_{s^o}| / \max_{\hat{s}^o \in \Omega^{S^o}} |\Omega^{FC}_{\hat{s}^o}|$, the normalized cardinality of $|\Omega^{FC}_{s^o}|$.

The issue of fulfilling requirement (ii) is that evaluating the impact of all combinations π on all network components is computationally expensive. Hence, evaluating the impact of *all* possible combinations is intractable in existing approaches. State-of-the-art approaches simplify the issue to identify all s^o for which at least 1 feasible shift combination exists, i.e. Ω^{FS^o}. The proposed approach evaluates *all* possible combinations and estimates $|\Omega^{FC}_{s^o}|$ of all $s^o \in \Omega^{FS^o}$.

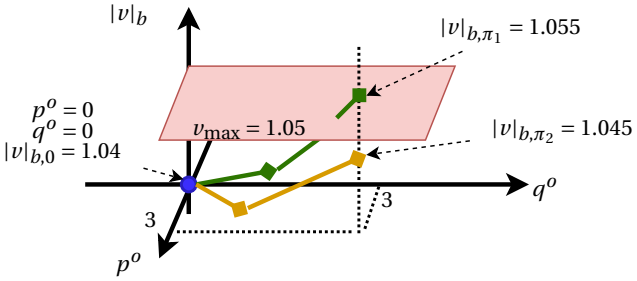


Figure 3.2: Example for 2 shift combinations (π_1 as ,π_2 as) from the initial OP ( ) to reach $p^o + \Delta p^o =$ 3MW, $q^o + \Delta q^o =$ 3MVAR, but only π_2 feasible for bus b due to the maximum voltage constraint ( ).

The proposed algorithm is TensorConvolution+ with input as the network topology, the initial OCs, and the FSPs. When adapting FAs for new OCs, the needed input is a subset of the network components' new voltage and loading magnitudes. Hyperparameters include the resolution $\delta p, \delta q$ of Ω^S, i.e. increments in Ω^S elements, and the sensitivity thresholds c_v, c_l. Fig.3.3 shows the algorithm's steps. TensorConvolution+ initially runs PF simulations to generate samples for requirements (i) and (ii). The algorithm decomposes the flexibility constrained by each network component (e.g., bus, line), and uses tensors and convolutions to efficiently process the samples for requirement (ii). Then, TensorConvolution+ returns the FA constrained by all network components. If all combinations fulfill the network constraints, i.e. $\Omega^{FC}_{s^o} = \Omega^{C}_{s^o} \forall s^o \in \Omega^{S^o}$, then convolutions are applied to deal with requirement (i). Fig.3.4 shows the output FA of the proposed algorithm. Each blue-colored pixel has a value of 1 and represents a reachable but not feasible OP for the network constraints. Each pixel with values between $(1,2]$ represents a reachable and feasible OP and its DFC, i.e., larger values have more feasible options than lower values. The area covered by all colored pixels represents the reachable set Ω^{S^o} from FSP offers, related to requirement (i). The feasible area represents Ω^{FS^o} and DFC, related to requirements (i), (ii).

3.2.1. REDUCING REQUIRED POWER FLOW SIMULATIONS

Computing all inputs to (3.6) requires estimating $v_{b,\pi}, l_{z,\pi}$ as:

$$v_{b,\pi} = v_{b,0} + \sum_{s \in H(\pi)} \Delta v_{b,s} \quad \forall b \in \Omega^B, \forall \pi \in \Omega^C_{s^o}, \tag{3.7}$$

$$l_{z,\pi} = l_{z,0} + \sum_{s \in H(\pi)} \Delta l_{z,s} \quad \forall z \in \Omega^L, \forall \pi \in \Omega^C_{s^o}, \tag{3.8}$$

where $v_{b,0}, l_{z,0}$ are the OC's values for $b \in \Omega^B$ and $z \in \Omega^L$. Variables $\Delta v_{b,s}, \Delta l_{z,s}$ are the shift s impacts on b, z, estimated using PF simulations. The number of PF simulations needed to explore all possible FSP shift combinations is (3.2). To simplify this combination complexity, the authors apply:

Assumption 1: The impact of each FSP's output shift on a network component is not affected by other FSPs' output shifts, i.e., $\Delta v_{b,\tilde{s}} \perp\!\!\!\perp \Delta v_{b,\hat{s}} \forall \tilde{s} \in \Omega^S_{\tilde{i}}, \forall \hat{s} \in \Omega^S_{\hat{i}}, \forall \tilde{i} \neq \hat{i} \in \Omega^{FSP}$, $\Delta l_{z,\tilde{s}} \perp\!\!\!\perp \Delta l_{z,\hat{s}} \forall \tilde{s} \in \Omega^S_{\tilde{i}}, \forall \hat{s} \in \Omega^S_{\hat{i}}, \forall \tilde{i} \neq \hat{i} \in \Omega^{FSP}$.

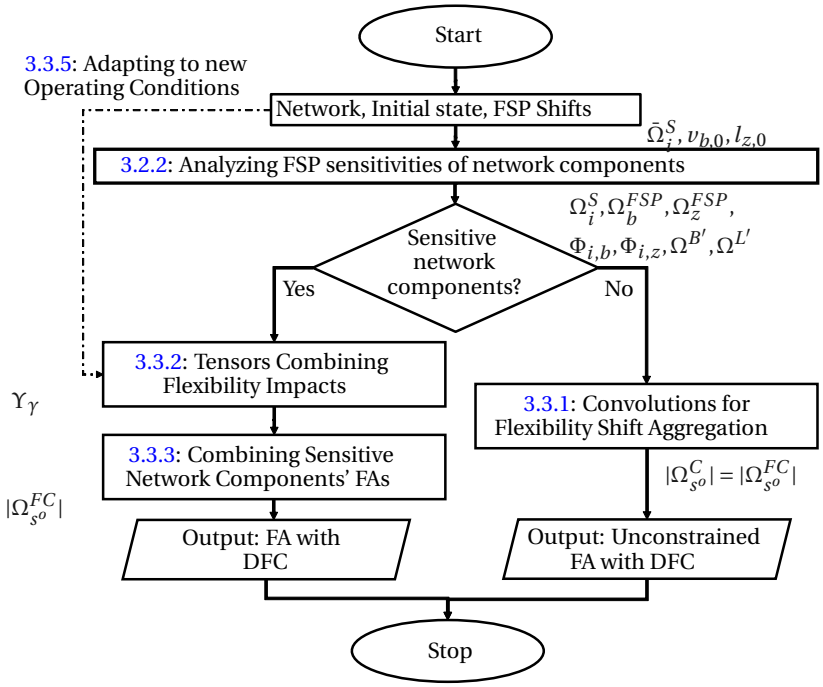


Figure 3.3: Overview of the proposed flexibility estimation algorithm.

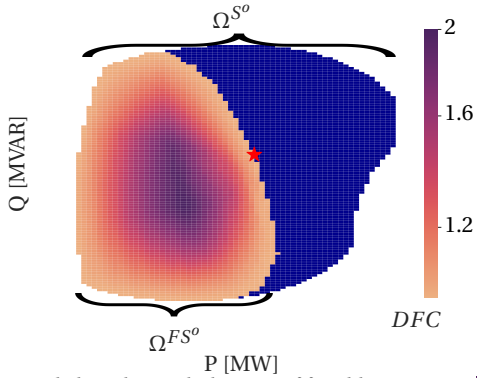


Figure 3.4: FA from the proposed algorithm with the DFC of feasible FA points (■) and not feasible FSP shift combinations (■) from the initial OP (★).

When adopting assumption 1, the proposed algorithm requires one PF simulation for each possible FSP shift. Hence, the required PF simulations are decreased from (3.2) to $\sum_{i\in\Omega^{FSP}}|\Omega_i^S|$.

3.2.2. ANALYZING FSP SENSITIVITIES OF NETWORK COMPONENTS

This section analyzes the sensitivities of network components to FSP shifts through the impacts $\Delta v_{b,s}, \Delta l_{z,s}$. This analysis further reduces the complexity $O\left((|\Omega^B|+|\Omega^L|)\cdot|\Omega_{s^o}^C|\right)$ in estimating (3.7) and (3.8). The analysis starts with:

- *Observation (a):* each network component's voltage or loading is not sensitive to all FSPs.
- *Observation (b):* not all network components can reach their voltage or loading limitations due to the FSP shifts.

Exploiting (a), the FSP sensitivity sets for each $b\in\Omega^B$ and $z\in\Omega^L$ are:

$$\Omega_b^{FSP} = \{i | c_v \leq \max_{s\in\Omega_i^S}(|\Delta v_{b,s}|), i\in\Omega^{FSP}\}, \tag{3.9}$$

$$\Omega_z^{FSP} = \{i | c_l \leq \max_{s\in\Omega_i^S}(|\Delta l_{z,s}|), i\in\Omega^{FSP}\}, \tag{3.10}$$

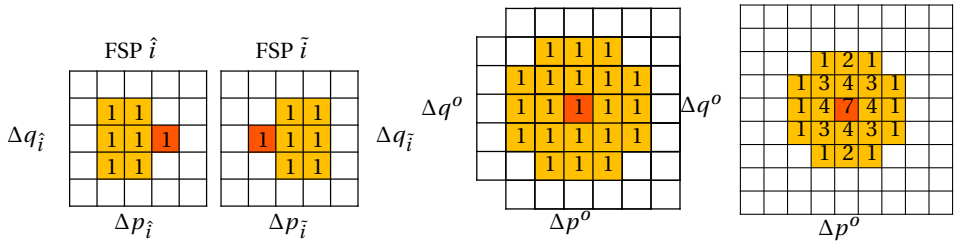
where c_v, c_l, are sensitivity thresholds. The FSPs that do not impact the constraints of a network component are $\Omega_\gamma^{FSP'} = \{i | i\in\Omega^{FSP}\setminus\Omega_\gamma^{FSP}\}\forall\gamma\in\Omega^B\cup\Omega^L$. For example, in a network with 2 feeders connected to the PCC, the components on the first feeder can be insensitive to shifts from FSPs connected to the second feeder. The sets in (3.9)-(3.10) replace Ω^C with Ω_γ^C, the set of combinations for which $\gamma\in\Omega^B\cup\Omega^L$ is sensitive to its constraints. This analysis reduces the constraint-evaluated combinations from $|\Omega_{s^o}^C|$ to $|\Omega_{\gamma,s^o}^C|\forall\gamma\in\Omega^B\cup\Omega^L$; the set of sensitive combinations per component as:

$$\Omega_{\gamma,s^o}^C = \{\pi | \sum_{s\in H(\pi)} s = s^o \forall\pi\in\Omega_\gamma^C\} \subset \Omega_\gamma^C. \tag{3.11}$$

In DN OCs with higher margin from the network constraints, $|\Omega_{\gamma,s^o}^C| << |\Omega_{s^o}^C|$.
Exploiting (b) the sets of non-sensitive components are the ones whose voltage or loading cannot reach the constraints when accumulating the highest impact from the FSPs, as:

$$\begin{aligned}\Omega^{B'} = \{b | (v_{b,0} + \sum_{i\in\Omega_b^{FSP}} \max_{s\in\Omega_i^S}(\Delta v_{b,s}) \leq v_{max})\wedge \\ (v_{min} \leq v_{b,0} + \sum_{i\in\Omega_b^{FSP}} \min_{s\in\Omega_i^S}(\Delta v_{b,s})), \forall b\in\Omega^B\},\end{aligned} \tag{3.12}$$

$$\begin{aligned}\Omega^{L'} = \{z | (l_{z,0} + \sum_{i\in\Omega_z^{FSP}} \max_{s\in\Omega_i^S}(\Delta l_{z,s}) \leq l_{max}), \wedge \\ (l_{min} \leq l_{z,0} + \sum_{i\in\Omega_z^{FSP}} \min_{s\in\Omega_i^S}(\Delta l_{b,s})), \forall z\in\Omega^L\}.\end{aligned} \tag{3.13}$$



(a) $F(\cdot)$ of $\hat{\imath}, \tilde{\imath} \in \Omega^{FSP}$. (b) Minkowski sum. (c) Convolution.

Figure 3.5: FAs for Minkowski sum and Convolution of $F(\cdot)$ for $\hat{\imath}, \tilde{\imath} \in \Omega^{FSP}$. The FAs include squares of feasible shifts(■), and initial operating points (■).

The analysis of observations (a) and (b), reduces the combinations evaluated in (3.7)–(3.8) by replacing $\Omega^C_{s^o}$ with Ω^C_{γ,s^o}, Ω^B with $\Omega^B \setminus \Omega^{B'}$ and Ω^L with $\Omega^L \setminus \Omega^{L'}$. Thus, the complexity becomes $O\left((|\Omega^B \setminus \Omega^{B'}| + |\Omega^L \setminus \Omega^{L'}|) \cdot |\Omega^C_{\gamma,s^o}|\right)$.

3.3. TENSOR CONVOLUTIONS

3.3.1. CONVOLUTIONS FOR FLEXIBILITY SHIFT AGGREGATION

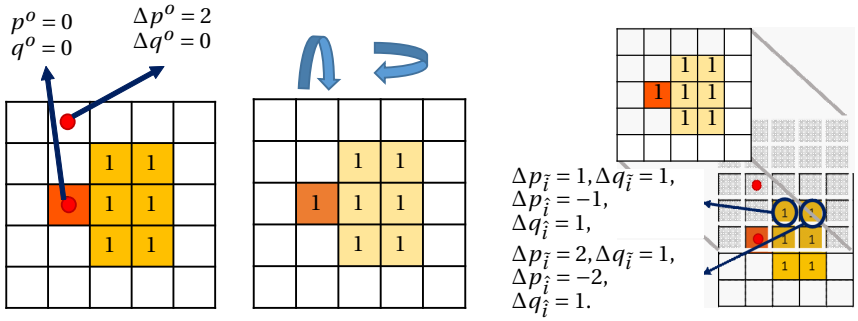
This section introduces convolutions, to aggregate shifts without considering network constraints. Convolutions can efficiently combine the flexibility sets, and accumulate the number of combinations leading to each s^o, i.e. $|\Omega^C_{s^o}|$. Relating, the Minkowski sum can be applied to efficiently combine flexibility sets without considering network constraints [91], e.g. combining feasible flexibility sets from multiple feeders connected to the PCC. However, the Minkowski sum does not consider how many combinations from the input sets lead to each s^o. In Fig.3.5(b), the Minkowski sum, and in Fig.3.5(c), the two-dimensional (2D) discrete convolution of two flexibility sets. In Fig.3.5(c), each resulting point includes the number of combinations reaching it.

The proposed algorithm aggregates shifts without considering constraints in two cases. First, when $\Omega^{B'} = \Omega^B, \Omega^{L'} = \Omega^L$, thus $\Omega^C_{s^o} = \Omega^{FC}_{s^o}$. Second, for each component $\gamma \in \Omega^B \cup \Omega^L$ where $1 \leq |\Omega^C_\gamma| < |\Omega^C|$, the algorithm explores all Ω^C combinations; the cartesian product between Ω^C_γ and all FSPs in $\Omega^{FSP'}_\gamma$, i.e. $\Omega^C = \Omega^S_1 \times \ldots \times \Omega^S_{|\Omega^{FSP'}_\gamma|} \times \Omega^C_\gamma$.

Let any 2 FSPs $(\hat{\imath}, \tilde{\imath})$ offering shifts $\Omega^S_{\hat{\imath}}$, and $\Omega^S_{\tilde{\imath}}$ respectively, as in Fig.3.5(a). The 2D discrete convolution for each shift $s^o = [\Delta p^o, \Delta q^o]^T$ is defined as:

$$(F_{\hat{\imath}} * F_{\tilde{\imath}})(\Delta p^o, \Delta q^o) = \sum_{\Delta p=-\infty}^{\infty} \sum_{\Delta q=-\infty}^{\infty} \big(F_{\tilde{\imath}}(\Delta p, \Delta q) \cdot F_{\hat{\imath}}(\Delta p^o - \Delta p, \Delta q^o - \Delta q)\big), \tag{3.14}$$

where $(\cdot * \cdot)(x, y)$ is a convolution of two functions for the input x, y, and $F_i \forall i \in \Omega^{FSP}$ is



(a) $F_{\tilde{i}}(\cdot)$. (b) $F_{\hat{i}}(\cdot)$. (c) $(F_{\hat{i}} * F_{\tilde{i}})(2,0)$ for $\hat{i}, \tilde{i} \in \Omega^{FSP}$.

Figure 3.6: Convolution of $F_{\hat{i}}(\cdot), F_{\tilde{i}}(\cdot)$ for $\hat{i}, \tilde{i} \in \Omega^{FSP}$ at $\Delta p^o = 2, \Delta q^o = 0$.

an indicator function as:

$$F_i(s) = \begin{cases} 1, & \text{if } s \in \Omega_i^S, \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

Fig.3.6 visualizes the convolution process for a single $\Delta p^o, \Delta q^o$. The circled overlapping pixels in Fig.3.6(c) represent all possible combinations reaching $p^o + \Delta p^o, q^o + \Delta q_o$. Useful properties from the convolution [92] are applied as:

1. Associativity and commutativity: The order of FSPs $\hat{i}, \tilde{i}$ (or more) does not affect the convolution output.

2. Impulse response: The convolution of an FA with a delta function results in the FA shifted by the delta offset.

The appendix B proves a third property stating: *The discrete convolution of shifts between k FSPs considers all possible discrete combinations between these k FSPs for s^o. The resulting value for each s^o is the sum of all combinations from the k FSPs that can reach it.*
The proposed approach describes the combined FA function for a sequence $\mathcal{K}$ of convolutions for k FSPs as:

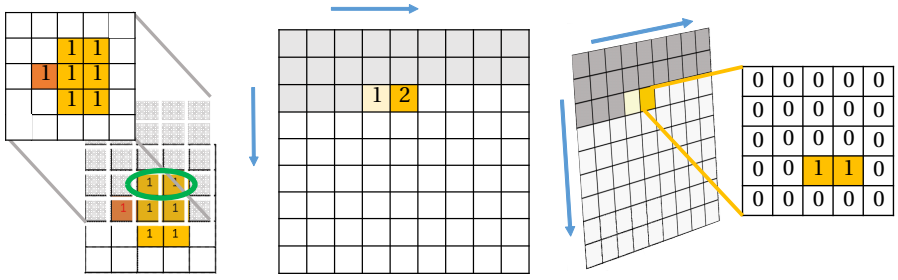
$$\tilde{F}_{\mathcal{K}}(s^o) = \begin{cases} m_{\mathcal{K}}(s^o), & \text{if } s^o \in \Omega_{\mathcal{K}}^S, \\ 0, & \text{otherwise,} \end{cases} \tag{3.16}$$

where:

$$\Omega_{\mathcal{K}}^C = \Omega_1^S \times \Omega_2^S \times \ldots \times \Omega_k^S, \tag{3.17}$$

$$\Omega_{\mathcal{K}}^S = \{s_{\mathcal{K}} | s_{\mathcal{K}} = \sum_{s \in \pi} s, \forall \pi \in \Omega_{\mathcal{K}}^C\} \subset \Omega^{S^o}. \tag{3.18}$$

$m_{\mathcal{K}}(s^o) \in \mathbf{N}$ is the number of shift combinations of FSP sequence $\mathcal{K} = 1, 2, \ldots k$ reaching s^o. For the remainder of the chapter, $F_{\hat{i}} * F_{\tilde{i}}$ corresponds to applying (3.14) for all $[\Delta p^o, \Delta q^o]^T \in \Omega_{\mathcal{K}}^S$, where $\mathcal{K} = \hat{i}, \tilde{i}$. This operation is efficient and widely available within computer vision and machine learning software libraries.



(a) Alignment of F_i, F_j. (b) Convolution result for alignment. (c) Tensor convolution result for alignment.

Figure 3.7: Convolution of $F(\cdot)$ for $i, j \in \Omega^{FSP}$ for an s^o. Out of all reachable shifts (■), 2 combinations are possible for the alignment (⬭).

3.3.2. TENSORS COMBINING FLEXIBILITY IMPACTS

This section describes the algorithm to estimate the FA of each constraint-sensitive network component, i.e., $b \in \Omega^B \setminus \Omega^{B'}, z \in \Omega^L \setminus \Omega^{L'}$. The proposed algorithm uses tensors to efficiently explore, represent, and process the information obtained through PF simulations.

Intuitively, as shown in Fig.3.6, during a 2D discrete convolution, every FSP combination reaching s^o is accumulated. However, applying (3.7)-(3.8) before the accumulation is needed to check whether each combination is feasible or not for the network constraints. Hence, the authors propose avoiding the summation step of convolution and storing the alignment of each step in new dimensions. For example, using convolution, the 2 combinations of Fig.3.7(a) result in an entry value of 2 for the matrix in Fig.3.7(b). Through the proposed tensor-convolution as in Fig.3.7(c), the entry for $\Delta p^o = 0, \Delta q^o = 2$ is a matrix of the element-wise multiplication between the F_j, and the shifted-flipped F_i. The resulting tensors store the information of which combinations π are available through this alteration.

After this tensor-convolution process, the combination of 2 FSPs results in a 4 dimensional tensor, and each additional FSP adds 2 dimensions. The function for this process is Ψ, and the tensor convolution operation is $\hat{*}$.

$$\Psi(F_1, F_2, ..., F_k) = F_1 \hat{*} F_2 ... \hat{*} F_k, \in \mathrm{R}^{2k}. \tag{3.19}$$

As shown in Sec.3.2.2 observations, the impactful FSPs for each component's constraints and the sensitive network components are limited, mitigating dimensionality issues. Nevertheless, if a limited RAM requires reduced tensor dimensionality, the algorithm can aggregate pairs of FSPs into one only for the component causing the issue. The algorithm determines which pair is closer through the electrical distance shown in [93]; the impedance of the lines between each FSP pair.

OBTAINING COMBINATION IMPACTS

For each restrictive network component $b \in \Omega^B \setminus \Omega^{B'}$, $z \in \Omega^L \setminus \Omega^{L'}$, for each of their impcatful FSPs $i \in \Omega_b^{FSP}$, $j \in \Omega_z^{FSP}$, the impact functions $\Phi_{i,b}(s), \Phi_{j,z}(s)$ are:

$$\Phi_{i,b}(s) = \begin{cases} \Delta v_{b,s}, & \text{if } s \in \Omega_i^S, \\ 0, & \text{otherwise,} \end{cases} \tag{3.20}$$

$$\Phi_{j,z}(s) = \begin{cases} \Delta l_{z,s}, & \text{if } s \in \Omega_j^S, \\ 0, & \text{otherwise.} \end{cases} \tag{3.21}$$

Estimating (3.7)-(3.8) is needed to check if the resulting $v_{b,\pi}$ are feasible for the network constraints. Thus, the algorithm needs to sum the associated Δv caused by every FSP within π for every π effective on b. The authors propose iteratively taking $i \in \Omega_b^{FSP}$ and applying:

$$T_{i,b} = \Psi(F_0, .., F_{i-1}) \hat{*} \Phi_{i,b} \hat{*} \Psi(F_{i+1}, .., F_{|\Omega_b^{FSP}|}), \in \mathrm{R}^{2|\Omega_b^{FSP}|}. \tag{3.22}$$

The above result, $T_{i,b}$, is the tensor of voltage impacts from i, on b, whose entries represent each π. $T_{i,b}$ does not include the voltage shifts caused by other FSPs. After obtaining $T_{i,b} \forall i \in \Omega_b^{FSP}$, the proposed algorithm performs element-wise tensor addition as:

$$\Gamma_b = \Sigma_{i \in \Omega_b^{FSP}} T_{i,b}, \in \mathrm{R}^{2|\Omega_b^{FSP}|}. \tag{3.23}$$

This addition aggregates the contribution to the voltage shift from all impactful FSPs for component b for each combination π, i.e., each Γ_b entry value is equal to $\sum_{s \in H(\pi)} \Delta v_{b,s}$ for a unique $\pi \in \Omega_{b,s^o}^C$. The boolean tensor $T_b^{bool} = F_0 \hat{*} ... \hat{*} F_{|\Omega_b^{FSP}|}$ shows which combinations π belong to Ω_b^C, i.e. T_b^{bool} entries are 1 where a combination exists and 0 where not. The proposed algorithm applies the following to classify each combination as feasible or not:

$$\Xi_b = \Lambda^v(\Gamma_b + v_{b,0} \cdot \mathbf{1}_{\Gamma_b}) \circ T_b^{bool}, \in \mathrm{R}^{2|\Omega_b^{FSP}|}, \tag{3.24}$$

$$\Lambda^v(v_b) = \begin{cases} 1, & \text{if } v_{min} \leq v_b \leq v_{max}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.25}$$

Where $\mathbf{1}_{\Gamma_b}$ is a tensor of ones, with the shape of Γ_b. The operation $\circ$ is the tensor Hadamard product. The addition $\Gamma_b + v_{b,0} \cdot \mathbf{1}_{\Gamma_b}$ estimates the resulting voltages at component b for every possible combination of flexibility shifts. The filter Λ^v returns 1 for all combinations within the network constraints. The Hadamard product sets 0 for all combinations not offered by the *FSPs*. The corresponding variables for loading components are $\Gamma_z, T_z^{bool}, \Lambda^l(l_z)$, and Ξ_z.

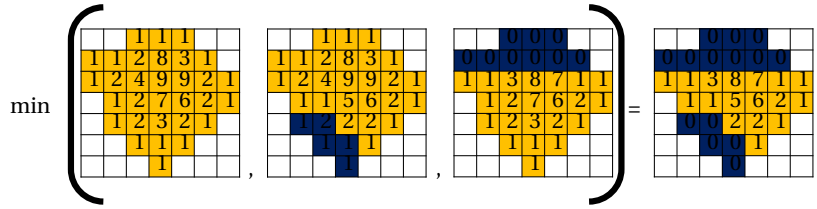


Figure 3.8: Element-wise minimum of FAs.

TENSOR TO FLEXIBILITY MAPPING

The Ξ_b, Ξ_z have dimensionalities of $2 \cdot |\Omega_b^{FSP}|, 2 \cdot |\Omega_z^{FSP}|$ respectively. The first 2 dimensions correspond to the PCC s^o; as in Fig.3.7(c). The rest constitute the constraint-validated combinations reaching each PCC s^o. Therefore, summating all tensor entries except the first 2 dimensions returns the flexibility area constrained by the component b or z. Intuitively, this operation restores the accumulation step of convolution. Using the Einstein summation convention allows easy and efficient summing operations over multiple dimensions of tensors in Python [94]. The function $C : \xi \rightarrow 2$ represents the summation of all tensor's Ξ entries from any ξ dimensions to the first 2 dimensions:

$$A_\gamma = C(\Xi_\gamma), \in \mathrm{R}^2, \tag{3.26}$$

where $\gamma \in \Omega^B \cup \Omega^L$ is any network component. Each A_γ entry corresponds to a different s^o. Each entry value of A_γ is the number of combinations in Ω_γ^C, feasible for γ for an s^o.

ADDING FSPS INSENSITIVE FOR COMPONENTS

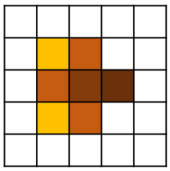
A_γ excludes contributions from FSPs causing negligible voltage or loading impacts on $\gamma \in \Omega^B \cup \Omega^L$. The proposed algorithm performs 2D convolution between A_γ and the F of all elements in $\Omega_\gamma^{FSP'} = \{j'_0, j'_1, ...\}$ to consider the feasibility of *all* possible combinations in Ω^C for component γ, as:

$$\Upsilon_\gamma = A_\gamma * F_{j'_0} * F_{j'_1} ..., \in \mathrm{R}^2. \tag{3.27}$$

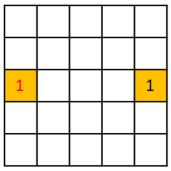
Let μ, τ be the row and column indices of $\Upsilon_\gamma \forall \gamma \in \Omega^B \cup \Omega^L$. The bijective function λ maps each s^o to a unique μ, τ, i.e. $\lambda : s^o \rightarrow (\mu, \tau)$. Each entry value of Υ_γ is the number of combinations in Ω^C feasible for γ for a unique s^o.

3.3.3. COMBINING SENSITIVE NETWORK COMPONENTS' FAS

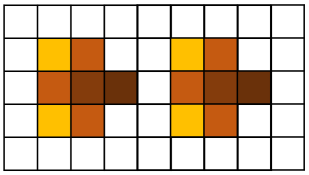
To combine the FA of all components into one, the authors apply:
Assumption 2: $|\Omega_{s^o}^{FC}| \approx \min_{\gamma \in \Omega^B \cup \Omega^L}(\Upsilon_{\gamma_{\lambda(s^o)}})$.
Using assumption 2 allows estimating the FA of each component before approximating the final FA at the PCC as their element-wise minimum. For instance, Fig.3.8 illustrates a scenario with 3 network components, one non-restrictive and two restrictive due to different constraints.



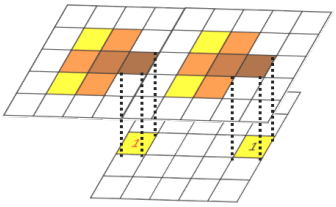
(a) Feasible combinations by other FSPs $\hat{\Upsilon}_\gamma$.



(b) Non-linear FSP F_{j_d}.



(c) $F_{j_d} * \hat{\Upsilon}_\gamma$.



(d) Dirac function property.

Figure 3.9: Convolution of a discrete variable's $F_{j_d}(\cdot)$ with the feasible combinations from other FSPs. The convolution result of Fig.3.9(c) is the same as displacing Fig.3.9(a) by the Fig.3.9(b) shifts as in Fig.3.9(d).

3.3.4. DIRAC FUNCTIONS FOR NON-LINEAR FSP

The proposed algorithm can deal with disconnected areas and non-linear FSPs, exploring the impulse response property introduced in Sec.3.3.1. Non-linear FSPs can cause disconnected areas as in [19], [38]. Non-linear FSPs could be on-load tap changers [19], [38] or generators/loads offering specific shift set points rather than a range of shifts [95]. Depending on the network component sensitivity or insensitivity to the non-linear FSPs, the algorithm performs different estimations to combine the non-linear FSP flexibility with the FA from the rest of the FSPs.

If the component $\gamma \in \Omega^B \cup \Omega^L$ is insensitive to the non-linear FSP $j_d \in \Omega_\gamma^{FSP'}$, the proposed algorithm initially performs (3.27), excluding j_d to get $\hat{\Upsilon}_\gamma$, as in Fig.3.9(a). To add the flexibility from non-linear FSPs one could convolve $\hat{\Upsilon}_\gamma$ with the indicator function of j_d, i.e, F_{j_d} of Fig.3.9(b), as in Fig.3.9(c). Alternatively, exploiting the Dirac function property of convolutions, the algorithm displaces $\hat{\Upsilon}_\gamma$ by each shift $\Delta p, \Delta q$ from j_d and sums these displaced results, as in Fig.3.9(d). The summation results to Υ_γ.

If j_d, is sensitive for $b \in \Omega^B$, the algorithm ignores the j_d impact on b until after (3.23). Then, for each j_d shift $s \in \Omega_{j_d}^S$ whose impact is $\Delta v_{b,s}$, the matrix $A_{b,s}^d$ is:

$$A_{b,s}^d = C\left(\Lambda^v(\Gamma_b + (v_{b,0} + \Delta v_{b,s}) \cdot \mathbf{1}_{\Gamma_b}) \circ T_b^{bool}\right). \tag{3.28}$$

The (3.28) encompasses adding the impacts of all linear FSP shifts to the impact of a non-linear FSP value, filtering the results based on the network constraints as in (3.24), (3.25), and summing feasible combinations for each s as in (3.26). Thus, each $A_{b,s}^d$ measures all feasible combinations between the linear FSPs and a non-linear FSP value s. Subsequently, the proposed algorithm sums $A_{b,s}^d \forall s \in \Omega_{j_d}^S$ displaced by the s offsets. The

summation result describes A_γ, used in (3.27). The (3.28) also holds for $z \in \Omega^L$, but with loading variables instead.

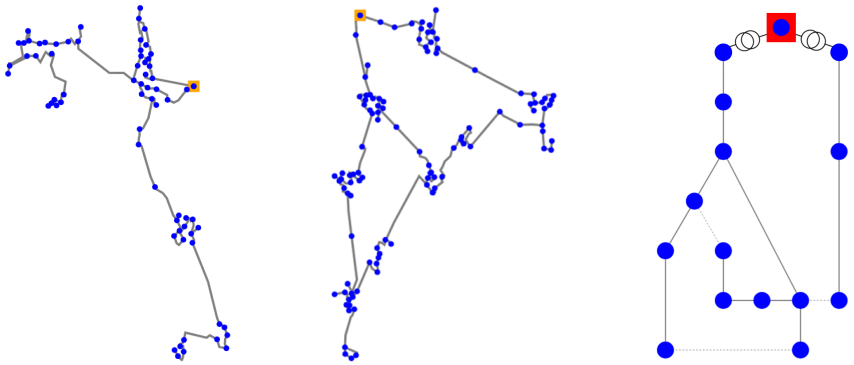
3.3.5. ADAPTABILITY TO PARTIALLY OBSERVABLE OPERATING CONDITIONS

Let FSPs that offer the same flexibility shifts for related OCs. In that case, the proposed algorithm assumes the shift impacts $\Delta v_{b,s} \forall s \in \Omega_i^S \forall i \in \Omega^{FSP}, \forall b \in \Omega^B$, $\Delta l_{z,s} \forall s \in \Omega_i^S \forall i \in \Omega^{FSP}, \forall z \in \Omega^L$ to be similar. Therefore, the tensors $\Gamma_b, \Gamma_z, T_b^{bool}, T_z^{bool}$ are also similar for these related OCs. The proposed algorithm can store the tensors $\Gamma_b, \Gamma_z, T_b^{bool}, T_z^{bool}$, and adapt the flexibility in subsequent FA estimations by applying Λ^v, Λ^l on the stored tensors and the new OCs $l_{z,0} \forall z \in \Omega^L \setminus \Omega^{L'}, v_{b,0} \forall b \in \Omega^B \setminus \Omega^{B'}$. The proposed algorithm applies tensor train decomposition to reduce the space needed to save multiple high-dimensional tensors; a method that allows efficient representation of high-dimensional tensors with a small number of parameters without losing significant information [96]. For example, let an FA be estimated for expected (e.g. day ahead) OCs. Let real-time measurement units be placed on the DN components that are sensitive to constraints and flexibility shifts, i.e., $\Omega^B \setminus \Omega^{B'}, \Omega^L \setminus \Omega^{L'}$. TensorConvolution+ stores the tensors of the expected OCs FA and uses the real-time measurements from the sensitive components to approximate the partially observable OCs FA.

3.4. CASE STUDIES

Fig.3.10 illustrates the test networks; the meshed medium voltage Oberrhein network's substations 0 (OB0) (70 buses) and 1 (OB1) (109 buses), and the radial CIGRE medium voltage network (15 buses). The authors modified OB0 and OB1 to get meshed networks and provide more challenging scenarios for the algorithm. These modifications led to initial OCs with a minimum voltage in OB0 of $0.95 p.u.$ and OB1 of $0.958 p.u.$. These modifications increased the sensitivity of network components to FSPs. The algorithm's inputs were the network and the locations of load and generator FSPs. The algorithm's parameters were the $\delta p, \delta q, c_l, c_v, l_{\max}, v_{\max}, v_{\min}$. In all case studies the loading constraint was $l_{\max} = 100\%$ and the threshold parameters were $c_v = 0.001 p.u., c_l = 1\%$ for the CIGRE network, $c_v = 0.0001 p.u., c_l = 1\%$ for OB0, and $c_v = 0.005 p.u., c_l = 1\%$ for OB1. The other inputs and parameters are referenced below and vary between the case studies. In all case studies, the FSPs were assumed to offer flexibility covering any p, q setpoint with apparent power less than their initial apparent power except in case study C where one FSP was non-linear and case study B.3.

Study A considered 3 FSPs in the CIGRE network, offering their full flexibility. The 3 FSPs were loads 3, 11 and generator 8. The FSP costs were 40 €/MW, 50 €/MW, and 60 €/MW. The exhaustive baseline approach performed PF simulations for all possible shift combinations between the FSPs for each $\delta p = 0.25$MW, $\delta q = 0.25$MVAR increment. The voltage constraints were $v_{\min} = 0.95 p.u.$, $v_{\max} = 1.05 p.u.$. The simulations recorded the cost and feasibility for the network constraints per combination. Study B.1 considered 2 FSPs in radial OB0, load 57, and generator 29 to compare the TensorColvolution+ performance to the exhaustive PF baseline with a $\delta p = 0.025$MW, $\delta q = 0.025$MVAR increment. The voltage constraints were $v_{\min} = 0.95 p.u.$, $v_{\max} = 1.05 p.u.$. Missing values were linearly imputed. Study B.2 considered 360 estimations of FAs for 2–15 FSPs on



(a) OB0 (b) OB1 (c) CIGRE network.

Figure 3.10: Test network lines (—), buses (●), high to medium voltage transformer stations (■), transformers (◎), external grid (■) .

the OB0, and OB1. Each estimated FA had a random set of FSPs between the network loads and generators. For the cases with $|\Omega^{FSP}| \geq 10$, the random FSPs were sampled from 2 sets relating to different network regions. The 240 estimations had narrow voltage constraints with $c_{\min} = 0.95 p.u., c_{\max} = 1.05 p.u.$, and 120 estimations wide voltage constraints with $c_{\min} = 0.9 p.u., c_{\max} = 1.1 p.u.$. In narrow voltage constraints estimations, around 400 pixels (as in Fig.3.14), and in wide voltage constraints around 670 pixels were explored. The number of pixels determined the values of $\delta p, \delta q$ from the total capacity of flexibility offers per scenario. For each estimated FA, 1000 samples of flexibility shifts were generated with a Monte Carlo baseline. Each feasible sample generated from the baseline was assigned to a pixel on the proposed algorithm's estimated FA. If the pixel was estimated feasible by the TensorConvolution+, then the estimation was correct. The percentage between the correctly assigned feasible pixels over the total number of feasible pixels explored is a_f. The percentage between the explored and not explored pixels is a_r. The 1000-sample sets were sampled from 2 conventional (uniform, Kumaraswamy) and 1 harder distributions, amounting to 3000 samples. The CIGRE network shows the effects of radial topologies and higher resolutions with 11 FSPs, loads $9, 14, 16, 17$, and generators $0, 1, 2, 3, 4, 5, 6$ and narrow voltage constraints. The output for the CIGRE network was around 10000 pixels with $\delta p = 0.01 MW, \delta q = 0.02 MVAR$. Study B.3 compares 3 FAs estimated using an OPF-based algorithm and TensorConvolution+. The FAs included 5 and 6 FSPs on the CIGRE network. Study C considered 7 FSPs offering any setpoint in their flexibility range, loads $12, 14, 16$ and generators $0, 1, 2, 3, 8$. The wind plant FSP (generator 8) only offered full curtailment (2 setpoints) to produce a disconnected area. The voltage constraints were set at 0.94–$1.06 p.u.$, a challenging case where not-feasible shifts exist, but both areas in the disconnected FA include feasible shifts. Study D considered 9 FSPs, loads $3, 5, 6, 9, 17$, and generators $4, 5, 6, 8$ in the CIGRE network. These FSPs varied in offered capacity, S from 0.03MVA to 1.5MVA. The step-size of $\delta p = 0.1$MW, $\delta q = 0.2$MVAR led to approximately 1500 pixels while neglecting 3 FSPs of $0.03, 0.03$, and 0.04MVA. The voltage constraints were $v_{\min} = 0.95 p.u.$, $v_{\max} = 1.05 p.u.$.

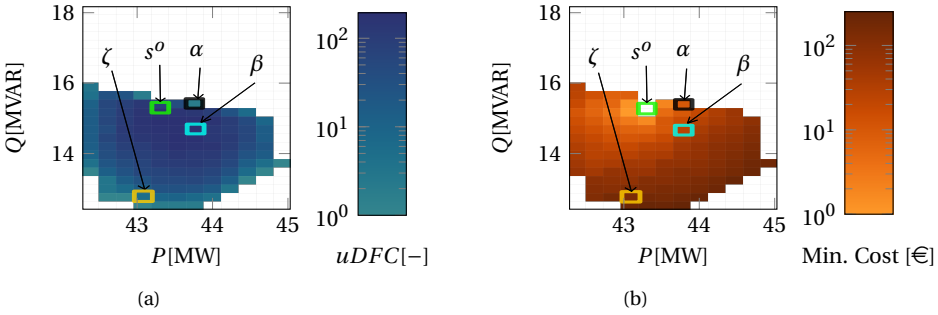


Figure 3.11: Selecting safe and cheap shifts with the unnormalized DFC (uDFC) in Fig.3.11(a). Additional information on shift minimum costs (Min. Cost) in Fig.3.11(b). Possible shifts include α (▭); very cheap, but not safe, β (▭); cheap and safe, and ζ: (▭) expensive and not safe.

Study E scenarios considered visually different FAs for the same FSPs between initial and altered OCs. The FSPs in the CIGRE network case were loads $3,5,6,9,17$ and generator 8, with $\delta p = 0.2MW, \delta q = 0.2MVAR$. For the altered OCs, the power factor of loads $0,7,8,12$ and 15 was reversed for the CIGRE network. In OB0, shifts were randomly sampled for all non-FSP generators and loads using normal distribution centered at the expected OC's values, with a standard deviation of 0.2. The FSPs in the OB0 case were loads $18,22$ and generators $26,50$, with $\delta p = 0.068MW, \delta q = 0.068MVAR$. The voltage constraints were $v_{\min} = 0.95p.u.$, $v_{\max} = 1.05p.u.$.
The algorithm's output is an FA Pandas data frame. The times referenced correspond to estimating the data frame. Simulations were performed on an Intel Core i7-1185G7 CPU with 16 GB RAM and an NVIDIA A100 GPU with 40GB VRAM. The GPU is available in Google Colab [97]. The algorithm's implementation in Python also included the PandaPower, SciPy, PyTorch, scikit-learn, and Numpy libraries.

3.4.1. DFC IMPROVING TSOS FLEXIBILITY SHIFT SELECTION

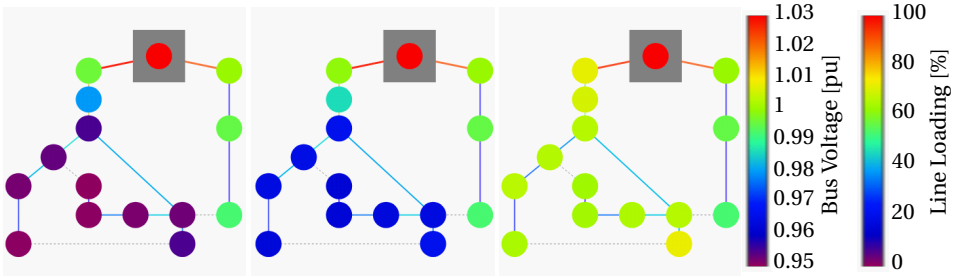
This case study exemplifies the improvement in flexibility shift selection using DFC. Fig.3.11 illustrates the simulation results on the CIGRE network, and Tab.3.1 summarizes the results for shifts α, β, ζ. Fig.3.12 illustrates the PF results for the cheapest flexibility shift combinations reaching α, β, ζ. The α shift's results are unsafe as multiple buses have approximately $0.95pu$ voltage magnitude, and only 1.1% of the available combinations for α are feasible. The few combinations available for ζ make this shift's combinations expensive, and less reliable (dependent on specific FSPs). The 193 feasible combinations for β make this selection less dependent on specific combinations. In addition, β has safer results for the network constraints than α and costs cheaper than ζ. Thus, using DFC, the TSOs can select safe and cost-efficient shifts.

3.4.2. ANALYSING TENSORCONVOLUTION+ PERFORMANCE

This case study visually and quantitatively evaluates the TensorConvolution+ performance in terms of FA DFC, range, and computational speed.

Table 3.1: Comparing shifts α, β, ζ with unnormalized DFC (uDFC), number of not-feasible combinations (NFC), feasible combinations percentage (FCP), and minimum cost(Min.Cost).

Shift	uDFC [-]	NFC[-]	FCP[%]	Min. Cost [€]
α	2	174	1.1	37.9
β	193	0	100	66.5
ζ	5	0	100	177.5



(a) PF for α. (b) PF for β. (c) PF for ζ.

Figure 3.12: PF results for cheapest flexibility combinations for shifts α, β, ζ, where network voltage constraints are $0.95 - 1.05 p.u.$

PERFORMANCE ANALYSIS FOR DFC

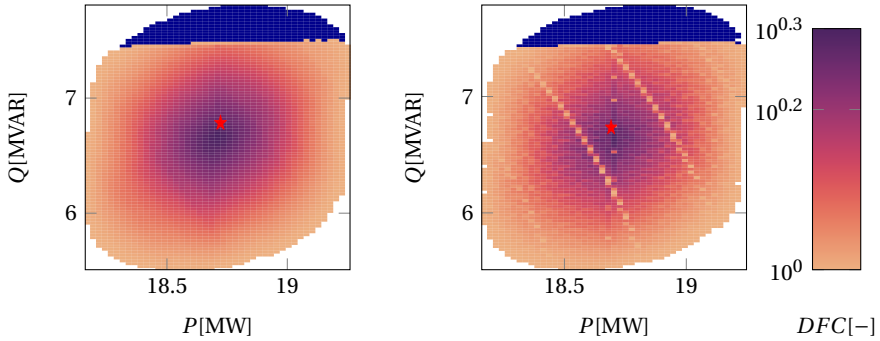
Fig.3.13 illustrates the resulting FA of the proposed approach and the ground truth, an exhaustive PF-based approach on the radial version of OB0. The root mean squared error between all pairs of pixels between Fig.3.13(a) and Fig.3.13(b) is 0.13, validating the observable high similarity between the two FAs.

To analyze the computational times using the CPU, the exhaustive PF-based approach needed 3 hours, 39 minutes, and 5 s for 480702 PF simulations. Increasing the number of FSPs increases the number of PF simulations needed largely. For this case study, when adding 1 more FSP, the number of PF simulations needed for the exhaustive PF-based approach is 1300032. This number highlights the challenge the proposed approach addresses. The proposed algorithm needed $42.8 s$, a speed-up of over 300 times.

PERFORMANCE ANALYSIS FOR FLEXIBILITY RANGE

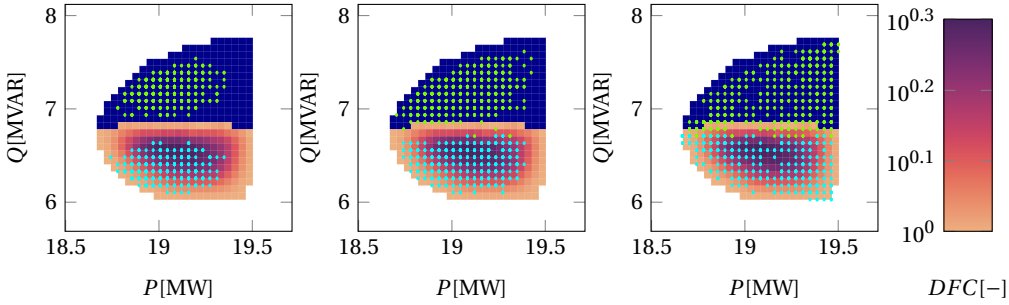
Tab.3.2 summarizes the average evaluation metrics of the scenarios on meshed OB0 and OB1 for narrow and wide voltage constraints. The metrics Ac_f and Ac_r are the mean a_f and a_r over the estimated FAs. These results show a great performance of TensorConvolution+ for the FA range. Fig.3.14 illustrates an example alignment of each distribution's samples with the algorithm's FA estimation, as used to estimate a_f and a_r for each of the 360 FAs.

Multiple cases required more than the available RAM in the narrow voltage constraints OB0 estimations for 6, 7 FSPs. In those cases, TensorConvolution+ aggregated at least one pair of FSPs to one FSP for at least one network component to reduce the tensor



(a) Proposed approach. (b) Exhaustive PF approach.

Figure 3.13: The FA for TensorConvolution+ in Fig.3.13(a) and PF-based approach in Fig.3.13(b). The FAs include the DFC ( ) for feasible shifts from the initial OP (★) and not-feasible shifts ( ).



(a) Kumaraswamy distribution. (b) Uniform distribution. (c) 'Hard' distribution.

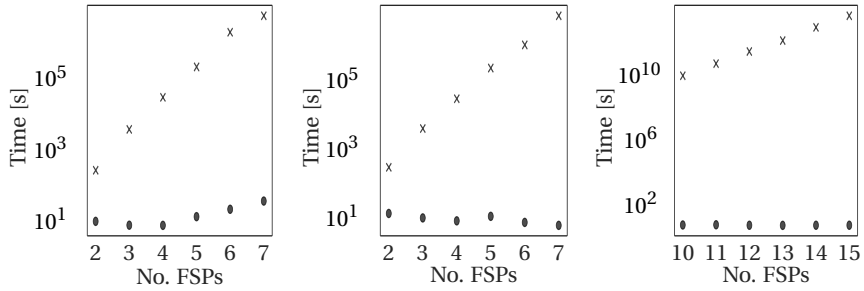
Figure 3.14: Monte Carlo-based results from different distributions aligned with the TensorConvolution+ output. Feasible TensorConvolution+ output pixels ( ). Not feasible TensorConvolution+ output pixels ( ). Feasible Monte Carlo samples ( ). Not feasible Monte Carlo samples ( ).

dimensions.

Fig.3.15 shows the computational time to estimate the FA for different FSP numbers. The authors multiplied the number of PF simulations needed for the exhaustive algorithm with the time the GPU spent for 1 PF simulation (0.038s for OB0 and 0.039s for OB1) to estimate the average time the baseline would need to estimate the 360 FAs. The exhaustive PF-based algorithm would be intractable for the majority of the FAs. The average time needed for the TensorConvolution estimations of Fig.3.15 varied between 6–36s and the average time over the 360 estimations was 11s. For a high resolution of 10000 pixels and 11 FSPs in the CIGRE network, the proposed algorithm needed 50s to estimate the FA. The maximum number of FSPs impacting a component for the CIGRE network's FA was 4 due to the radial network structure. All estimations were performed on the GPU.

Table 3.2: Evaluation of Algorithm for OB0, OB1 Networks (Net.), with narrow and wide voltage constraints (VC).

			Hard		Uniform		Kumaraswamy	
VC	**Net.**	$\lvert\Omega^{\mathbf{FSP}}\rvert$	$Ac_f\%$	$Ac_r\%$	$Ac_f\%$	$Ac_r\%$	$Ac_f\%$	$Ac_r\%$
Narrow	OB0	2–7	99	99	99	100	99	99
Narrow	OB1	2–7	99	99	100	100	100	100
Wide	OB0	10–15	99	99	99	99	98	98



(a) OB0 with narrow voltage constraints.

(b) OB1 with narrow voltage constraints.

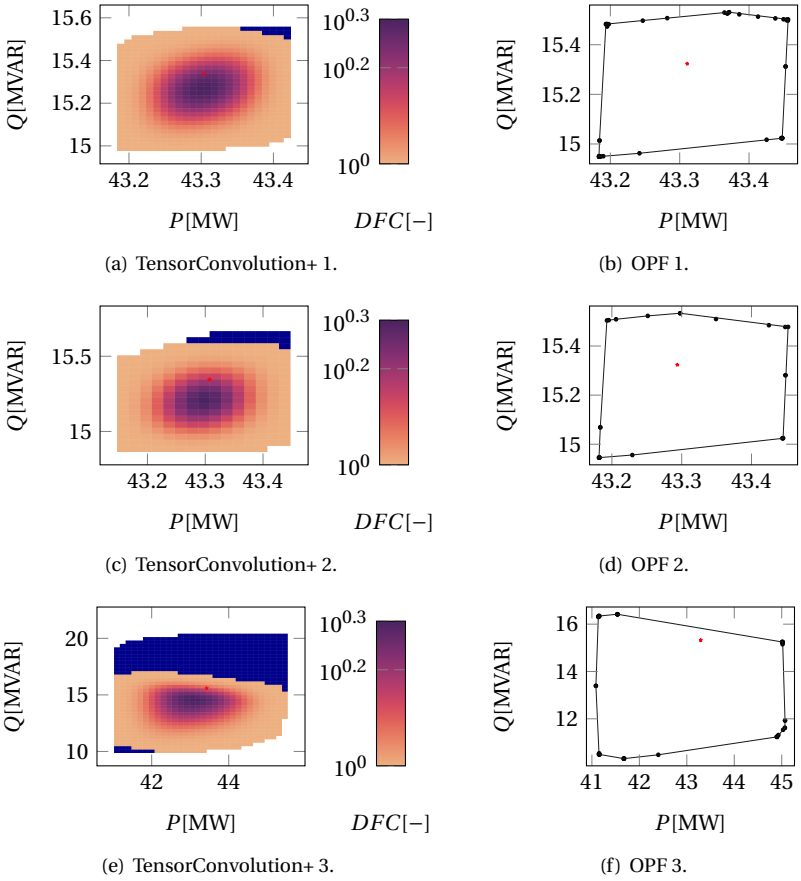
(c) OB0 with wide voltage constraints.

Figure 3.15: Computational time of TensorConvolution+ (⬮) and estimated computational time for exhaustive PF-based approach (X) per FSP amount.

COMPARISON WITH OPF-BASED METHOD

An OPF-based algorithm was employed to compare with the TensorConvolution+ estimations. The OPF-based algorithm employed weighted-sum multi-objective optimization with objectives $\max(\theta\Delta p^o + (1-\theta)\Delta q^o)$, $\max(-\theta\Delta p^o + (1-\theta)\Delta q^o)$, $\max(\theta\Delta p^o - (1-\theta)\Delta q^o)$, and $\max(-\theta\Delta p^o - (1-\theta)\Delta q^o)$. For each objective, θ varied between 0–1 with step size as a hyper-parameter. Due to convergence issues, the transforming loading limitations were ignored in the OPF algorithm. The FSP flexibility in this section was considered square where the active and absolute reactive power were between 0 and the nominal power of the FSP.

TensorConvolution+ and the OPF-based algorithm were compared in three cases on the CIGRE network. The cases varied in resolutions and network sensitivity to constraints. First, a case with FSPs the loads 14, 16 and generators 2, 4, 6. The first case resolution was $\delta p = 0.01MW, \delta q = 0.02MVAR$ ($\approx$ 690 pixels) and a θ with 0.1 increments for the OPF-based algorithm (44 OPFs). The second case included FSPs the loads 14, 16 and generators 2, 4, 5, 6. The resolution for the second case was $\delta p = 0.02MW, \delta q = 0.04MVAR$ ($\approx$ 300 pixels), and a θ of 0.2 increments (24 OPFs). The third case included FSPs the loads 3, 5, 6, 17, and the generator 8. The third case resolution was $\delta p = 0.15MW$, $\delta q = 0.3MVAR$ ($\approx$ 1050 pixels), and a θ with 0.1 increments (44 OPFs). The results of Fig.3.16 illustrate the FAs obtained for the three cases. The two algorithms agree on the shape of the FAs. However, the TensorConvolution+ algorithm also explores the inner area feasibility and estimates the DFC. Regarding computational burden, using the GPU for



(a) TensorConvolution+ 1.

(b) OPF 1.

(c) TensorConvolution+ 2.

(d) OPF 2.

(e) TensorConvolution+ 3.

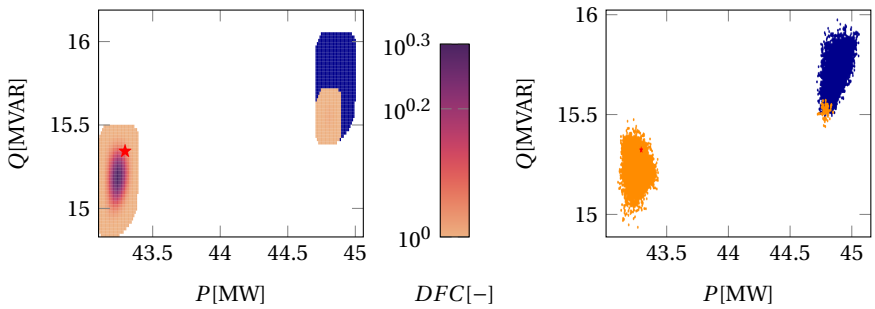
(f) OPF 3.

Figure 3.16: Comparing FAs from TensorConvolution+ and OPF based algorithm.

cases 1,2,3, TensorConvolution+ required $9s, 5.4s, 23.3s$, and the OPF-based algorithm required $37.6s, 21s, 36.3s$ respectively.

3.4.3. ESTIMATING DISCONNECTED FLEXIBILITY AREAS

This case study showcases the proposed algorithm's capability to estimate disconnected FAs and to deal with non-linear FSPs. Fig.3.17(a) illustrates the proposed algorithm's result compared to the result from 10000 samples of Fig.3.17(b) using the Monte-Carlo-based algorithm with "Hard" distribution. The PF-based algorithm was not capable of effectively exploring the disconnected areas, as the exact limits of feasible areas are unclear. The proposed algorithm was capable of estimating the FA, the range, and DFC in $8.5s$ using the CPU or $7.4s$ using the GPU.



(a) TensorConvolution+. (b) Monte-Carlo-based.

Figure 3.17: Disconnected FA Predicted by TensorConvolution+ and the Monte-Carlo-Based algorithms.

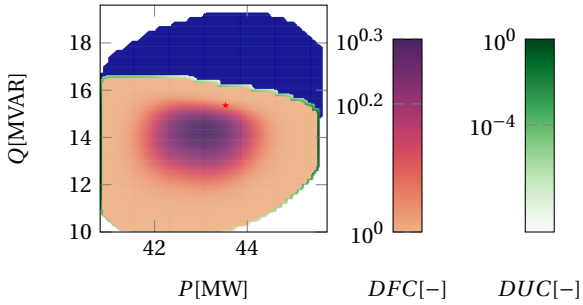
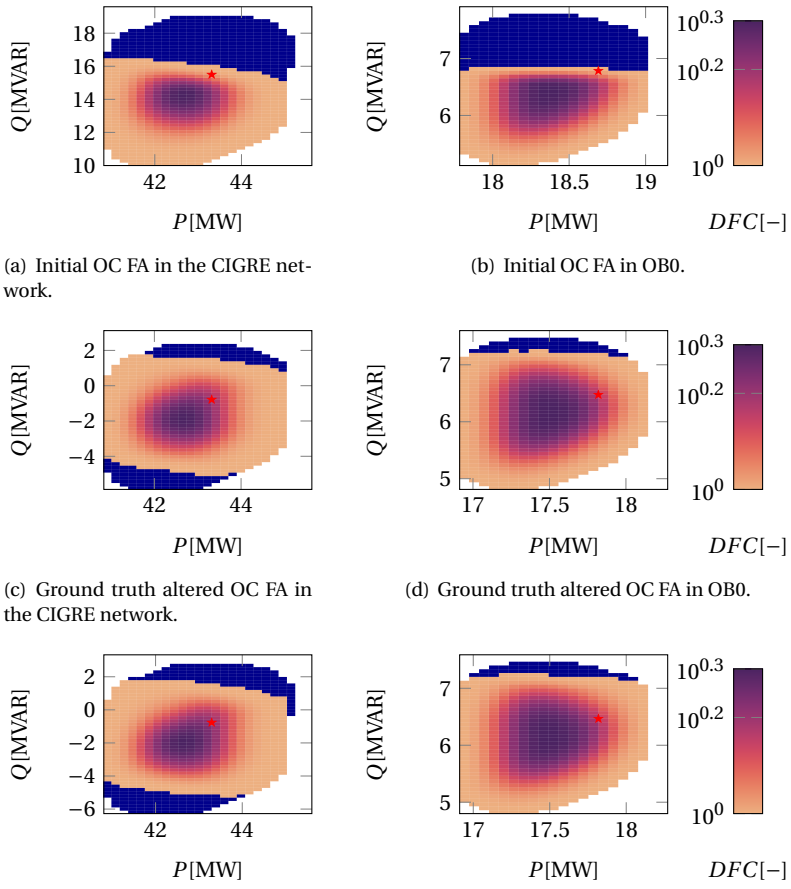


Figure 3.18: FA with DFC for feasible combinations ( ), density of uncertain combinations (DUC) for feasibility ( ), and not feasible FSP shift combinations ( ) from initial OP (★).

3.4.4. INCLUDING FLEXIBILITY FROM SMALL FSPS

This study shows the proposed algorithm's capability to include contributions from small FSPs. The increments $\delta p, \delta q$ between FSP shifts determine the FA resolution. In cases with large differences in the flexibility between the larger and smaller FSPs, reducing the resolution is impractical. Thus, the algorithm initially neglects all FSPs offering flexibility lower than $\delta p, \delta q$ and estimates the FA from the rest. Then, using bi-linear spline interpolation increases the resolution of the estimated FA to add the smaller FSPs. The algorithm convolutes the enhanced FA with the previously neglected FSPs' shifts. The additional area obtained in the last result constitutes the uncertain FA (not tested for the network constraints). In this study's results, after neglecting 3 small FSPs, the algorithm increased the pixels of the evaluated area 5 times and aggregated the previously neglected FSPs. Fig.3.18 visualizes the resulting FA. The uncertain FA addition process caused 11.09s delays using the GPU, with the rest of the FA estimation process needing 53.2s.



(a) Initial OC FA in the CIGRE network.

(b) Initial OC FA in OB0.

(c) Ground truth altered OC FA in the CIGRE network.

(d) Ground truth altered OC FA in OB0.

(e) Adapted altered OC FA in the CIGRE network.

(f) Adapted altered OC FA in OB0.

Figure 3.19: Adaptability of FAs from initial OCs to altered OCs.

3.4.5. ADAPTING FAS FOR PARTIALLY OBSERVABLE OCS

This case study showcases the algorithm's capability to adapt FAs for altered partially known OCs. For the CIGRE network, the algorithm used the tensors computed for Fig.3.19(a), and the partial initial OCs of Fig.3.19(c) to estimate the FA in Fig.3.19(e). The OC change from Fig.3.19(a) to Fig.3.19(c) increased the CIGRE network's buses' sensitivity to the over-voltage constraints. For OB0, the algorithm used the tensors computed for Fig.3.19(b), and the partial initial OCs of Fig.3.19(d) to estimate the FA in Fig.3.19(f). The OC change from Fig.3.19(b) to Fig.3.19(d) reduced the OB0 component sensitivities to FSP shifts. The adapted areas of Fig.3.19(e) and Fig.3.19(f) approximate the ground truths of Fig.3.19(c) and Fig.3.19(d). However, the adapted FAs required the initial OCs from limited components; 13 components in the CIGRE network, and 11 in OB0.

Using the GPU, estimating Fig.3.19(a) and Fig.3.19(b) and storing the tensors required

286s and 636s. The ground truths of Fig.3.19(c) and Fig.3.19(d) required 13.2s and 17.6s respectively. The adapted areas of Fig.3.19(e) and Fig.3.19(f) required 6.8s and 8.3s, respectively. Adapting FAs using the computed values for prior FAs is a capability absent in existing algorithms. The required computing time for TensorConvolution+ can be reduced to approximately half through this capability.

3.4.6. DISCUSSION

The proposed algorithm can evaluate the feasibility for the network constraints for all combinations between the FSP shifts. The studies show that this evaluation can benefit system operators in selecting shifts. Shifts with higher DFC offer more feasible flexibility options for system operators. Bigger DFC relates with a higher margin from the network constraints (Fig.3.12) and lower costs (Tab.3.1). The results show that TensorConvolution+ is the most computationally efficient algorithm for FA estimation with DFC among the algorithms studied. Simulations between the exhaustive PF-based alternative showed that the proposed algorithm could improve the computational speed over 300 times for 2 FSPs. The speed improvement increases with FSPs (Fig.3.15) as the required PFs for the exhaustive approach are (3.2). Most existing algorithms do not evaluate all flexibility combinations but focus on estimating the FA range. The proposed algorithm has high confidence in estimating the FA range, with an average $Ac_f = 99\%$. Identifying which FA regions are feasible and which are not is important for system operators to adopt FAs. Existing algorithms can have issues estimating disconnected FAs, causing estimation delays or incorrectly assuming in-between regions as feasible [19]. The proposed algorithm makes use of the impulse response property of convolution (Sec.3.3.1) to estimate disconnected areas with low computational burden. The results show that the algorithm can estimate disconnected FAs in 7.4s with distinct feasible regions (Fig.3.17). The results demonstrated that the proposed algorithm perform for meshed and radial network topologies, which was a challenge for existing algorithms.

The parameters $\delta p, \delta q$, and c_v, c_l influence the proposed algorithm's performance. Lowering $\delta p, \delta q$, increases the FA resolution and computational burden. For a total offered flexibility P_{tot}, recommended $\delta p = 0.05 \cdot P_{tot}, \delta q = 0.1 \cdot Q_{tot}$ result to approximately 400-pixel FA. Lowering c_v, c_l increases the complexity of (3.7)-(3.8). Recommended values of $0.001 \geq c_v \geq 0.0001, 1 \geq c_l \geq 0.1$ are based on the initial OC minimum margins from the constraints. The proposed algorithm's main limitation is the memory usage to store all flexibility combinations in tensors. As FSPs and constraint-sensitive DN components increase, this limitation becomes more notable. Aggregating FSPs for the components requiring unavailable memory can reduce this limitation. Alternatively, using memory-efficient software can mitigate the limitation. Another limitation concerns the need for complete network observability for (non-adapted) FA estimation.

3.5. CONCLUSION

This chapter develops an approach to estimate the flexibility of distribution systems for TSO-DSO coordination. This approach has a near-term practical value for power system operators. This chapter introduces applying tensors and convolutions for the flexibility estimation task, utilizing their useful properties. The proposed density of feasible

combinations (DFC) improves deciding on operating points that guarantee higher flexibility. The proposed algorithm makes computing and identifying these operating points tractable. The tractability is realized by minimizing the required PF simulations, applying convolutions, and using tensors. Convolutions aggregate flexibility from FSPs, including DFC. Tensors store the impacts of FSP shift combinations to determine the combination feasibility. Moreover, the algorithm applies the Dirac function to represent discrete FSPs. This application allows estimating disconnected FAs with low computational burden.

Case studies on 15−, 70−, and 109−bus systems show the proposed algorithm's performance on meshed and radial networks for connected and disconnected FAs. The results show high numerical confidence in the FA range and DFC. The algorithm estimated the FA around 300 times faster than the alternative approach. The algorithm's average duration of 11 s renders the approach promising for further development toward near-real-time TSO-DSO coordination.

Future work can investigate dealing with limited observability in distribution systems. The investigation will explore deep learning-based approaches or techniques from system identification that may advance this flexibility estimation approach. Subsequent work will pursue further improvement of the proposed approach's memory efficiency. Future research will also explore algorithms that can adapt to changing network topologies. Changing network topologies alters the impacts of FSPs on network components. Thus, approximating the impact alterations should allow adapting FAs for changing topologies.

4

SELECTION FOR FLEXIBILITY AREAS USING PROBABILISTIC MACHINE LEARNING UNDER MEASUREMENT UNCERTAINTY

What we observe is not nature itself, but nature exposed to our method of questioning.

Werner Heisenberg

Coordination between transmission system operators (TSOs) and distribution system operators (DSOs) can support TSOs in using distribution system (DS) flexibility while ensuring feasible operation. Flexibility areas can support TSO-DSO coordination, aggregating the total feasible flexibility within the DS. However, existing real-time estimation approaches do not consider the limited measurements within DS. This chapter proposes a Bayesian neural network (BNN) to estimate the operating conditions that bound the operational flexibility, including epistemic and aleatoric uncertainties. These uncertainties stem from the limited real-time measurements in DSs and the measurement noise. TSOs can select a threshold that confirms a probability of safety, considering uncertainty margins. The chapter also provides FA estimation in DS topologies with 2 points of common coupling (PCC) with the transmission system. Case studies in the CIGRE and Oberrhein networks compare the proposed BNNs to baseline statistic-based approaches for forecast and measurement uncertainty in FAs. The case studies show the proposed FA estimation

Parts of this chapter are in: **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"Selection for Flexibility Areas using Probabilistic Machine Learning Under Measurement Uncertainty"*, Under Review, 2025.

under various safety margins and systems with 2-PCC. Case studies also assess various measurement noise levels and evaluate model performance for different DS topologies.

4.1. INTRODUCTION

Flexibility from service providers (FSPs) in DSs can ensure a resilient and efficient transmission system (TS) and DS operation. FSPs can modify their active and reactive power net injection upon request by the TSOs or DSOs to avoid or address TS or DS issues, such as congestion.

Aggregated DS flexibility areas (FAs) can facilitate the coordination between DSOs and TSOs for DS flexibility. DSOs can inform TSOs about the total feasible flexibility available within DSs using FAs, without requiring the exchange of details on DS topology, and operating conditions. A flexibility combination is feasible if the resulting operating condition fulfills all power flow constraints in the DS, and line loading and bus voltages do not exceed rated values. By summarizing the capabilities of FSPs within DS, FAs enable TSOs to access the flexibility needed for balancing supply and demand, managing congestion, and enhancing system stability at the transmission level while ensuring DS feasibility. This TSO-DSO coordination approach simplifies the DSO-TSO interaction, ensures data privacy, and reduces the complexity of operational coordination while still using the potential of DS flexibility to support grid reliability and efficiency.

Existing FA estimation approaches mainly require the operating condition (OC) of the DS as input, as [21]–[25]. However, DSs can have limited real-time observability with measurement units on limited network components [1], [27]–[29], e.g. 5–40% in medium voltage (MV) DS [98]. Thus, DSOs might not know the exact physical feasibility margins from all DS constraints due to this limited observability. These margins are important for DSOs to evaluate the feasibility of flexibility combinations. Although some existing approaches address forecasting uncertainty in day-ahead FA estimations [18], [20], [30], existing FA approaches neglect the limited observability in real-time FA estimation.

This chapter proposes a new approach for FA estimation, using probabilistic BNNs to assess the risk for non-feasible flexibility due to limited DS observability and DS constraint approximation error. BNNs extend conventional neural networks by including uncertainty estimation in the predictions [99], [100], capturing inherent uncertainty from the available measurements and the uncertainty in model parameters. These probabilistic estimates indicate the model prediction confidence and problem uncertainty, which can be important for robustness in tasks involving critical infrastructure, such as the power grid. The main tractable approaches in BNNs include variational inference [77], Monte Carlo dropout [76], and deep ensembles [78]. The proposed approach enables the TSOs to select the safety percentage in the resulting area. A lower safety returns a larger FA for TSOs to select an operating point, whereas a higher safety better secures feasibility on the selected operating point. Nevertheless, the proposed approach also includes the density of feasible combinations for each operating point and the given safety. Thus, for a low-safety FA estimation, a higher-density operating point provides more reliability in finding alternative flexibility combinations.

FA estimation approaches apply steady-state simulations, typically using power flows (PF) [20], [24], or optimal power flows [18], [22], [23]. PF-based approaches typically

apply Monte-Carlo sampling for a predetermined number of combinations. OPF-based approaches apply different multi-objective optimization (MOO) techniques. These approaches typically rely on an initial DS forecast or approximated state to evaluate the feasibility of different flexibility combinations. To deal with uncertainties from renewable sources in aggregated flexibility area estimation, [30] applied a chance-constrained method, and [18], [101] applied robust optimization. [20], [102] included the probability distributions of forecast errors to determine the probability of feasibility, and [103] focused on forecast uncertainty generation. The approaches of [18], [20], [30], [101]–[103] consider the forecast uncertainties rather than real-time observability uncertainty and map these uncertainties to the network constraints using the power-flow relationships. The proposed approach approximates the network constraint margin uncertainty using BNNs, considering the real-time measurements and limited DS observability.

Existing FA estimation approaches rely on TS-DS interconnections with a single point of common coupling (PCC) [18], [20], [22]–[24]. However, DS ring topologies with 2 PCCs are increasingly adopted and can improve the DS losses [104]. The main challenges for estimating FA of TS-DS connected with multiple PCCs are the dependencies between the PCCs [38] and the complexity of representing these dependencies. This chapter provides an approach to estimate and represent network flexibility with 2 PCCs.

Distribution system state estimation (DSSE) algorithms consider the challenge of limited observability in DSs by estimating the system's state using available measurements and pseudo-measurements. These algorithms primarily estimate nodal voltage magnitudes and angles [14], [28], [105], [106], though some also estimate nodal active and reactive powers [107] or additional metrics as line loading [108]. DSSE approaches applied Bayesian models for non-Gaussian pseudo-measurement uncertainties [107], to fuse measurements with varying sampling rates [105], or to generate data samples to train deep learning-based DSSE models [106], [109]. Bayesian models for DSSE often model uncertainties but typically provide a single expected state estimate [105], [107]. These DSSE approaches generally focus on estimating voltage magnitudes and angles. The proposed FA estimation approach applies BNNs to estimate the probabilistic DS constraint margins, accounting for the available measurements to provide a FA that is constrained by the safety margins selected by the TSO. The proposed approach also includes real-time interface (RTI) measurement considerations. RTI is a platform in the Netherlands that mandates DER with a capacity higher than 1MW to provide real-time measurements to system operators.

This chapter develops an approach for FA estimation with BNNs to consider the estimation and DS observability uncertainties. This chapter analyses the performance of BNN approaches considering the impact of network sizes, assumptions about noise levels, and data distributions for the power system task. The proposed FA approach also includes an algorithm for networks with 2 PCC. Therefore, the main contributions are:

1. FA estimation approach that considers the uncertainty for the DS constraint margins due to limited DS observability and estimator uncertainty.

2. BNN model structures to estimate FAs considering real-time uncertainties and data distributional changes.

3. Approach approximating and representing FAs in networks with 2-PCCs.

The case studies use the CIGRE 15−bus MV network and Oberrhein 70−, 109−, and 179−bus MV networks. The case studies compare the proposed BNN model to alternative approaches for FA uncertainty, and demonstrate the proposed approach for FA estimation in DS with 1 and 2 PCCs.

4.2. Limited Observability in Steady State Flexibility Area Estimation

4.2.1. Problem Introduction

Let a DS state described by $X(t) \in \mathrm{R}^{n\times 2}, V(t) \in \mathrm{R}^{n}, L(t) \in \mathrm{R}^{k}$ at time t. The matrix $X(t)$ is the active and reactive power injected at each of the n nodes, $V(t)$ is the voltage magnitude at each n, and $L(t)$ is the loading at each of the k lines and transformers. The FSPs can linearly modify $X(t)$ by increasing or reducing the injection on a subset of the DS nodes $\hat{n}$. However, the impact on the network component loading and voltage is non-linear. Thus, a combination of flexibility shifts from FSPs at t is $U(t,\tau) \in \Omega^{U}(t,\tau) \subset \mathrm{R}^{\hat{n}\times 2}$, where $\Omega^{U}(t,\tau)$ is the set of all flexibility combinations at t, and τ is the duration until the new state is achieved. Applying $U(t,\tau)$ would result in a new state after τ as:

$$X(t+\tau) = X(t) + B \cdot U(t,\tau), \tag{4.1}$$
$$V(t+\tau) = V(t) + F^{v}(X(t), V(t), U(t,\tau)), \tag{4.2}$$
$$L(t+\tau) = L(t) + F^{l}(X(t), L(t), U(t,\tau)), \tag{4.3}$$

where $B \in \{0,1\}^{n\times \hat{n}}$ is a binary matrix with 1 on each FSP's corresponding node and 0 otherwise, e.g. if only the 5^{th} FSP is connected on the 1^{st} node, the first row of B will be 1 only on the fifth element. $F^{v}(\cdot), F^{l}(\cdot)$ are the non-linear functions of flexibility impacts on network component voltage and loading, respectively.
A network includes voltage constraints $c_v^{\min}, c_v^{\max}$, and loading constraints $c_l^{\max}$. The objective of a steady-state FA estimation algorithm is to describe the feasible area (considering DS constraints) of active and reactive power exchange at π PCCs between the TSO and DSO, given all available flexibility combinations $\Omega^{U}(t,\tau)$. A single PCC's active and reactive power exchange is $x_{PCC}(t)$. As shown in [1], [110], an additional dimension for the density of feasible combinations can further demonstrate the reliability in achieving a PCC operating point. Reliability in terms of a plethora of alternatives to achieve said operating point. Thus, the FA objective is to identify $Y(t+\tau) \subset \mathrm{R}^{\pi\times 3}$, as:

$$\begin{aligned} Y(t+\tau) = \Big\{ & y(t+\tau) = F^{A}\big(X(t), L(t), V(t), U(t,\tau)\big) \big|, \\ & U(t,\tau) \in \Omega^{U}(t,\tau), c_v^{\min} \le V(t+\tau) \le c_v^{\max}, \\ & L(t+\tau) \le c_l^{\max} \Big\}, \end{aligned} \tag{4.4}$$

where $y(t+\tau) \in \mathrm{R}^{\pi\times 3}$ describes the active power, reactive power, and density of feasible combinations (DFC) [110] for each PCC operating point. Function $F^{A}(\cdot)$ applies each $U(t,\tau) \in \Omega^{U}(t,\tau)$ on (4.1)-(4.3) and aggregates the feasible combinations for each $x_{PCC}(t+\tau)$ as [110].

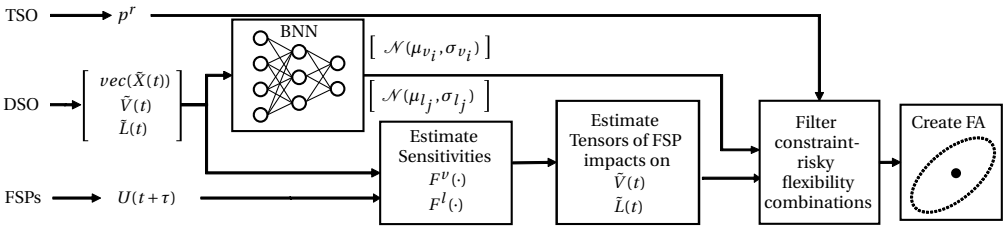


Figure 4.1: Overview of proposed FA estimation process.

4.2.2. FA UNCERTAINTY

Existing approaches mainly apply (4.1)-(4.3) by solving the power flow relationships. However, DSOs do not deterministically know $X(t), V(t), L(t)$ to estimate the FAs. Depending on the intended FA estimation speed, the $X(t), V(t), L(t)$ source for uncertainty can be:

1. In the day-ahead estimation of FAs, DSOs forecast the DS operating conditions, which can differ in real-time.

2. In the real-time estimation of FAs, DSOs only measure a subset of the network components.

Addressing the day-ahead FA estimation forecast uncertainty, existing approaches [18], [20], [30] include the standard deviation in DER injection forecasts for $X(t)$ to provide robust or chance constrained FAs. However, for limited observability in real-time, existing approaches neglect the resulting FA error [1]. This research addresses the impact of real-time limited DS observability on the FA estimation task.
FA estimation uncertainty can be split into aleatoric and epistemic. Aleatoric uncertainty is inherent in data and cannot be reduced, e.g., measurements and pseudo- measurements have noise that causes uncertainty. Epistemic uncertainty relates to the model's structure, parameters, or assumptions due to limited data. Different data distributions and patterns can be absent in the model development, however appear in model deployment. Existing FA estimation approaches for forecasts focus on aleatoric uncertainty. Considering this uncertainty, a safety margin for the DS loading and voltage constraints can alleviate any potential impact from the miscalculation of the initial DS conditions.

4.2.3. PROPOSED FA ESTIMATION APPROACH

Fig.4.1 shows the proposed FA estimation. The DSO provides the measured and pseudo-measured DS variables and the percentage p^r for safety margin. FSPs provide their offered flexibility, to obtain the $U(t,\tau)$. The proposed approach estimates $F^v(\cdot), F^l(\cdot)$ and the tensors of flexibility impacts on the constrained voltage and loading variables using the modified TensorConvolution+ algorithm of [110]. The BNN estimates the distributions of the constrained DS voltage and loading variables using the available measurements and pseudo-measurements. Using these distributions, the p^r, and the estimated flexibility impact of each FSP combination, the proposed approach evaluates the feasibility of each flexibility combination. The approach then aggregates the filtered feasible

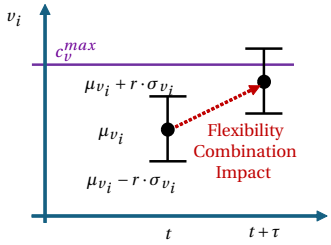


Figure 4.2: Feasibility evaluation example.

combinations as in [110] and returns the flexibility area. Thus, the proposed approach adds minor computational burden to [110], mainly for the BNN estimation. Fig.4.2 illustrates how the considered uncertainty and accepted safety by the TSO can influence the flexibility considered feasible. For the TSO-selected safety and uncertainty levels in Fig.4.2, the impact from the flexibility combination could result in over-voltage. Thus, in this example, the proposed algorithm would consider the combination non-feasible.
The proposed approach first assumes a local approximation of voltage and loading sensitivity to a flexibility action. If two DS states $\{\hat{X}(t), \hat{V}(t), \hat{L}(t)\}, \{\tilde{X}(t), \tilde{V}(t), \tilde{L}(t)\}$ are within a small ball of radius ϵ, then $\forall U(t,\tau) \in \Omega^U(t,\tau)$:

$$F^v(\hat{X}(t), \hat{V}(t), U(t,\tau)) \approx F^v(\tilde{X}(t), \tilde{V}(t), U(t,\tau)) \tag{4.5}$$

$$F^l(\hat{X}(t), \hat{L}(t), U(t,\tau)) \approx F^l(\tilde{X}(t), \tilde{L}(t), U(t,\tau)) \tag{4.6}$$

Thus, this local approximation assumption represents the sensitivity functions as similar for close DS states.
Let Ω^b be the set of DS buses excluding the PCC buses, Ω^λ the set of DS lines and transformers, i.e. $|\Omega^\lambda| = k$. Let $\bar{X}(t), \bar{V}(t), \bar{L}(t)$ be the expected network state values considering the measured and pseudo-measured variables, i.e., measured variables have lower noise. Considering the limited observability, any estimated values for voltage and loading $\bar{V}(t), \bar{L}(t)$ will have some mismatch $\varepsilon_v^e, \varepsilon_l^e$ from the actual voltage and loading values $V(t), L(t)$. The proposed probabilistic estimator represents this imperfect estimation and task's stochasticity by estimating the distribution for each bus voltage $v_i(t) \forall i \Omega^b$, and line or transformer loading $l_j(t) \forall j \in \Omega^\lambda$, given the measured and pseudo-measured values. Thus, the proposed probabilistic model estimates a a normal distribution for each $v_i(t), l_j(t)$ as $\mathcal{N}(\mu_{v_i}(t), \sigma^2_{v_i}(t)), \forall i \in \Omega^b, \mathcal{N}(\mu_{l_j}(t), \sigma^2_{l_j}(t)), \forall j \in \Omega^\lambda$, where $\mu_{v_i}(t), \sigma^2_{v_i}(t)$ are the model's expected value and standard deviation for the voltage magnitude of bus i at t, and $\mu_{l_j}(t), \sigma^2_{l_j}(t)$ are the model's expected value and standard deviation for the loading of line or transformer j at t. Using these standard deviations, the proposed approach applies the quantile function $H(\cdot)$ (inverse of cumulative distribution function) to select a safety level for the network constraints. The quantile function takes a probability and a distribution as inputs and returns the margin r. The probability of a sample being inside the margin r equals the input probability p^r. For example for $p = 84\%$, the $H(\cdot)$ of a normal distribution will return $r = 1$, for the margins $\mu_{v_i}(t) \pm r \cdot \sigma_{v_i}(t)$. Hence, 84% of samples from the input distribution would be within the $\mu_{v_i}(t) \pm \sigma_{v_i}(t)$. Through $H(\cdot)$, the TSOs can select the safety percentage for the FA, p^r%, being the probability of

the actual network state being within the estimated margins, modifying (4.4), (4.2), and (4.3) to:

$$\begin{aligned} Y(t+\tau) = \{ & y(t+\tau) = F^A(\bar{X}(t), \bar{L}(t), \bar{V}(t), U(t,\tau)|, \\ & U(t,\tau) \in \Omega^U(t,\tau), \\ & c_v^{\min} \leq \mu_{v_i}(t+\tau) - r \cdot \sigma_{v_i}(t) \forall i \in \Omega^b, \\ & \mu_{v_i}(t+\tau) + r \cdot \sigma_{v_i}(t) \leq c_v^{\max} \forall i \in \Omega^b, \\ & \mu_{l_j}(t+\tau) + r \cdot \sigma_{l_j}(t) \leq c_l^{\max} \forall j \in \Omega^\lambda \}, \end{aligned} \tag{4.7}$$

$$r = H(p^r), \tag{4.8}$$

$$\mu_{v_i}(t+\tau) = \mu_{v_i}(t) + F^v(\bar{X}(t), \bar{V}(t), U(t,\tau)), \forall i \in \Omega^b, \tag{4.9}$$

$$\mu_{l_j}(t+\tau) = \mu_{l_j}(t) + F^l(\bar{X}(t), \bar{L}(t), U(t,\tau)), \forall j \in \Omega^\lambda. \tag{4.10}$$

4.2.4. TWO-PCC FA ESTIMATION

Considering two PCCs in FA estimation would require aggregating feasible points in 4 dimensions, the active and reactive power per PCC. Visualizing these 4 dimensions can be challenging, while this increased dimensionality can worsen the implementation complexity or computational burden for the FA estimation. The proposed approach considers that the TSO can select and decouple the flexibility for active and reactive power. Alternatively, estimating two single-PCC FAs, one for each PCC can represent the problem while neglecting dependencies between the PCC operating points.

The proposed approach modifies $y(t+\tau)$ of (4.7) selecting the first two dimensions as (i) active powers of the two PCCs and (ii) reactive powers of the two PCCs. The third dimension is the density of feasible combinations. This approach initially runs power flows to estimate $F^v(\cdot), F^l(\cdot)$ as in Fig.3.3. However, the sensitivity matrices' (and impact Tensors') first two dimensions are the TSO's selection of active or reactive powers. If the TSO considers active and reactive power flexibility from the two PCCs, then the DSO can perform 2 simulations, selecting two of the 4 dimensions for each simulation. However, this selection will decouple the two sets of dimensions.

4.3. BAYESIAN NEURAL NETWORK BASED APPROACH

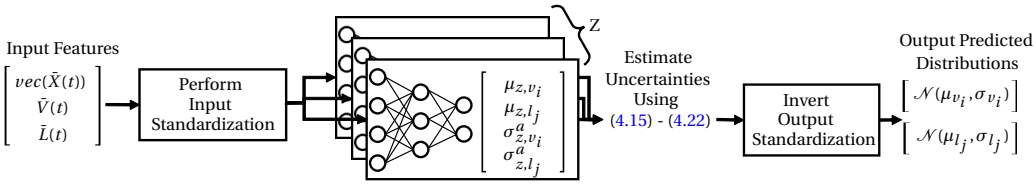


Figure 4.3: Overview of proposed BNN model output estimation process.

Figure.4.3 illustrates the proposed BNN model process in estimating the bus voltage magnitude and the line and transformer loading distributions. The BNN model estimates the distributions per datapoint Z times, each time with different model parame-

ters to approximate the epistemic uncertainty. For each parameter setting $z \in Z$ and input features $\chi = [vec(\bar{X}), \bar{V}, \bar{L}] \in \mathrm{R}^{3 \cdot n + k}$, the BNN estimates the output means μ_{z,v_i}, μ_{z,l_j} and aleatoric standard deviations $\sigma^a_{z,v_i}, \sigma^a_{z,l_j}$ for bus voltages and line or transformer loading. The loading variables have different scales and variability from the voltage and power injection variables. Hence, the proposed approach first performs standard scaling on all input features.

The proposed supervised BNN uses a training dataset D with $|D|$ data points. During training, the model processes the dataset D in batches. The model uses these batches to compute predictions, compare them to the true labels using a loss function, and update its parameters via backpropagation. Each datapoint $d \in D$ includes input features χ_d and targets $\phi_d = [v_i \forall i \in \Omega^b, l_j \forall j \in \Omega^\lambda] \in \mathrm{R}^{|\Omega^b|+|\Omega^\lambda|}$. $vec(\cdot)$ vectorizes an input matrix and $|\cdot|$ is a set cardinality.

Conventional supervised feed-forward neural networks (FNNs) return point predictions for the above targets ϕ_d given the input features χ_d, and trained parameters W. These FNNs approximate the underlying input-output relationship using the available training data. However, these point predictions are also impacted by the inherent noise within the training data and the data distributions, which might differ from the real-life applications. Therefore, understanding what the models do not know can be critical for ML [111].

BNNs consider the aleatoric and epistemic uncertainties in their estimations. To consider these uncertainties, BNNs do not make point predictions but rather predict distributions of possible outputs for the given inputs:

$$p(v_i|\chi, D) = \int p(v_i|\chi, W) p(W|D) dW, \forall i \in \Omega^b, \tag{4.11}$$

$$p(l_j|\chi, D) = \int p(l_j|\chi, W) p(W|D) dW, \forall j \in \Omega^\lambda, \tag{4.12}$$

where $p(\cdot)$ is a probability distribution, $p(v_i|\chi, W), p(l_j|\chi, W)$ are the likelihoods of the outputs given the inputs and model parameters, $p(W|D)$ is the posterior distribution over the model parameters. The integral represents the marginalization over all possible model parameters W. The v_i, l_j in (4.11) and (4.12) are the values from the dataset targets, split to return one normal distribution per target.

4.3.1. EPISTEMIC UNCERTAINTY

To consider the Epistemic uncertainty, BNNs do not have a single setting of parameters but rather use different settings of parameters weighted by the posterior probabilities [100], the $p(W|D)$ term. As (4.11), (4.12) can be computationally challenging, main approaches such as variational inference[77], Monte Carlo dropout [76], and deep ensembles [78] approximate (4.11),(4.12). Variational inference uses a variational posterior $q(W|D)$ to approximate the true posterior $p(W|D)$, thus optimizing the parameters of the distribution $q(W|D)$. Monte Carlo dropout applies dropout during training and inference time. During inference, the BNN model estimates the output distributions for the same inputs multiple times, with the activated dropout resulting in different neurons being deactivated each time. Therefore, for Z estimations for the same inputs, Z

different parameter settings are used, approximating (4.11),(4.12) with:

$$p(v_i|\chi,D) \approx \frac{1}{Z}\sum_{z=1}^{Z} p(v_i|\chi,W_z), \quad W_z \sim q(W|D), \forall i \in \Omega^b, \tag{4.13}$$

$$p(l_j|\chi,D) \approx \frac{1}{Z}\sum_{z=1}^{Z} p(l_j|\chi,W_z), \quad W_z \sim q(W|D), \forall j \in \Omega^\lambda. \tag{4.14}$$

Deep ensembles train Z independent FNNs, resulting in Z different parameter settings, applying (4.13), (4.14).

4.3.2. ALEATORIC UNCERTAINTY

To consider and estimate aleatoric uncertainty, BNNs include standard deviation $\sigma^a_{z,v_i}, \sigma^a_{z,l_j}$ for each predicted output in the $p(v_i|\chi,W), p(l_j|\chi,W)$ terms of (4.11),(4.12), and (4.13),(4.14). The proposed BNN for FA estimation considers aleatoric uncertainty as heteroscedastic [111], i.e., dependent on each data point rather than being constant for all data, e.g., varying environmental factors can impact measurements and pseudo-measurements differently between prediction instances.

4.3.3. BNN MODEL OUTPUT DISTRIBUTION ESTIMATION

The mean value, and the epistemic, aleatoric, and total standard deviation per output are:

$$\mu_{v_i} = \frac{1}{Z}\sum_{z=1}^{Z} \mu_{z,v_i}, \quad \forall i \in \Omega^b, \tag{4.15}$$

$$\sigma_{v_i} = \sqrt{(\sigma^a_{v_i})^2 + (\sigma^e_{v_i})^2}, \tag{4.16}$$

$$\sigma^a_{v_i} = \sqrt{\frac{1}{Z}\sum_{z=1}^{Z} (\sigma^a_{z,v_i})^2}, \quad \forall i \in \Omega^b, \tag{4.17}$$

$$\sigma^e_{v_i} = \sqrt{\frac{1}{Z}\sum_{z=1}^{Z} (\mu_{z,v_i} - \mu_{v_i})^2}, \quad \forall i \in \Omega^b, \tag{4.18}$$

$$\mu_{l_j} = \frac{1}{Z}\sum_{z=1}^{Z} \mu_{z,l_j}, \quad \forall j \in \Omega^\lambda, \tag{4.19}$$

$$\sigma_{l_j} = \sqrt{(\sigma^a_{l_j})^2 + (\sigma^e_{l_j})^2}, \tag{4.20}$$

$$\sigma^a_{l_j} = \sqrt{\frac{1}{Z}\sum_{z=1}^{Z} (\sigma^a_{z,l_j})^2}, \quad \forall j \in \Omega^\lambda, \tag{4.21}$$

$$\sigma^e_{l_j} = \sqrt{\frac{1}{Z}\sum_{z=1}^{Z} (\mu_{z,l_j} - \mu_{l_j})^2}, \quad \forall j \in \Omega^\lambda. \tag{4.22}$$

(4.16), (4.17), (4.18), (4.20), (4.21), (4.22), use variances before estimating standard deviations, as variances can be additive for independent estimations. (4.17),(4.21) estimate

the aleatoric variances for each target as the average variance from the Z BNN output standard deviations [111]. (4.18), (4.22) estimate the epistemic variances for each target as the average variance of the mean predictions for each target distribution [111]. (4.16),(4.20) add aleatoric and epistemic variances to obtain the total variance [111].

4.3.4. MODEL TRAINING LOSS FUNCTION

As the BNN estimates a distribution per output, the negative log-likelihood (NLL) function measures how well the estimated distributions represent the actual distribution generating the data [78], [112], [113]. NLL considers the confidence in the aleatoric uncertainty, and the error between the predicted mean and the target values as:

$$L_{NLL}(\theta) = \frac{1}{|D|}\sum_{d=0}^{|D|}\left(\frac{\phi_d - \mu_{\chi_d}}{2\cdot(\sigma^a_{\chi_d})^2} + \frac{1}{2}log(\sigma^a_{\chi_d})^2\right) \tag{4.23}$$

The BNN outputs, $\mu_{\chi_d} = [\mu_{z,v_i} \forall i \in \Omega^b, \mu_{z,l_j} \forall j \in \Omega^\lambda]^T, \sigma^a_{\chi_d} = [\sigma^a_{z,v_i} \forall i \in \Omega^b, \sigma^a_{z,l_j} \forall j \in \Omega^\lambda]^T$ include voltage and loading variables that differ in scale and deviation. For VI, the loss can also include a regularization term on the Kullback–Leibler divergence. Different scales and variances for a subset of outputs can impact the loss function to focus on the subset that produces a higher numerical loss, e.g., a $0.5p.u.$ voltage mismatch is severe, whereas a 0.5% loading mismatch is minor. Hence, the proposed approach performs standard scaling in the target outputs during training, to ensure (4.23) is not impacted by the different output units. This scaling results in standardized (4.15)-(4.22). Thus, the proposed BNN output estimation (Fig.4.3) performs inverse standardization on the outputs to obtain $p.u.$ voltage and % loading scaling.

4.4. CASE STUDIES

The case studies used the CIGRE MV network, the networks from the separated Oberhein substations 0 (Ob0), 1 (Ob1), and the connected Oberrhein network (Ob) as visualized in Fig.4.4. All case studies considered the PCC buses, the HV-MV transformers, and the buses bellow the HV-MV transformers as observable (v, I, P, Q measurements). In the CIGRE network, the wind turbine with 1.5MW capacity would require RTI measurements. For Oberrhein networks, all buses connected to 3 lines or more were considered observable.

The case study in Sec.4.4.1 included the CIGRE network of Fig.4.4 in a meshed (closed switches) and radial topology (open switches) with 4 FSPs. For each FSP, Sec.4.4.1 considered all shifts $[\Delta P, \Delta Q]$ with steps 0.1MW constrained by the FSP's initial output. For each shift, 100 different DS operating conditions were sampled by changing the power injection $X(t)$ of all non-observable loads and generators with mean 0 and standard deviation $\sigma^{pm} \in \{1, 2, 5, 10, 20, 50\}$% for pseudo-measurement errors.

The case study in Sec.4.4.2 compares the proposed BNN model's performance to modified uncertainty evaluations from FA estimation literature. Appendix C includes the comparative analysis performed on BNN structures: a variational inference (VI), Monte Carlo dropout (MCD), and deep ensemble (DE) models. The MCD model had the highest consistency under different test set distributions.

For the baseline design, existing forecast-based evaluations map the uncertainty of nodal injections to the network constraints using the power flow relationships [20], [30], [102], [103]. For example, [30] assumed a mathematical linear relationship between power injection forecast errors and deviations in voltage magnitudes and angles. However, the [30] formula is rather complex, includes participation factors, and excludes the loading variables. The [20] approach runs multiple power flows for different forecasts to approximate the probability of feasibility. This case study approximates the constraint uncertainty mapping using power flows and data statistics. This statistical approximation considers a forecast DS operating condition per hour for each of the Fig.4.4 systems. For each hour forecast, the statistical approximation estimates 100 "actual" operating conditions with σ^{pm}, and ψ^{ld}, ψ^{pv} the load, and PV correlation, respectively. The average standard deviation from the 100 samples returns the $\sigma_{v_i} \forall i \in \Omega^b, \sigma_{l_j} \forall j \in \Omega^\lambda$ for each forecast condition. For each network, the approximated $\sigma_{v_i} \forall i \in \Omega^b, \sigma_{l_j} \forall j \in \Omega^\lambda$ used datasets with $\sigma^{pm} = 20\%, \psi^{ld} = 60\%, \psi^{pv} = 70\%$, like the BNN training datasets. The case study of Sec.4.4.2 also considers a second baseline, where the measured buses included $\sigma^m_\eta, \sigma^m_\eta$ power and loading measurement standard deviation, instead of σ^{pm}, to create a measurement-based statistical approach. The baseline test datasets for each network vary the $\sigma^{pm} \in \{10, 20, 50\}\%$ and $\sigma^m_\eta = \sigma^m_l \in \{1, 2, 5\}\%$ as the BNN test datasets. For each baseline model's test set evaluation, the estimated mean resulted from the power flow solutions, whereas the standard deviations were the training set hourly standard deviations.

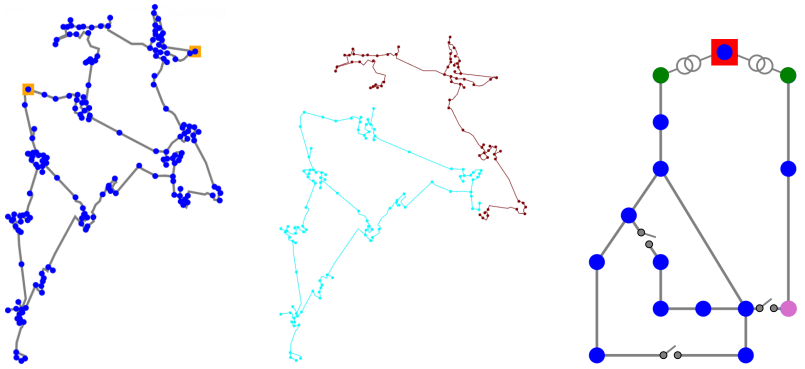
The evaluation metrics for the BNNs and the baselines were (i) the prediction interval coverage probability [114]–[118] for 95% (PICP_{95}):

$$\text{PICP}_{95} = \frac{1}{|\tilde{D}|} \sum_{d=1}^{|\tilde{D}|} \begin{cases} 1 & \text{if } \phi_d \in [\mu_{\chi_d} \pm 1.96 \cdot \sigma_{\chi_d}] \\ 0 & \text{otherwise} \end{cases} \tag{4.24}$$

where $\tilde{D}$ are the validation and test sets, and $\sigma_{\chi_d} = [\sigma_{v_i} \forall i \in \Omega^b, \sigma_{l_j} \forall j \in \Omega^\lambda]^T$. The PICP_{95} metric shows how well the model confidence reflects the actual test data, i.e., PICP_{95} should be as close to 0.95, whereas $\text{PICP}_{95} > 0.95$ shows under-confident model, and $\text{PICP}_{95} < 0.95$ overconfident model. An overconfident model estimates too narrow uncertainty intervals. Additional evaluation metrics are (ii) the root mean squared error (RMSE) between the mean predictions and the actual values [78] (should be as close to zero as possible), (iii) the negative log-likelihood [78], [100], [112] (should have the lowest value).

The case study in Sec.4.4.3 estimates safety-constrained FAs in the OB1 with, $c_v^{max} = 1.05, c_v^{min} = 0.95$, and 5 FSPs. The active and reactive power discretization steps were 0.05MW, 0.1MVAR, respectively. The $\sigma_{v_i} = 0.01 p.u. \forall i \in \Omega^b$, $\sigma_{l_j} = 4\%$ for all the lines and 2% for the measured transformers. The FSP flexibilities were any power setpoint within each FSP's nominal apparent power.

The case study in Sec.4.4.4 estimates safety-constrained FAs in the OB with 2 PCC, with 7 FSPs. The active and reactive power discretization steps were 0.03MW, 0.03MVAR, respectively. The uncertainties were the same as Sec.4.4.3. Due to the initial DS conditions being outside the $0.95 - 1.05 p.u.$ limits, the constraints were $c_v^{min} = 0.9 p.u.$ and $c_v^{max} = 1.1 p.u.$. The FSP flexibilities for the active power scenarios were any setpoints



(a) OB. (b) Ob0 ■ and Ob1 ■. (c) CIGRE MV.

Figure 4.4: Test network lines (—), buses (●), HV-MV transformer stations (■), transformers (⌀), external grid (■), CIGRE observable buses (●), RTI observable buses (●), switches (⌀). (b) shows how Ob separates into Ob0 and Ob1.

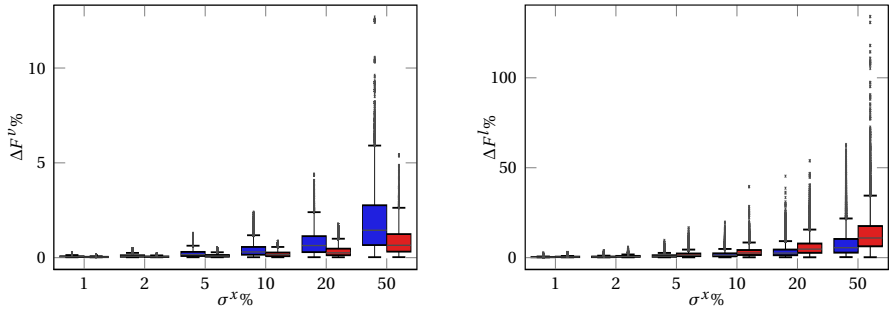
below the initial active power. For reactive powers, the flexibilities were between −150% and 150% from the initial reactive powers for the loads and between −50% and 50% from the initial active power outputs for the generators (as the initial setpoints had 0 reactive power).

4.4.1. SENSITIVITY LOCAL APPROXIMATION

This case study analyses the impact of different pseudo-measurement deviation levels on the sensitivity between DS operating conditions, for the validity of the assumption in Sec.4.2.3. The FA estimation algorithm approximates the sensitivity of network buses $F^v(\cdot)$ and lines $F^l(\cdot)$ to FSP shifts using measured and pseudo-measured values $\{\hat{X}(t), \hat{V}(t), \hat{L}(t), U(t,\tau)\}$ instead of the actual, unobserved values $\{\tilde{X}(t), \tilde{V}(t), \tilde{L}(t), U(t,\tau)\}$, assuming F^v, F^l to remain approximately constant for pseudo-measured operating conditions close to the real operating conditions. Fig.4.5 compares the absolute percentage difference in sensitivities between actual and pseudo-measured conditions $(\Delta F^v\%, \Delta F^l\%)$. The simulations only included $F^v(\tilde{X}(t), \tilde{V}(t), U(t,\tau)) > 0.0005 p.u., F^l(\tilde{X}(t), \tilde{L}(t), U(t,\tau)) > 0.5\%$, to avoid instabilities from low-sensitive components to FSP shifts.

The results had average line loading sensitivites $F^l(\tilde{X}(t), \tilde{L}(t), U(t,\tau))$ between −16.52% and 9.23%, and average voltage sensitivities $F^l(\tilde{X}(t), \tilde{L}(t), U(t,\tau))$ between $-0.017 p.u.$ and $0.019 p.u.$ for the radial network. For the mesh network, the average loading and voltage sensitivities were $[-4.32, 5.13]\%$, and $[-0.009, 0.008] p.u.$, respectively. From the results in Fig.4.5, the sensitivities remain approximately consistent, with an average deviation less than 5% for loading and less than 1% for voltage with pseudo-measurements with 20% or less standard deviation. For 50% deviation, the loading sensitivities deviate by an average of ≈ 11% for loading in the meshed network.

ΔF^l is higher in the meshed network, whereas ΔF^v is higher in the radial network. Addressing the first observation, in radial networks, the load flow paths are relatively limited; in mesh networks, the power injection deviations can redistribute the load flow



(a) (b)

Figure 4.5: Voltage and loading sensitivity deviations for operating conditions with different pseudo-measurement standard deviation for the CIGRE radial (■) and mesh (■) topologies.

paths. Addressing the second observation, the limited paths in radial networks could mean that a different power injection on a bus would impact all upstream or downstream bus voltages.

4.4.2. BNN COMPARISON WITH BASELINES

Fig.4.6 compares the PICP_{95}, RMSE, and NLL between the baselines and the proposed BNN of Sec.4.3 for varying test set noise deviations and networks. The results indicate that the network structure and noise levels impact the BNN less than the baselines. In PICP_{95}, the BNN outperforms the baselines with an average $\text{PICP}_{95}\,v = 0.91$, and $\text{PICP}_{95}\,l = 0.94$ compared to the measurement-based $\text{PICP}_{95}\,v = 0.73$, and $\text{PICP}_{95}\,l = 0.78$, and forecast-based $\text{PICP}_{95}\,v = 0.67$, and $\text{PICP}_{95}\,l = 0.77$. The baseline models consistently underestimate the uncertainty, with values significantly lower than 0.95, except in the lower-noise scenarios for loading variables. In these scenarios, $0.9 \leq \text{PICP}_{95}\,l \leq 0.95$. However, this means that the uncertainty estimation for loading can be accurate if the noise is half of the one considered during estimation. The baseline exclusion of epistemic uncertainty in the estimations contributes to this significant underperformance compared to the BNN. As Fig.4.7 shows, the epistemic uncertainty can be a substantial part of the total uncertainty. Fig.4.6(c)-Fig.4.6(f) show that the BNN drastically improves the voltage and loading RMSE and NLL compared to the baselines. Further, BNN is less impacted by the different noise levels and networks than the baselines. The BNN has, on average, more than 4 times improved voltage RMSE and more than 3 times improved loading RMSE compared to the baselines.

Fig.4.7 shows example MCD predictions on the vertical axis with measurement inputs on the horizontal axis. The uncertainties of Fig.4.7 have a 95% confidence margin (2 standard deviations). To improve the visualization clarity, 90% of the datapoints were removed in Fig.4.7 in densely populated regions. The figures show that epistemic uncertainty depends on the dataset's number of examples. In the regions with more concentrated datapoints (closer to 0 in the horizontal axes), the epistemic uncertainty is smaller than the less-populated regions. This difference highlights that more training examples

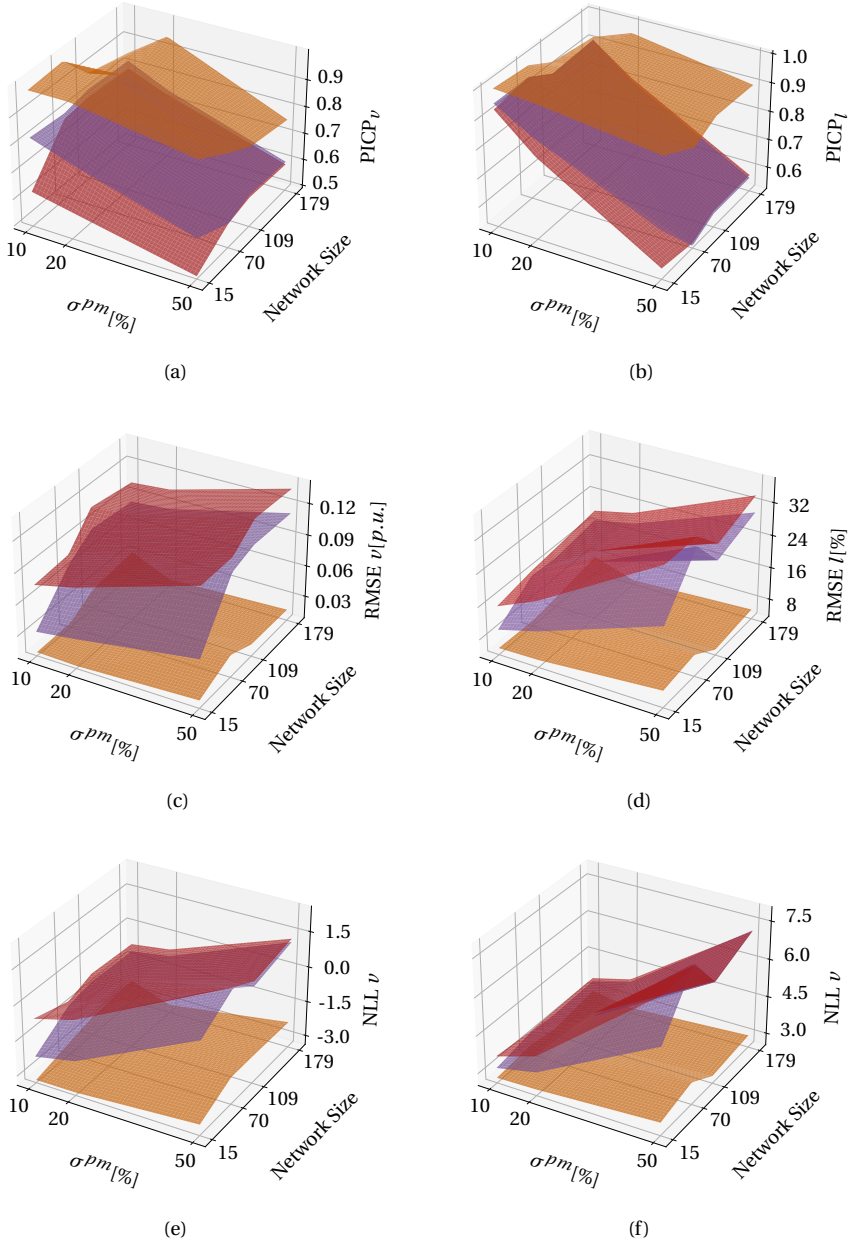
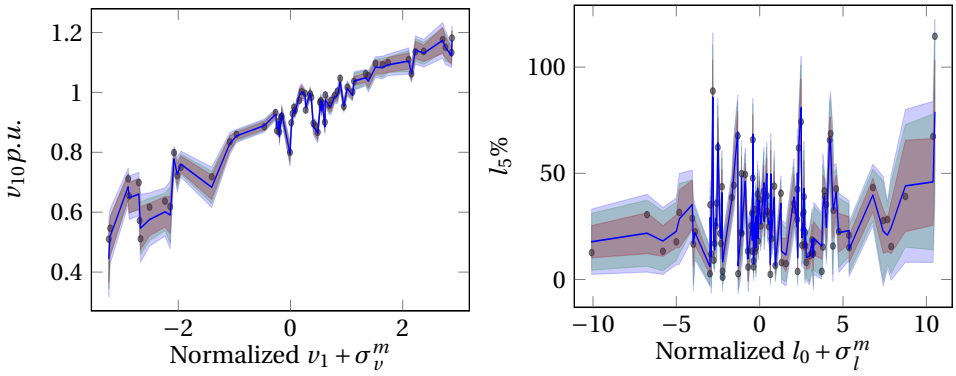


Figure 4.6: Comparison of forecast-based statistical baseline (), measurement-based statistical baseline () and MCD BNN () models under varying pseudo-measurement (and measurement) noise and network size. The network size corresponds to the number of buses.

can reduce the model uncertainty in the parameter space and the epistemic uncertainty. Fig.4.7 also illustrates that the aleatoric uncertainty varies between datapoints.



(a) (b)

Figure 4.7: Voltage and loading predictions by MCD model for CIGRE DS with mean (—), epistemic uncertainty (▬), aleatoric uncertainty (▬), total uncertainty (▬), and true data (•).

Considering the computational burdens, the baselines would not require additional estimations as the FA estimation would solve the power flows to estimate the sensitivities in Fig.3.3. However, the MCD BNN would add a negligible prediction duration, as Appendix C Tab.C.2 shows (between $0.09 - 0.3s$).

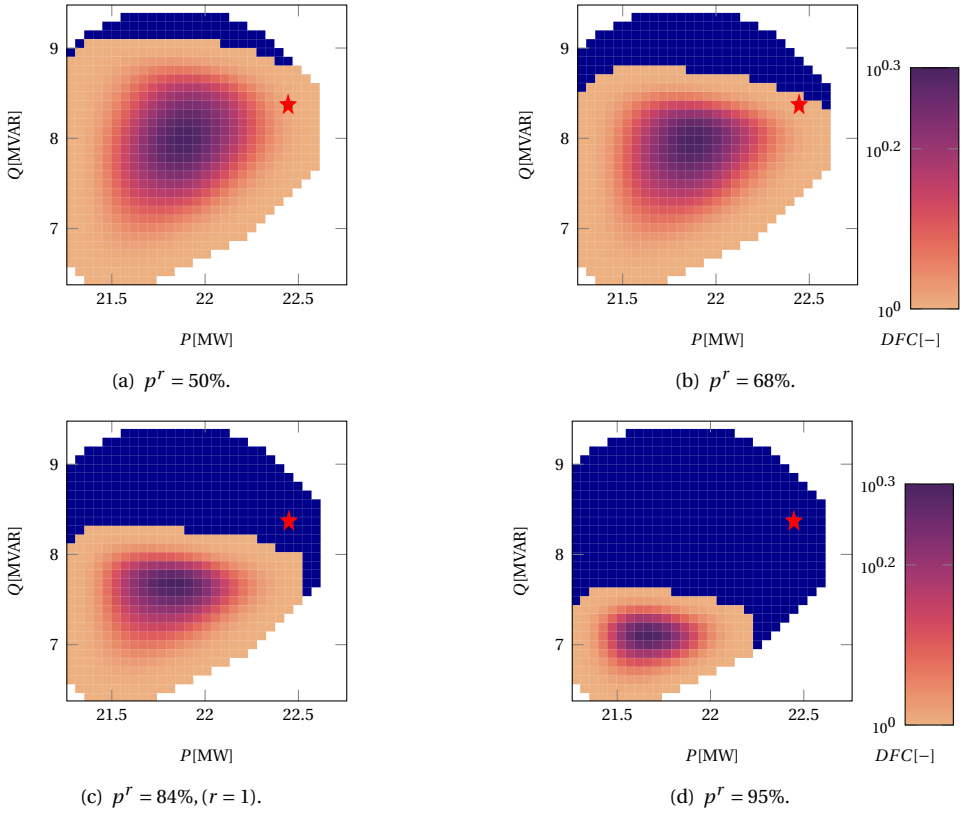
4.4.3. MEASUREMENT UNCERTAINTY-CONTROLLED FA ESTIMATION

This section shows the proposed FA esimation process of Sec.4.2.3. Using the BNN mean prediction and uncertainty, the TSO can select a safety margin on the FAs with the p^r input as in Fig.3.3. Fig.4.8 shows the proposed approach for FA estimation with different safety levels, $p^r = 50\%, 68\%$ ($\pm$ the estimated standard deviation), 84% ($r = 1$), and 95% ($\pm 2\times$ the estimated standard deviation). DFC is the density of feasible combinations for each FA point. The increasing safety margin further restricts the FA feasible space. The uncertainty levels in Fig.4.8(c)-4.8(d) show that even the initial DS operating condition can be out of the selected safety margins.

The FA estimation durations using the Google Colab's A100 GPU required 15.6, 16.2, 17.1, and $18.7s$ for Fig.4.8(a)-4.8(d), respectively. The difference in computational time as the margins increase is due to the increased number of network buses and lines that can reach the network constraints.

4.4.4. TWO-PCC FA ESTIMATION

This case study shows the proposed FA estimation approach for DS with 2-PCCs as explained in Sec.4.2.4. Fig.4.9 shows the resulting FAs for the 2−PCCs when considering only active power in Fig.4.9(a) or only reactive power flexibility in Fig.4.9(b). The diagonal shapes in Fig.4.9 show the correlation between the two PCCs' power exchange, e.g., a flexibility shift to decrease the active power at PCC_1 can also reduce the active power at PCC_2 for the majority of flexibility combinations. However, some combinations (low DFC) can also shift each PCC's power unequally, e.g., decreasing the active power in



(a) $p^r = 50\%$. (b) $p^r = 68\%$.

(c) $p^r = 84\%, (r = 1)$. (d) $p^r = 95\%$.

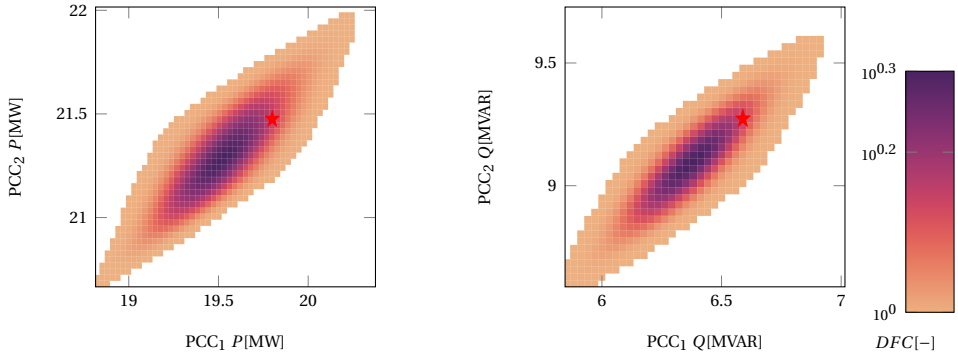
Figure 4.8: Safety-Constrained FAs of OB1 5 FSPs with different safety probabilities p^r. Feasible shifts (■). Not feasible FSP shift combinations (■). Initial operating point (★).

Fig.4.9(a) of PCC_1 by 0.5MW and maintaining the active power of PCC_2 at 21.5MW. Using these FAs, the TSO can select any feasible active or reactive power setpoint for the correlated PCCs. The FA estimation required 11.1s for Fig.4.9(a) and 7.5s for Fig.4.9(b) with the A100 GPU.

4.4.5. DISCUSSION

Through the proposed approach, TSOs can consider real-time uncertainties when accounting for the available flexibility, unlike prior FA approaches. These real-time uncertainties are due to limited measurements, input noise, and model parameters. Using BNNs, the estimated uncertainties can generalize better to different noise levels and networks compared to the statistical baselines extended from [20], [30], [102], [103].

The results of Sec.4.4.2 show the proposed BNN structure improves the uncertainty estimation in all metrics compared to the statistical baselines. The baselines underestimate the uncertainty in the majority of tests ($PICP_{95} < 0.95$). The measurement and



(a) Active Powers.

(b) Reactive Powers.

Figure 4.9: Safety-Constrained FAs of OB with 7 FSPs for active power flexibility support in (a) and reactive power flexibility support in (b). Feasible shifts ( ). Initial operating point (★).

pseudo-measurement noise increase impacts the baselines significantly more than the MCD BNN in all metrics. In terms of computational burden, MCD requires low training times (0.8–1.2 hours) and prediction duration 0.1–0.3s). The computational burden added to [110] for FA estimation is mainly the BNN prediction duration, with FA estimations requiring between 7.5–18.7s.

The results of Appendix C indicate BNN approaches can perform well for the power-system task, maintaining a relatively consistent PICP_{95}, RMSE, and NLL when trained in different networks. Between VI, MCD, and DE BNNs, the MCD has, on average, the best l PICP_{95}, RMSE (for v and l), and NLL (for v and l).

A limitation of the proposed approach is the underconfidence in test sets with lower noise than the training sets ($\text{PICP}_{95} > 0.95$). To mitigate this limitation in case of low pseudo-measurement and measurement noise, DSOs can utilize real-life data to train another BNN MCD in ≈ 1 hour.

4.5. CONCLUSIONS

System operators can use the proposed approach to select safety levels in real-time FA estimations, considering the limited observability in DS. Aiming for real-time operation, the proposed approach is fast and considers the uncertainty from limited real-time DS measurements. In addition, with the proposed approach, TSOs can select active and reactive power flexibility aggregation in DS with 2-PCCs.

The proposed BNN structure improves the accuracy of the estimated uncertainty and RMSE compared to the baselines. The proposed BNN improves the RMSE compared to baselines by an average $\approx 4\times$ for voltage and $\approx 3\times$ for loading. The average MCD BNN test uncertainty using PICP_{95} shows slight overconfidence with 0.91 for voltage and 0.94 for loading, unlike the baselines that show significant overconfidence.

Future work includes evaluating the BNN performance in real data from distribution networks and studying the impact of alternative measurement locations in DS.

5

Machine Learning-based Method to Support TSO-DSO Adaptive Coordination with Active Power Management for Instability Prevention

An ounce of prevention is worth a pound of cure

Benjamin Franklin

Coordination between power system operators can improve the power system stability and effectively deploy resources in distribution systems (DS). The research work of this chapter provides a coordination method to mitigate the impact of dynamic events on transmission systems (TS). The proposed method uses a machine learning (ML)-based model to estimate the collective dynamic response of DS under varying TS dynamic properties, DS operating conditions, and share of inverter based resources (IBRs). In addition, the ML-based model enables TS operators (TSOs) to provide feedback to DS operators (DSOs) for controlling the IBRs' active power output to prevent post-fault instabilities. The proposed TSO-DSO coordination method includes a risk-based active power setpoint optimizer for instability prevention. The proposed method uses existing measurement and IBR control platforms available in DS and estimates the post-fault DS dynamic response considering IBR active power control actions. Case studies on synthetic models of TS and DS covering the Zeeland province in the Netherlands illustrate the application of the proposed coordination and the instability risk mitigation when optimizing IBR setpoints.

5.1. Introduction

Inverter-based resources (IBRs) in distribution systems (DSs) affect the overall power system dynamic stability. DS IBRs operating in grid-forming (GFM) mode can support the dynamic stability of transmission systems (TSs) that suffer from low inertia. Transmission system operators (TSOs) need to anticipate and evaluate the impact of controlling DS IBRs on system stability as IBRs in DSs replace TS resources. However, including active components, distributed generation, and IBRs in DSs makes modeling, maintaining, and co-simulating TSs and DSs challenging [120].

Coordination between TSOs and distribution system operators (DSOs) is a topic of increasing interest [10], [121], [122]. As distributed energy resources (DER) flexibility becomes necessary, TSOs and DSOs must coordinate to avoid causing issues to one another [122]. Existing TSO-DSO coordination methods differ in the central entity and data exchange requirements [10]. DSO-centric coordination allows low data exchange between TSOs and DSOs using flexibility areas [1], [18], [36]. However, flexibility areas evaluate and estimate the available flexibility through steady-state simulations. Evaluating the flexibility for dynamic stability is underrepresented in TSO-DSO coordination methods.

TSOs must ensure their system's security, withstanding imminent contingencies without service interruptions [123], [124]. The typical industry practice is deterministic, where all probable contingencies are treated with equal risk [124]. However, risk-based security assessment becomes more applicable as the contingency probability, risks, and consequences differ [123], [124]. Probabilistic approaches can consider the contingency likelihood, probability, and consequences of instability [123]. The proposed coordination helps TSOs assess the risks and costs of probable contingencies compared to flexibility costs.

Different share between grid-following (GFL) and GFM inverters impacts the dynamic response of DSs [125]. GFM inverters improve the dynamic performance of power systems with enhanced response to load [126] and fault [125] events. However, GFM inverters have stability issues after faults due to current saturation [126]–[128]. Thus, controlling GFM IBRs to increase the headroom between the IBR output and the current limitation improves the system stability under disturbances. Different TS inertia alters the response of the DS [129], [130], and the system inertia is highly variable [130], [131].

Methods for DS aggregated dynamic response representation (ADRR) aim to support dynamic studies for TSOs and DSOs by alleviating the DS modeling and simulating complexity. Existing ADRR approaches do not support controlling DS GFM IBR outputs and typically consider fixed TS dynamic properties throughout the ADRR development. Considering TSs with fixed, high inertia in ADRR modeling does not accurately represent the DS response for TSOs to assess the impact of controlling GFM IBR setpoints. Thus, DS ADRR models should account for the variable TS dynamic properties and support controlling the DS GFM IBR setpoints.

This research proposes a TSO-DSO coordination method that enables TSOs to evaluate and request flexibility from GFM IBRs for dynamic support. Flexibility corresponds to

Parts of this chapter are in: **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"Machine Learning-based Method to Support TSO-DSO Adaptive Coordination with Active Power Management for Instability Prevention"*, International Journal of Electrical Power & Energy Systems, 2025 [119].

the DS resources whose operation can be altered by TSOs or DSOs to support system operation [132]. A DS ADRR is a key component in enabling this coordination, accounting for the variable TS dynamic properties, types of IBRs, and IBR setpoints. The proposed TSO-DSO coordination requires measurements currently available to system operators, identifies the GFM IBR setpoints minimizing instability risks, and includes a classifier identifying extreme instabilities.

DS ADRR approaches mainly use system identification approaches that rely on measured or simulated data [46], [120]. The scarcity of data with large disturbances limits the potential for measurement-based ADRR models. System identification-based ADRR approaches can diversify to black box and grey box. Black box approaches use data to fit a model of an unknown, typically machine learning (ML)-based structure [52], [133]. Grey box approaches assume a DS equivalent structure, e.g., a load, a synchronous, and a static generator, and use data to fit the structure's parameters [46], [54]. Grey box models are easier to integrate with software for dynamic simulations. However, the grey box modeling selection of a specific system structure and parameters can impair the model representation for variable operating conditions. To allow representing various operating conditions, [33], [134] provide different ADRR model parameters for different operating condition clusters. With increased diversity in IBRs and DERs in DSs, grey box approaches require more complex structures with several components [135], [136], whereas ADRR models should ease the DS modeling and simulation complexity. Applying a grey box approach can be challenging for a ring system topology [136], and including external variables in grey box model structures can be challenging in mapping the external variables and structure components. Black box methods do not have these restrictions in operating conditions, IBR diversity, external variables, or system topology. As ML algorithms become more prominent, integrating ML models into dynamic simulation software becomes anticipated. ML-based ADRR models explored long-short-term memory (LSTM) recurrent neural networks (RNNs) [52], Gaussian process models [133], and artificial neural networks [53]. RNNs are developed for sequential data but cannot effectively include static features. [133] included static features with repeated power outputs and voltage inputs. The proposed black box approach utilizes RNNs but also considers static features for the initial operating conditions, IBR share, and TS inertia.

DS models used in ADRR approaches typically assume knowledge of the IBR types, as GFL [120], [134]. However, GFM inverters become increasingly used [137]. To account for differences in DS generation types, [134] model considers the mixture of synchronous and asynchronous generation but assumes asynchronous DER. [34] modeled GFM and GFL IBR types and used the DS IBRs and synchronous generator inertia to provide an ADRR model with a more accurate response. [34] assumed the DSOs know detailed characteristics of the IBRs, e.g., the inertia, droop, and damping constants, and a high TS inertia. [136] modeled GFM IBRs but also assumed known IBR characteristics and developed a complex ADRR model structure with one branch for each GFM inverter counteracting the ADRR appeal for DS model simplification. This research represents IBRs with GFL and GFM types and estimates these types using measurements without assuming detailed IBR parameter knowledge.

Preventive dynamic security assessment studies require dynamic models of the underlying systems to identify remedial actions [138]. These security assessment studies con-

sider equally credible contingencies [139], [140] or include contingency probabilities for instability risks [123]. These studies typically focus on TS operating conditions to minimize instability risks. The proposed method focuses on controlling DS IBRs and the impact of DS IBR dynamics that preventive dynamic security assessment studies mainly exclude.
This chapter proposes a TSO-DSO coordination method where TSOs can use the ADRR model to select and evaluate flexibility setpoints for the IBRs. Therefore, the main contributions are:

1. Developing a TSO-DSO coordination method for dynamic stability based on the estimation of the aggregated dynamic response of DS using measurements available to system operators.

2. Extension of the method to also include an algorithm for risk-based evaluation and selection of TSO actions to prevent instability.

3. Proposing an approach representing the aggregated DS dynamic response accounting for variable dynamic system properties, the aggregated power output, and the type of primary control of IBRs, combining sequential and non-sequential features.

Case studies are performed on the synthetic system model of the Zeeland region in the Netherlands. The case studies showcase the classification of IBR types, active power setpoint optimizer, and ADRR model performance with regression and high instability classification. The following subsections are Sec.5.2 TSO-DSO coordination; the proposed coordination approach and key components, Sec.5.3 DS aggregated dynamic response representation for variable TS dynamic properties and IBR share; the proposed ADRR model structure, Sec.5.4 case studies; the case studies, and Sec.5.5 conclusion.

5.2. TSO-DSO COORDINATION

As the share of DER increases, IBR flexibility can help TSOs achieve operational stability. Assessing IBR flexibility for dynamic events is challenging for TSOs as:

1. TSOs would need real-time information on the available flexibility in DSs from each IBR type,

2. TSOs would need models to consider the post-disturbance dynamic response of interconnected DSs,

3. TSOs would need functionality to select DS IBR setpoints to improve the TS dynamic stability.

To mitigate these challenges, this chapter introduces the coordination illustrated in Fig.5.1. The DSO applies the IBR classifier algorithm to inform the ADRR model about the real-time DS operating condition. The TSO uses the operating condition to evaluate the stability of the post-fault system. If the ADRR detects instability, the GFM IBR active power setpoint optimizer identifies the risk-minimizing setpoint for the GFM IBRs and informs the DSO. The DSO validates that the requests respect DS network constraints and sends

the validated requests to the IBRs through RTI. If TSO requests are invalid for the constraints of the DS network, the TSO is informed to explore an alternative risk mitigation without altering IBR setpoints.

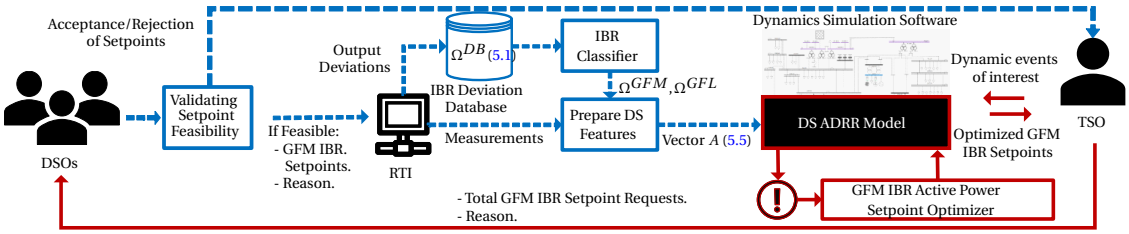


Figure 5.1: Overview of the proposed TSO-DSO coordination with DSO-operated actions (⇢) and TSO-operated actions (→).

5.2.1. ESTIMATING THE REAL-TIME IBR OUTPUTS IN DSS

DSOs have limited measurements in DSs [29]. This section describes how the proposed coordination considers the available measurements through the existing Dutch real-time interface (RTI) project to estimate the real-time IBR outputs in DSs. In 2020, Dutch DSOs and the TSO developed the RTI project to improve real-time communication between large DERs and DSOs, TSOs [141]. Through RTI, DERs with capacity 1–50$[MW]$, connected to the medium voltage, are required to measure and exchange information with the DSOs about the active power, reactive power, phase voltages, and phase currents. These limits correspond to type B generators in the European network code on requirements for generators [142]. Example DER in MVs for RTI include PV systems and wind turbines [141].

RTI measurements can have intervals of less than 60$[s]$. Using RTI, the DSOs and TSO can request new power setpoints from DER to ensure their system's stability. The DSOs and TSO must include the reason for the requests.

The proposed coordination creates a database of IBR output deviations provided by RTI for DSOs to classify each IBR as GFM or GFL. These deviations are load change disturbances that occur continually [124], i.e., a sudden load consumption change. The database includes the wind speed deviation as uncertainty caused by variable weather conditions. These wind deviation values are attainable from real-time regional weather radars. Set Ω^{DB} characterizes the database as:

$$\Omega^{DB} = \{(\Delta|v|_{jl}, \Delta i_{jl}, \Delta P_{jl}, \Delta Q_{jl}, \sigma_l) | j \in \Omega^{DER}, l \in \Omega^{\Phi}\}, \tag{5.1}$$

where Ω^{DER} is the set of DER providing measurements through RTI, and Ω^{Φ} is the set of load disturbances. $\Delta|v|_{jl}, \Delta i_{jl}, \Delta P_{jl}, \Delta Q_{jl}$ are the voltage, current, active, and reactive power changes of the j DER terminals after the l load event. σ_l is the wind speed standard deviation measurement during each l event. These σ_l measurements should be based on regions between wind turbines, as wind measurements for each individual turbine might be unavailable to DSOs.

While the IBR share impacts the DS response to dynamic events [125], DSOs are unaware of specific DS IBR types of operation. Thus, in the proposed coordination, the DSO ap-

plies the IBR classifier in Alg.1 using Ω^{DB} to classify each IBR as GFM or GFL, resulting in the sets Ω^{GFM} and Ω^{GFL} respectively. The proposed approach assumes that DSOs know which DERs are synchronous machines, Ω^{SM}, e.g., gas and steam turbines. This assumption differs from alternative approaches that assume knowing the IBR types and DER parameters [34], [136].

Algorithm 1 Classifying IBR types

Input: $\Omega^{DB}, \Omega^{DER}, \Omega^{SM}$
Output: $\Omega^{GFM}, \Omega^{GFL}$
$\Omega^{GFM} \leftarrow \emptyset, \Omega^{GFL} \leftarrow \emptyset$
for $j \in \Omega^{DER} \setminus \Omega^{SM}$ **do**
  $\psi \leftarrow 0$
  for $(\Delta P_{jl}, \sigma_l) \in \Omega^{DB}$ **do** ▷ from (5.1)
    $\psi \leftarrow \psi + \Psi(\Delta P_{jl}, \sigma_l)$ ▷ using (5.2)
  end for
  if $\psi > 0$ **then**
    $\Omega^{GFM} \leftarrow \Omega^{GFM} \cup \{j\}$
  else
    $\Omega^{GFL} \leftarrow \Omega^{GFL} \cup \{j\}$
  end if
end for
return $\Omega^{GFM}, \Omega^{GFL}$

Following DS disturbances such as load changes, GFM IBRs typically reduce or increase the output power to mitigate the disturbance and maintain stable voltage and frequency [143]. In contrast, GFL IBRs typically continue to inject power following the system voltage and frequency [143]. GFL IBRs, including the ones used in the case studies, can also offer frequency support with droop, with the GFM IBR response typically being stronger [144]. The active power threshold diversifying GFM from GFL responses is $\Delta\widehat{P}$[%]. The classifier takes a weighted assignment $\Psi \in [-1, 1]$ on the IBR type after each event l in line 5 of Alg.1, as:

$$\Psi(\Delta P_{jl}, \sigma_l) = \begin{cases} tanh(\frac{|\Delta P_{jl}| - \Delta\widehat{P}}{\sigma_l}), & \text{if } \sigma_l \neq 0, \\ sign(|\Delta P_{jl}| - \Delta\widehat{P}), & \text{otherwise,} \end{cases} \tag{5.2}$$

where $\Psi > 0$ for GFM type, and $\Psi < 0$ for GFL type. The Ψ magnitude is the weight for the class assignment confidence. The classifier considers wind speed variability during each event as a source of uncertainty. Wind deviations can cause changes in IBR outputs unrelated to the load events. Thus the recorded ΔP_{jl} are:

$$\Delta P_{jl} = \Delta\bar{P}_{jl} + \epsilon(\sigma_l)[\%], \tag{5.3}$$

where $\Delta\bar{P}_{jl}$ is the change caused by the actual inverter type of IBR j responding to event l. The $\epsilon(\sigma_l)$ is the uncertainty caused by the wind variability. Inverters and controllers of wind turbines typically include damping mechanisms to reduce output deviations due

to weather conditions [145]. These mechanisms reduce the impact of $\epsilon(\sigma_l)$ on ΔP_{jl}. Nevertheless, an event l with low σ_l compares to a low $\epsilon(\sigma_l)$.
Since Ω^{DB} stores multiple load events, Alg.1 gets $\phi = |\Omega^{\Phi}|$ samples to classify each IBR as GFM or GFL. The class assignment from each sample has a weight of confidence from (5.2). Alg.1 sums the weighted class assignments for each IBR. If the accumulated Ψ from all events is non-negative, then the IBR is classified as GFM in lines 7 – 8 of Alg.1. Otherwise, the IBR is classified as GFL in lines 9 – 10 of Alg.1. Let a classification metric with a probability of correct classification $\hat{\rho}_l^{\gamma} \in [0,1]$ for a single load event l. Given ϕ independent load events, the probability of the correct class being predicted for an IBR for the majority of events is:

$$\hat{\rho}_{\phi}^{\gamma} = \sum_{n=\frac{\phi+1}{2}}^{\phi} \binom{\phi}{n} (\hat{\rho}_l^{\gamma})^n (1-\hat{\rho}_l^{\gamma})^{\phi-n} \in [0,1], \tag{5.4}$$

which is the cumulative distribution function of the binomial distribution.
The DSO uses the $\Omega^{GFM}, \Omega^{GFL}$ outputs from the IBR classifier and the constant flow of real-time measurements from RTI to output vector A of parameters that could impact the DS response to dynamic events:

$$\begin{aligned} A = [& \sum_{\hat{j}\in\Omega^{GFM}} P_{\hat{j}}, \sum_{\hat{j}\in\Omega^{GFM}} Q_{\hat{j}}, \sum_{\tilde{j}\in\Omega^{GFL}} P_{\tilde{j}}, \sum_{\tilde{j}\in\Omega^{GFL}} Q_{\tilde{j}}, \\ & \sum_{\tilde{j}\in\Omega^{SM}} P_{\tilde{j}}, \sum_{\tilde{j}\in\Omega^{SM}} Q_{\tilde{j}}, P_{\hat{j}} \forall \hat{j} \in \Omega^{GFM}, Q_{\hat{j}} \forall \hat{j} \in \Omega^{GFM}, \\ & S_{\hat{j}} \forall \hat{j} \in \Omega^{GFM}]^T. \end{aligned} \tag{5.5}$$

Where P is active power, Q is reactive power, and S is apparent power.

5.2.2. TSOs SELECTING IBR SETPOINTS

This section analyzes the GFM IBR setpoint optimizer of Fig.5.1, which enables TSOs to select IBR setpoints to minimize instability risks for the post-fault response of the interconnected DS and TS. ADRR models can address the second challenge of Sec.5.2, where a DSO provides a DS ADRR model that the TSO can deploy to perform dynamic simulations. The proposed ADRR model includes the IBR setpoints as inputs for the TSO to evaluate the impact of flexibility actions on the DS response and to address the third challenge of Sec. 5.2.
Unlike small load disturbances, large disturbances in TSs, such as line faults, are less frequent but can cause instability and high costs for TSOs [124]. The TSO can use the proposed coordination process to avoid instabilities caused by large disturbances. In Fig.5.1, the TSO informs the DSO about the setpoint request for all IBRs and the dynamic event (reason for request). The DSO validates DS constraint-feasibility for the TSO requests using distribution system state estimation (DSSE). DSSE approaches consider measured nodes and pseudo-measured (or forecasted) nodal power injections to estimate the DS state [28]. In the proposed approach, the DSO can apply DSSE using RTI measurements for the IBRs excluded in the TSO requests, the new setpoints for GFM IBRs in the TSO requests, and pseudo-measurements for the remaining DS nodes to estimate the DS state. If the requested state respects the DS constraints, the DSO forwards

the requests to the GFM IBRs through RTI. Alternatively, the DSO informs the TSO that the setpoints are invalid, so the TSO can explore alternative risk-mitigation sources from neighboring DSOs or TS-connected resources. RTI requires the DSOs to provide the IBR with this reason for the request.

The current saturation issue of droop-controlled GFM inverters causes instability if the active power output of the inverter exceeds the limit, as analyzed in [127]. In the proposed approach, TSOs can simulate disturbances in TS with ADRR representing the DS, considering the potential increase in GFM IBR active power after probable disturbances. If these simulations indicate that an increase in the active power output of GFM IBR would cause instability due to current saturation, reducing the GFM IBR setpoints could result in greater headroom, avoiding reaching maximum current. The proposed approach compares the risks associated with the probability of disturbance and the costs of instability to avoid unnecessary setpoint reductions.

A large disturbance event e has a probability of occurrence $\hat{\rho}^o(e) \in [0,1]$. If event e causes instability, the potential costs for the TSO can be $C^\iota(e)$ [€], caused by loss of load in the area or outage costs [124]. Alternatively, the costs of adjusting the GFM IBR setpoints to prevent instability can be $C^\eta(e)$[€/%]. The active power setpoint for GFM IBR is P^s[%]. The $\hat{\rho}^\iota(e,\zeta,A,P^s) \in [0,1]$ is the probability of instability for TS operating condition ζ, and DS attributes of vector A, after e, if P^s is applied. $\hat{\rho}^\iota(e,\zeta,A,P^s)$ is independent of $\hat{\rho}^o(e)$. Thus, the risk for the TSO, $R(e,\zeta,A,P^s)$, for e is:

$$R(e,\zeta,A,P^s) = \hat{\rho}^o(e)\cdot\bar{R}(e,\zeta,A,P^s) + (1-\hat{\rho}^o(e))\cdot C^\eta(e)\cdot(100-P^s)[€], \tag{5.6}$$

$$\bar{R}(e,\zeta,A,P^s) = C^\iota(e)\cdot\hat{\rho}^\iota(e,\zeta,A,P^s) + C^\eta(e)\cdot(100-P^s)[€]. \tag{5.7}$$

Where $\bar{R}(e,\zeta,A,P^s)$ is the risk for the TSO if e occurs. Thus combining (5.6) and (5.7) leads to:

$$R(e,\zeta,A,P^s) = \hat{\rho}^o(e)\cdot C^\iota(e)\cdot\hat{\rho}^\iota(e,\zeta,A,P^s) + C^\eta(e)\cdot(100-P^s)[€]. \tag{5.8}$$

Let $\overline{P^s}$[%], the maximum setpoint addressing the potential instability event. TSOs can either neglect the potential impact of e or apply $\overline{P^s}$ by assessing the risks of these actions as:

$$P^{s*} = \text{argmin}_{P^s\in\{\overline{P^s},100\}}\Big(R(e,\zeta,A,P^s)\Big)[\%], \tag{5.9}$$

where P^{s*}[%] is the minimum-risk setpoint.

The GFM IBR active power setpoint optimizer of Fig.5.1 and Alg.2 identifies the minimum-risk setpoint P^{s*}, of resolution δP[%], for GFM IBR to avoid the potential instability from the large disturbance e. The optimizer first identifies the maximum feasible setpoint $\overline{P^s}$ and then applies (5.9) to obtain P^{s*}. To estimate $\overline{P^s}$, Alg.2 performs a tree search for a list of setpoint levels Γ^s. The inputs $\bar{\Lambda}$ are the TSO stability thresholds, i.e., limits for frequency, voltage, rate of change of frequency (RoCoF), and active and reactive power deviations. Λ are the simulated values for the above thresholds. P^s_{min}[%] is the minimum

setpoint evaluated by the algorithm. $\tilde{A}$ is the value for A if P^s is applied. The function $F^s(A,P^s)$ applies the selected setpoints P^s to the GFM IBR attributes of A, representing controlling the GFM IBR setpoints on line 6 of Alg.2. $F^{sim}(e,\zeta,A)$ performs the dynamic simulation considering e and the DS and TS operating condition on line 7 of Alg.2. If no setpoint can solve the instability, the $\overline{P^s}=100$ from line 3 of Alg.2 shows that none of the changes provided a solution, thus (5.9) returns no action on line 25 of Alg.2. The setpoint optimizer does not intend to control the voltage or frequency post-fault response to specific follow trajectories, but rather reduce the risks for post-fault instability, considering the ADRR and TS simulated responses.

Algorithm 2 GFM IBR Active Power Setpoint Optimizer

Input: $e,\zeta,A,\delta P,\bar{\Lambda},P^s_{min},\hat{\rho}^o(e)$
Output: P^{s*}
1: $\Gamma^s \leftarrow [P^s_{min}, P^s_{min}+\delta P, P^s_{min}+2\delta P,..,100]$
2: $\overline{P^s_{max}} \leftarrow 100$
3: $\overline{P^s} \leftarrow 100$
4: **while** $\Gamma^s != [\ ]$ **do**
5: $P^s \leftarrow \Gamma^s[\lfloor \frac{\text{length}(\Gamma^s)}{2} \rfloor]$ ▷ Get the median value
6: $\tilde{A} \leftarrow F^s(A,P^s)$
7: $\Lambda \leftarrow F^{sim}(e,\zeta,\tilde{A})$
8: **if** $\Lambda <= \bar{\Lambda}$ **then**
9: **if** $P^s+\delta P < P^s_{max}$ **then**
10: $\Gamma^s \leftarrow [P^s+\delta P,..,P^s_{max}]$
11: **else**
12: $\Gamma^s \leftarrow [\ ]$
13: **end if**
14: $\overline{P^s} \leftarrow P^s$
15: $P^s_{min} \leftarrow P^s$
16: **else**
17: **if** $P^s_{min} < P^s-\delta P$ **then**
18: $\Gamma^s \leftarrow [P^s_{min}, P^s_{min}+\delta P,..,P^s-\delta P]$
19: **else**
20: $\Gamma^s \leftarrow [\ ]$
21: **end if**
22: $P^s_{max} \leftarrow P^s$
23: **end if**
24: **end while**
25: $P^{s*} \leftarrow$ (5.9) with $e,\zeta,A,\overline{P^s}$
26: **return** P^{s*}

In the proposed coordination, TSOs can identify potential communication failures by detecting disruptions in the information flow between RTI measurements and vector A input for the DS ADRR, as in Fig.5.1. In such cases, TSOs should use previously measured vector A instances until communication is restored. If simulations indicate instability, TSOs should only consider instability preventive measures outside the DS controllable

devices until RTI communications are re-established. Given the preventive nature of the proposed coordination, the GFM IBR response time to TSO setpoint requests can be on the scale of seconds, e.g., 30s.

5.3. AGGREGATED DYNAMIC RESPONSE REPRESENTATION

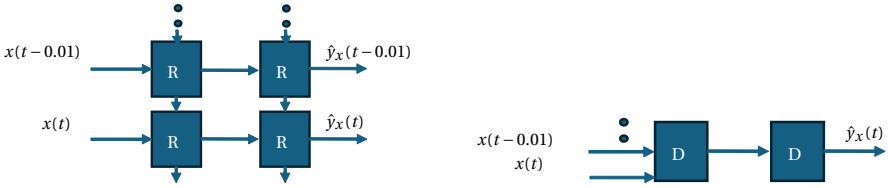
This section describes the proposed DS ADRR model, which comprises an ADRR regressor and an ADRR classifier. The DS ADRR model, as shown in Fig.5.1, enables the TSO to simulate the post-fault response of the DS considering the setpoints for DS IBRs. The proposed coordination of Fig.5.1 uses the IBR Classifier and vector A from (5.5) to inform the DS ADRR model on the DS condition. The setpoint optimizer of Sec.5.2.2 allows the TSO to optimize and evaluate GFM IBR setpoints using the DS ADRR model.

The proposed ADRR model includes a regressor representing the DS response to a large TS disturbance in a simulation and a classifier to detect high instabilities and terminate the simulation process. Power systems simulation software such as PowerFactory rely on iterative numerical methods. Highly instable conditions can cause no results or non-convergence issues in the numerical methods. These issues can further delay the simulations, lead to possibly erroneous results, or terminate the simulations early. Extreme spikes caused by these instable conditions can dominate optimization losses and reduce the overall ADRR model accuracy. The proposed ADRR structure identifies these highly instable conditions to inform TSOs about the high instability. The proposed ADRR model considers a regressor that outputs the DS active and reactive powers at the DS-TS point of common coupling (PCC) and a classifier that identifies high instabilities and terminates the simulation.

5.3.1. ADRR REGRESSION MODEL

Existing ADRR approaches typically do not require non-sequential features or do not use RNNs that focus on sequential features. Sequential features $x(t)$ are essential to capture the temporal dependencies of dynamic simulations. The non-sequential features z, which include vector A from (5.5), provide information on the initial operating condition and sensitivity of the DS to the IBR outputs. Sequential features depend on the dynamic simulation time-step t, e.g., total DS active power output $P(t)$. The simulation time-step does not impact non-sequential features, e.g., initial DS active power output. As shown in Fig.5.2, RNN layers retain the outputs from prior inputs to effectively capture the $x(t)$ temporal dependencies. Emerging RNN architectures such as LSTMs and gated recurrent units have gated mechanisms to keep long- and short-term dependencies in the sequential variables. To capture temporal dependencies, a feedforward neural network (FNN) would require input vectorized sequences of variables that can be less effective than RNNs.

The proposed model of Fig.5.3 uses the RNN specialization for sequential features $x(t)$ and applies a linear dense layer on non-sequential features z. The sequential input feature matrix $x(t)$ includes the PCCs' voltage magnitudes $|v|(t)$ and angles $\theta(t)$, the fre-



(a) RNN. (b) FNN.

Figure 5.2: NN dealing with sequential features. In (a), RNN layers (R) retain prior iteration outputs. In (b), dense layers (D) must include the sequence of inputs to identify temporal dependencies.

quency $f(t)$ and their previous values in a window μ as:

$$x(t) = [\hat{x}(t), \hat{x}(t-\tau), .., \hat{x}(t-\mu\cdot\tau)], \tag{5.10}$$

$$\hat{x}(t) = [|v|_{PCC_1}, |v|_{PCC_2}, \theta_{PCC_1}, \theta_{PCC_2}, f]^T(t), \tag{5.11}$$

where $\tau[s]$ is the simulation time-step, and PCC_1, PCC_2 are the 2 PCCs in ring DS topologies. The non-sequential inputs are the PCCs' initial power outputs P^0_{PCC}, Q^0_{PCC}, the TS inertia characteristic constant H, and the features vector A from (5.5):

$$z = [P^0_{PCC_1}, P^0_{PCC_2}, Q^0_{PCC_1}, Q^0_{PCC_2}, H, A^T]^T. \tag{5.12}$$

As Fig.5.3 shows, the proposed model performs feature fusion, concatenating the dense and RNN layer outputs. A FNN uses the combined outputs to predict the DS response. The vector of observed active and reactive power at each PCC at each time step t is:

$$y(t) = [P_{PCC_1}, P_{PCC_2}, Q_{PCC_1}, Q_{PCC_2}]^T(t). \tag{5.13}$$

For a single-PCC DS, y would exclude the features and outputs for PCC_2. For DSs with more PCCs, y would expand the features and outputs for each additional PCC, e.g., $|v|_{PCC3}$.
The function describing the proposed model of Fig.5.3 is:

$$\hat{y}(t) = F^{FN}_{w_{FN}}\left(F^{FF}(\hat{y}_x(t), \hat{y}_z)\right). \tag{5.14}$$

Where $\hat{y}_x(t)$ is the output of the RNN layers and $\hat{y}_z$ is the output of the dense layer. F^{FF} is the feature fusion function, concatenating the RNN and FNN model outputs. $F^{FN}_{w_{FN}}$ is the FNN model that takes the fused features as inputs, and outputs the $\hat{y}(t) \in \mathrm{R}^4$; to approximate the observed $y(t)$. The w_{FN} are the learnable weights of the FNN model.
The proposed RNN, as in [52], deploys Ξ LSTM RNNs in parallel with their outputs being concatenated in vector $\hat{y}_x(t)$ as:

$$\hat{y}_x(t) = [\hat{y}_{RNN_0}, .., \hat{y}_{RNN_\Xi}]^T(t), \tag{5.15}$$

$$\hat{y}_{RNN_\xi}(t) = F^{RNN}_{w_\xi}(x(t), \hat{y}_{RNN_\xi}(t-1)) \forall \xi \in [0, \Xi], \tag{5.16}$$

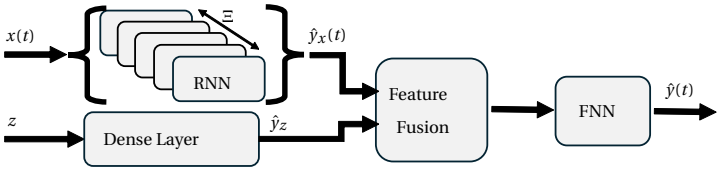


Figure 5.3: Proposed ADRR Regression Model.

where $F^{RNN}_{w_\xi}$ is the ξ^{th} RNN model and w_ξ the learnable weights. The dense layer output is:

$$\hat{y}_z = F^{DL}_{w_z}(z), \tag{5.17}$$

where $F^{DL}_{w_z}$ is the dense layer and w_z the learnable weights.
The regression objective is to find parameters $w = \{w_{FN}, w_\xi \forall \xi \in \Xi, w_z\}$ to minimize the error between the model outputs and observed DS outputs. Hence, the loss function L is the root mean squared error (RMSE) between y and $\hat{y}$ as:

$$L = \frac{1}{|D| \cdot \kappa} \sqrt{\sum_{d \in D} \sum_{k=0}^{k=\kappa} (y(k \cdot \tau) - \hat{y}(k \cdot \tau))^2}, \tag{5.18}$$

where $|D|$ is the length of dataset D with samples $d = (x(0), .., x(\kappa \cdot \tau), y(0), .., y(\kappa \cdot \tau), z) \in D$. κ is the number of time steps for each simulation.

5.3.2. Classification of Instabilities

High instabilities can be present in simulations with low TS inertia. ADRR models could inform TSOs and help them avoid these high instabilities. Unlike related literature, the proposed ADRR model also includes a classifier for high instabilities during the simulation to inform the TSO. Excluding the high instability classifier means including simulations with extreme values and fluctuations that could dominate the regression model's loss and learning process.
Classifying extreme instabilities requires analyzing sequential data to detect high fluctuations. However, assigning each sequence value as a unique feature in an FNN can be inefficient. The most prominent methods for time-series classification first process each signal to get low dimensional features and apply ML-based techniques to classify using these low dimensional features [146]. The proposed approach applies feature-based (FB) time series classification [146].
The classifier first gets the input χ that includes the simulation results between the $0.5[s]$ before, and $1.5[s]$ after the disturbance as:

$$\chi = \begin{bmatrix} \hat{x}(t_e - 0.5), \hat{x}(t_e - 0.5 + \tau), .., \hat{x}(t_e + 1.5) \\ \hat{y}(t_e - 0.5), \hat{y}(t_e - 0.5 + \tau), .., \hat{y}(t_e + 1.5) \end{bmatrix}, \tag{5.19}$$

where $t_e[s]$ is the event time. Since the proposed coordination uses ADRR to simulate potential events as in Fig.5.1, t_e is an input to the simulation. Considering the typically

small τ in dynamic simulations, χ can be a high dimensional matrix. The proposed approach applies the feature extractor function F^{FE} to analyze the time series of each signal in χ and get a low dimensional vector π as:

$$\pi = F^{FE}(\chi). \tag{5.20}$$

The F^{FE} estimates correlation coefficients across all signals. For each signal, F^{FE} estimates the variance, discrete Fourier transform coefficients, the mean power of frequency bands from the signal power spectrum, and the number of peaks in the signal. The classifier FNN uses the π feature vector and labels the output as highly instable or not:

$$\hat{y}_\chi = F^{CL}_{w_\chi}(\pi) \in \{0,1\}, \tag{5.21}$$

where $F^{CL}_{w_\chi}, w_\chi$ are the FNN classifier and its weights.

5.4. CASE STUDIES

The case studies showcase the proposed coordination's components in classifying IBR types, selecting IBR setpoints to minimize post-fault instability risks, providing ADRR, and classifying highly instable conditions.

5.4.1. TEST SYSTEM AND SYNTHETIC DATA GENERATION

The system used includes the synthetic TS and DS models of the Zeeland province in the Netherlands, shown in Fig.5.4. The diagram includes the DS of which the ADRR model is developed (surrounded by red dashed lines), the TS, the DER of a second DS (surrounded by blue dashed lines), and the external grid for the rest of the TS. The TS has 150$[kV]$ and 380$[kV]$ buses. The DS has 52.5$[kV]$ and 10.6$[kV]$ buses. The model is in DigSilent Powerfactory. TSO TenneT provided the TS and [147], [148] the DSs. The DSs were modified to include the 2024 DER capacity and load consumptions. The DS line and transformer capacity were increased by 4 times from [147] to accommodate the increased DER capacity. The system includes GFM IBR (Droop Control System), GFL IBR (WECC WT Control System Type 4A), and synchronous machines.
Two sets of data were generated: one for the IBR classifier and one for the ADRR model. The IBR classifier dataset includes 93 load event simulations, while the ADRR model dataset includes 1545 fault-event simulations (faulted lines were randomly sampled from the 9 options in Fig.5.4). Appendix D includes details on data generation.

CASE STUDY SETTINGS

The case study of Sec.5.4.2 investigates the threshold for the IBR classifier, comparing the application of $\Delta\widehat{P}$ to other RTI measurements $\Delta\widehat{i}$, and $\Delta\widehat{Q}$. A combination metric is also compared to $\Delta\widehat{P}$, which classifies an IBR as GFM or GFL if two or more thresholds are exceeded from $\Delta\widehat{P}$, $\Delta\widehat{i}$, and $\Delta\widehat{Q}$.
The case study of Sec.5.4.3 uses the test system including the DS, excluding the ADRR. The risk example assumed $C^{\iota}(e) \approx 400k$[€] due to outage, $C^{\eta}(e) \approx 100$[€/%], and the fault e happening once every 2 years.
The case study of Sec.5.4.4 compares the proposed ADRR model to baselines:

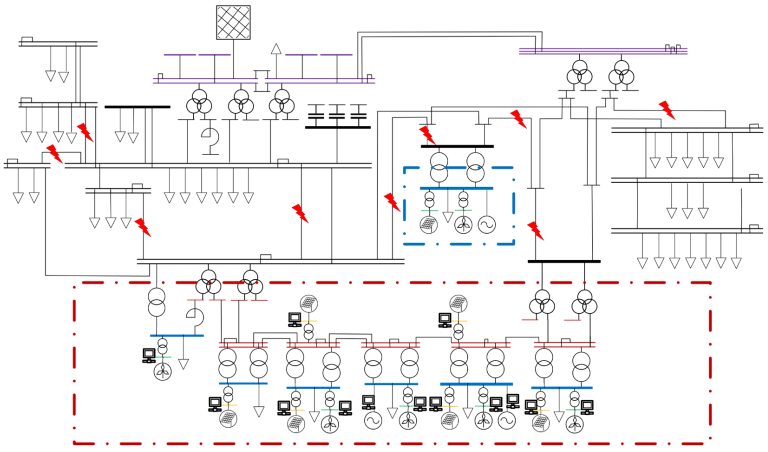


Figure 5.4: Test system diagram, with the DS, external grid, neighbouring DS, loads, PV modules, wind turbines, synchronous machines, RTI measurement locations. The indicate the simulated fault event locations.

(i) An FNN as [53], showcasing the impact of excluding LSTM in ADRR. This model includes the ADRR classifier.

(ii) An LSTM including sequential and repeated non-sequential input features followed by an FNN, showcasing the impact of excluding feature fusion and dense layer for non-sequential features in ADRR. This model is a natural extension of [52], [149] to include non-sequential features. This model includes the ADRR classifier.

(iii) The proposed structure, excluding the ADRR classifier. This model studies the impact of including extreme values observed in highly instable samples in ADRR.

(iv) An LSTM followed by an FNN, excluding non-sequential features, such as [52], [149]. This model studies the impact of excluding non-sequential features in ADRR.

The model comparison metric was the RMSE on a test set of simulated events.
The case study of Sec.5.4.5 analyzes the proposed ADRR classifier and compares it with baselines:

1. A classifier with additional features, the 5 coefficients when fitting the signal with the autoregressive moving-average (ARMA) model [150]. ARMA features are commonly used for FB time series classification as dynamic features [146].

2. A distance-based (DB) classifier [146], using dynamic time warping to estimate the distance between the signals as a similarity score. Using this similarity score, a k-nearest neighbors classifier predicted if each test sample was highly instable.

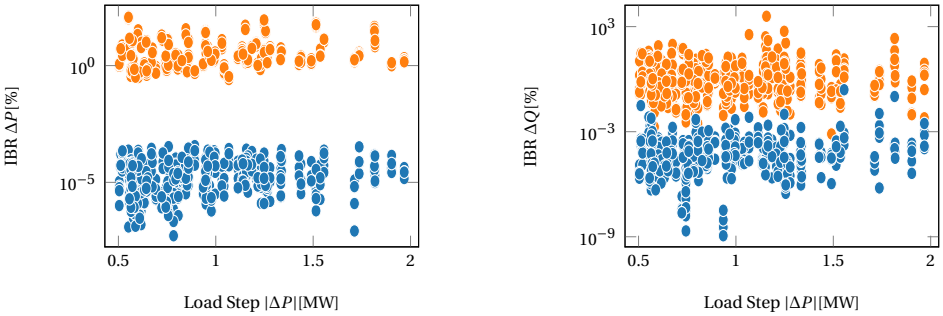
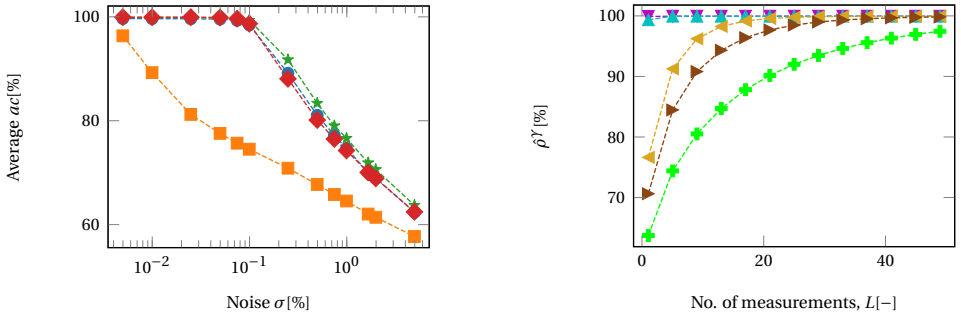


Figure 5.5: Changes in p, q on the GFM (●) and GFL (●) IBR terminals after load steps of different magnitudes. The $\Delta P, \Delta Q$ axes are in logarithmic scale.

The precision metric is the ratio of correctly classified samples as highly instable to all samples classified as highly instable. The recall metric is the ratio of samples correctly classified as highly instable to all highly instable samples. The accuracy metric is the ratio of correct predictions to all predictions.
Appendix D includes more details on model structure and training for the case studies.

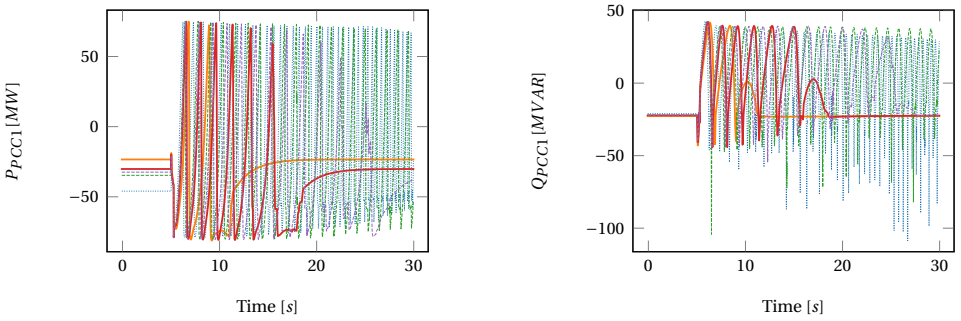
5.4.2. CLASSIFYING IBR TYPES

This case study analyses the selection of threshold $\Delta\widehat{P}$ of Sec.5.2.1. The threshold classifies each IBR as GFM or GFL using the pre- and post-load-event active powers. This study also explored the application of a current ($\Delta\widehat{i}$) threshold, a reactive power ($\Delta\widehat{Q}$) threshold, or the combination metric to classify an IBR as GFM or GFL.
Fig.5.5 shows a scenario where the weather conditions did not impact the IBR outputs within the measurement interval, i.e., $\epsilon(\sigma_l) = 0$. From this scenario, the changes in active power signified a clear distinction between the GFM and GFL responses. Thus, for $\Delta\widehat{P}$, any value between $[0.004, 0.255]$[%] returned a classification accuracy of $ac = 100$[%]. The value $\Delta\widehat{P} = 0.255$[%] was selected, as small threshold values are prone to weather and measurement noise. From the changes in the current and reactive power, there was a difference between the GFM and GFL responses, but no thresholds could separate the two types for all data. A $\Delta\widehat{i} = 0.222$[%] led to $ac = 99.8$[%], and $\Delta\widehat{Q} = 0.0084$[%] led to $ac = 98.9$[%].
Random noise from external weather factors impacting the measured signals affects the probability of correct classification by the above thresholds. Fig.5.6(a) shows the impact of different noise levels, σ, on the percentage of correctly classified events. Each point shows the average ac from 200 noise samples. The average ac for $\Delta\widehat{Q}$ deteriorated faster with the increase of σ, as $\Delta\widehat{Q} << \Delta\widehat{i}, \Delta\widehat{P}$ and a smaller threshold is more sensitive to noise. The results of Fig.5.6(a) indicate that the combination metric offers no advantage compared to $\Delta\widehat{P}$, which outperforms all thresholds. Therefore, the Algorithm 1 used $\Delta\widehat{P}$ as a metric. Let a probability of assigning the correct class to an IBR from a single event approximate the percentage of correct classifications as in Fig.5.6(a). Then, $\hat{\rho}^{\gamma}_{\phi} > \hat{\rho}^{\gamma}_{l}$ as



(a) Uncertainty impact. (b) Binomial distribution CMF.

Figure 5.6: In (a), the impact of added noise on IBR class prediction from a single event, using $\Delta\widehat{Q}$ (■), $\Delta\widehat{P}$ (★), $\Delta\widehat{I}$ (●) or the combination metric (◆). In (b), the contribution from using class assignments from multiple events as in (5.4), with added noise in ΔP with $\sigma = 0.01$[%] (▼), $\sigma = 0.1$[%] (▲), $\sigma = 1$[%] (◀), $\sigma = 2$[%] (▶), and $\sigma = 5$[%] (✚).



(a) (b)

Figure 5.7: The response of the DS at the PCC 1 after the same fault with setpoints of $P^s = 100$[%] (·····), $P^s = 50$[%] (——), $P^s = 75$[%] (- - - -), $\overline{P^s} = 65$[%] (——), and $P^s = 70$[%] (----)

ϕ increases. Fig.5.6(b) shows the contribution of multiple measurements with similar uncertainty levels in increased probability for correct prediction, using (5.4).

5.4.3. GFM IBR ACTIVE POWER SETPOINT OPTIMIZER

This case study investigates the setpoint optimizer of Alg. 2 in Sec.5.2.2. The setpoint optimizer simulates the DS response under different reduced GFM outputs to identify the maximum feasible setpoint and evaluate the risk of applying the setpoint. Fig.5.7 shows the simulated responses of different GFM setpoints. When reduced to 65[%] or 50[%], the post-disturbance DS response did not breach the stability constraints for $RoCoF$, under-/over frequency, or under/over voltage. Thus, $\overline{P^s} = 65$[%]. Without any action, $C^{\iota}(e) \approx 400k$[€], whereas applying (5.9) would suggest requesting $\overline{P^s} = 65$[%] if $\hat{\rho}^{o}(e) \geq 0.9$[%]. Deterministically controlling the setpoints whenever instability can happen ($\hat{\rho}^{o}(e) =$

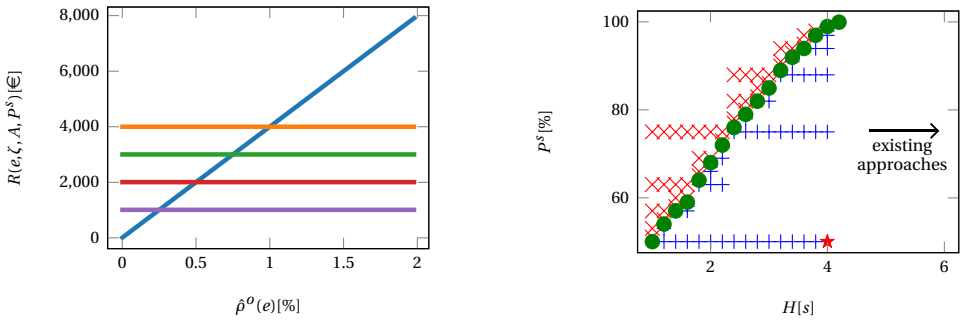


Figure 5.8: In (a) example risks for different fault event probabilities. Risks without reduction (—), or with reduction if $\overline{P^s} = 60[\%]$ (—), if $\overline{P^s} = 70[\%]$ (—), if $\overline{P^s} = 80[\%]$ (—), or if $\overline{P^s} = 90[\%]$ (—). In (b), the setpoint optimizer's search for maximum feasible $\overline{P^s}[\%]$ with $\delta P = 1[\%]$ for different levels of $H[s]$. Feasible (+), not-feasible (×, ★), and $\overline{P^s}$ (●).

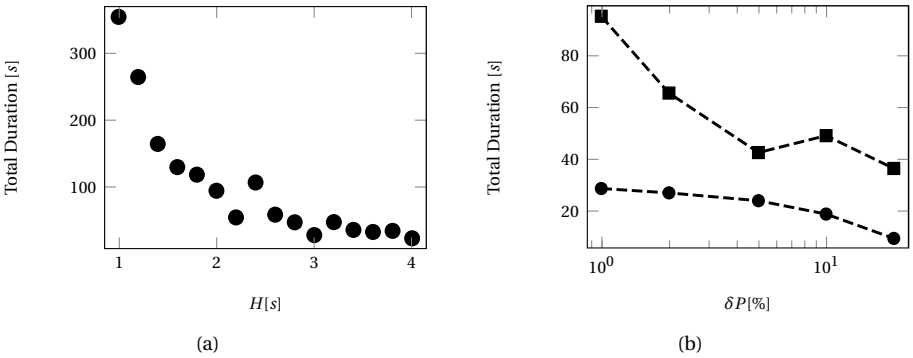


Figure 5.9: In (a), the total duration for the setpoint optimizer with $\delta P = 1[\%]$ for different TS H. In (b), the duration for simulations with varying δP for TS $H = 2[s]$ (■), and $H = 3[s]$ (●). δP axis is in logarithmic scale.

100[%]), could for example require applying $\overline{P^s} = 65[\%]$, 40 times per year, resulting in $\approx 280k$[€]. Probabilistically applying (5.9) only when $\hat{\rho}^o(e) > 0.9[\%]$ (e.g. 20 times per year) would result in $\approx 140k$[€].

Fig.5.8 shows how a low TS H requires a larger GFM IBR active power reduction to avoid reaching saturation after the disturbance. In TSs with low inertia, disturbances lead to larger frequency deviations and more volatile dynamic responses. GFM IBRs, which aim to regulate frequency and voltage through active and reactive power injection, can face higher demands after these disturbances. These demands can push the GFM IBRs to their current saturation limits. Upon reaching these limits, GFM IBRs can lose synchronous stability [127].

At $H = 4[s]$, a low setpoint value of 50[%] let to instable conditions. Therefore, for higher inertia $H \geq 3.5[s]$, setting $P^s_{min} \geq 25[\%]$ can avoid such issues and reduce the iterations of Alg.2. As TS H increases, the post-fault frequency deviations are less volatile. Therefore,

GFM IBRs require lower active power changes to address frequency deviations. Existing ADRR approaches represent TS with high inertia. Thus, in existing approaches, IBRs do not apply significant power output adjustments to control the system frequency and voltage and do not reach their saturation current.

Considering the computational burden, the duration for the GFM IBR active power setpoint optimizer can vary depending on the instability levels and resolution δP. Fig.5.9 shows the durations spent under varying TS inertia and δP using an i7-1185G7 CPU. For highly instable simulations, PowerFactory might need more numerical iterations, step adaptations, or matrix re-estimations, causing delays per simulation. Thus, for the cases with low TS inertia in Fig.5.9(a), the setpoint optimizer needed more time as more simulations were not feasible. However, the simulations in Fig.5.9(a) considered a $\delta P = 1$[%]. TSOs can increase the resolution to reduce the number of iterations for the setpoint optimizer. Fig.5.9(b) shows how increasing δP can reduce the optimizer duration. However, a δP increase can result in larger GFM IBR setpoint reductions, which could increase the costs for TSOs. As the proposed coordination considers preventive actions, this computational burden is not restrictive for TSOs.

5.4.4. ADRR REGRESSION MODEL

This case study compares the proposed ADRR regression model of Sec.5.3.1 to the baselines through the RMSE metric for predictions of the ADRR test set. Fig.5.10 shows example model predictions and RMSEs. Tab.5.1 compares the proposed model to baseline ADRR models. Tab.5.1 shows the proposed structure outperforms the fine-tuned baselines. Model (i) shows a higher mismatch between training and test errors, indicating a potential for overfitting compared to the other models.

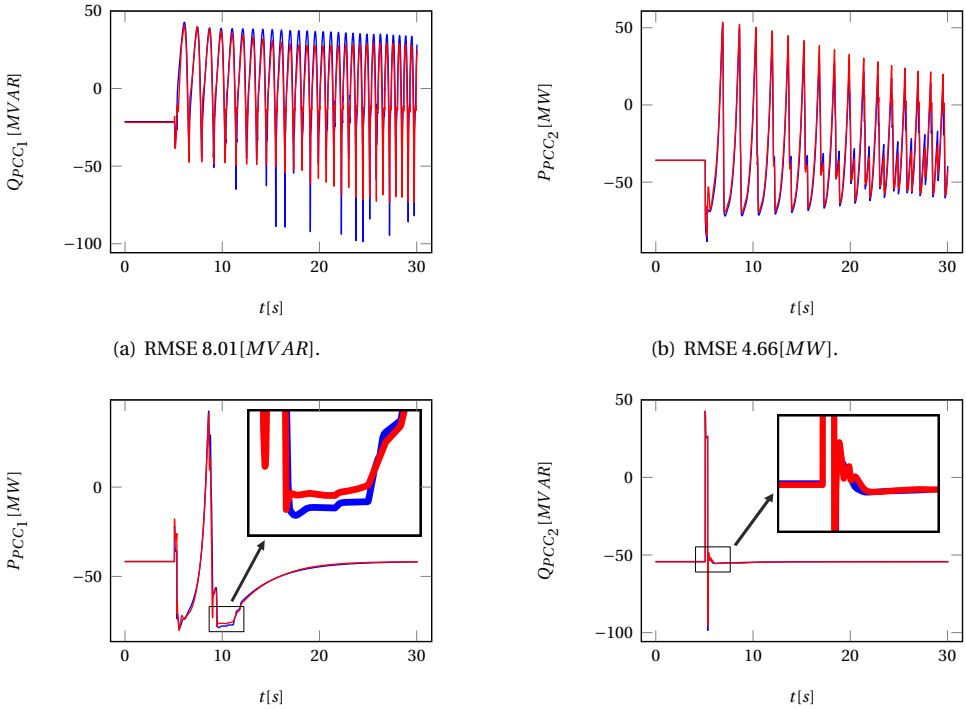
The training duration for model (ii) was more substantial than the rest, as the number of non-sequential features is larger than the sequential, and replicating their values is inefficient. All model training times allow frequent (e.g., monthly) re-training intervals.

The proposed ADRR model structure of Fig.5.3 outperforms the baseline structures. The model (i) structure is as [53], model (iv) structure is as [52], [149], whereas model (iii) is as [52], [149] extended to include non-sequential features in the LSTM inputs. Models (i), (ii), (iv), unlike [52], [53], [149], also include the classifier as it improves the RMSE, as shown by model (iii) performance compared to the proposed model. Comparing the proposed model with model (i), including RNNs is important for sequential features. Replicating and considering the non-sequential features as sequential is inefficient and can deteriorate the performance, as in model (ii). This result is expected as the non-sequential features outnumber the sequential. Removing the non-sequential features also deteriorates the performance, as in model (iv). Excluding the ADRR classifier for high instabilities burdens the regression model with high value and fluctuation prediction, deteriorating the model performance as with model (iii).

This case study also estimated the DS response using the proposed ADRR model for the setpoint optimizer task of Fig.5.7. Fig.5.11 shows the ADRR model response led to the exact conclusions as the simulated DS response. These conclusions were the maximum feasible setpoint of 65[%], with 50[%] being feasible, and 70, 75, 100[%] non-feasible. The ADRR outputs of Fig.5.11 approximate the simulations of Fig.5.7.

Table 5.1: Alternative Structures Performances

Model	Training RMSE	Validation RMSE	Test RMSE	Training [hours]
Proposed	0.83	0.98	0.94	8
(i)	1.06	1.29	1.29	7
(ii)	3.98	4.26	3.61	18
(iii)	2.96	3.2	3.03	8
(iv)	2.21	2.39	2.32	8



(a) RMSE 8.01$[MVAR]$.

(b) RMSE 4.66$[MW]$.

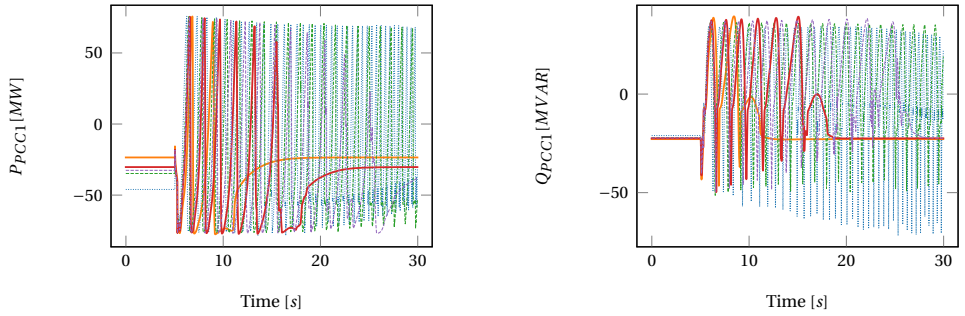
(c) RMSE 2.29$[MW]$.

(d) RMSE 0.27$[MVAR]$.

Figure 5.10: Example RMSE levels from model predictions (▬), compared to the simulated signals (▬).

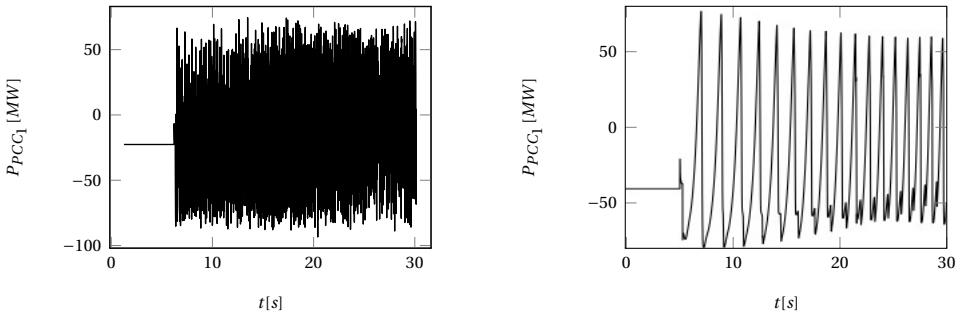
5.4.5. HIGH INSTABILITY CLASSIFICATION

This case study investigates the high instability classifier of Sec.5.3.2. Fig.5.12 shows examples of highly instable and regressor samples. Tab.5.2 compares the performance of the proposed FB model, the alternative FB model with added ARMA features, and the DB model. The prediction time indicates the delay each ADRR classifier adds to the ADRR, including the sample feature extraction. This classification delay only happens once per simulated event. Including the ARMA features does not improve the model accuracy, precision, or recall, but highly delays the prediction and feature extraction processes. The delay the proposed FB model adds once for each dynamic simulation (0.3$[s]$) is mi-



(a) (b)

Figure 5.11: The response of the ADRR model at PCC 1 after the same fault with setpoints of $P^s = 100[\%]$ (·····), $P^s = 50[\%]$ (——), $P^s = 75[\%]$ (- - - -), $\overline{P^s} = 65[\%]$ (——), and $P^s = 70[\%]$ (- - - -)



(a) Highly Instable Simulation. (b) Regressor Instable Simulation.

Figure 5.12: Example samples classified as highly instable (a) and not (b).

nor. The FB models outperform the DB in accuracy and recall. The 100[%] precision indicates that all models correctly identified all non-highly-instable simulations in the test set. The proposed high instability classifier shows the highest accuracy with the lowest delays among the options.

5.4.6. DISCUSSION

Considering the challenges for TSOs assessing the IBR flexibility for dynamic events, the case studies show the applicability of the proposed coordination. The IBR classifier can identify GFM and GFL IBRs considering ΔP using RTI measurements. The ADRR model represents the post-disturbance DS response. The GFM IBR active power setpoint optimizer can use the ADRR model's GFM IBR setpoint inputs to reduce instability risks. The identified setpoints when using the proposed ADRR model are similar to the simulations with the DS. Insights from the simulated responses are:

1. The system can withstand faults without reducing the GFM IBRs setpoints when

Table 5.2: Classification Model Performances

Model	Test Accuracy [%]	Test Precision [%]	Test Recall [%]	Feature Extraction and Training [minutes]	Prediction [s]
FB	99.7	100	98	0.63	0.3
FB with ARMA	99.7	100	98	110	4.7
DB	99.0	100	94.1	0.11	0.57

the TS inertia is high.

2. GFM IBRs require a headroom to avoid reaching their saturation current when the TS inertia is low.
3. The GFM IBR maximum feasible setpoint increases as the TS inertia increases.
4. Reducing the GFM IBR setpoints multiple times to prevent instabilities can cost less than suffering outages due to instabilities.
5. Probabilistically reducing the setpoints, considering instability risks, reduces the costs of preventing instability.

This research does not refute the GFM IBRs' positive impact on system stability and TSO support by providing inertia and fast frequency response. Nevertheless, this research offers a coordination process to limit the instability risks from the IBRs' current saturation issues. Using the proposed coordination, TSOs can also inform DSOs and IBR about the reason for the requested setpoint, as RTI requires.
The proposed coordination utilizes existing RTI frameworks for automated control of IBR by DSOs. Existing RTI applications consider capacity management, but future applications include voltage quality, system protection, and recovery applications [141] relating to the proposed application. The proposed communication between the TSO and DSOs does not require extensive sensitive information, as the ADRR does not include DS models. TSO setpoints for IBR can have values between $0-99\%$, and the reason can be encoded to a single digit. Thus, communication requirements are minimal.
The proposed application of the ADRR model does not directly control safety components, but rather informs TSOs of possible actions to prevent instability, involving human oversight. Therefore, to the best of the author's knowledge, the proposed coordination should comply with the European Union's AI Act [151]. Nevertheless, TSOs and DSOs should proceed with further considerations and evaluations before considering the proposed coordination to further ensure compliance with the AI Act and other European Union and national laws.
The test set included scenarios unseen during ADRR training, which varied in operating conditions, TS inertia, fault locations, and fault durations. The proposed ADRR results indicated a low RMSE for this test set and variability of scenarios. However, topological or structural changes in DSs can challenge data-driven ADRR models. Therefore, for structural and topological changes, additional data generation and ADRR model retraining or calibration are needed. A further limitation, the proposed approach did not

co-simulate the ADRR model and transmission network, due to the absence of ML-based models in PowerFactory dynamic simulations. Therefore, DSOs and TSOs can validate and adopt the proposed coordination when dynamic simulation software accepts ML-based models.

5.5. CONCLUSION

The proposed TSO-DSO coordination method allows TSOs to simulate the DS response to potential faults, evaluate the impact of different GFM IBR setpoints, and control these setpoints to prevent instabilities. This application addresses one of the key challenges recognized by the European DSO Entity and ENTSO-E [9]; the system's resilience through shared risk assessment, which requires risk assessment models, data-sharing, and coordination across stakeholders. The coordination uses RTI to classify the IBR types, identify the DS condition, and control the IBR setpoints. The proposed ADRR approach accounts for the transmission system's dynamic properties and share of IBR types that impact the DS response to dynamic events. The proposed ADRR model also allows TSOs to modify DS GFM IBR setpoints and analyze their impact. The proposed risk-based evaluation and setpoint selection can identify the TSO's minimum-risk setpoints.

In the Netherlands, the IBR response to dynamic events can impact the stability in TS with low inertia (e.g., below 4[s]). Thus, controlling the IBR setpoints under low-inertia TSs reduces instability risks, considering the probability of a dynamic event. Including the TS dynamic properties, DS operating condition, and IBR share improves the prediction performance by 59[%] (RMSE reduced from 2.32 to 0.94).

Future work includes exploring flexibility areas in the proposed coordination process to inform TSOs on the potential feasibility of the IBR control setpoints on the DS steady-state constraints. The adaptability of ADRR models to different topologies should also be studied. More diverse GFM and GFL models within the DS will also be explored in future work. The impact of GFM inverter inertial response provision will be considered and analyzed in subsequent research. Later work will also explore the expansion of grey box models for comparative analysis with the proposed ADRR model. Subsequent studies will explore the preventive capabilities of the proposed coordination for cascading faults.

6

CONCLUSION

This research addressed key challenges in TSO–DSO coordination by focusing on the estimation and activation of flexibility from DS. This research's outputs include algorithms to estimate the aggregated DS system flexibility, including reliability metrics with computational efficiency. A probabilistic method was introduced to quantify the uncertainty associated with these flexibility estimates, ensuring more robust decision-making under real-world variability. Additionally, this research proposed a novel approach for TSO-DSO coordination, considering the dynamic stability of the interconnected TSs and DSs. This work also released an open-source Python package that implements methods to estimate and visualize aggregated DS flexibility areas. This tool aims to support transparency, reproducibility, and practical uptake by grid operators, FSPs, and researchers.
The societal and industrial relevance of this work is significant. With the growing penetration of renewable energy and electrification of demand, grid flexibility is becoming a cornerstone for ensuring secure and cost-effective power system operation. The proposed methods offer system operators a practical pathway to assess and unlock the flexibility potential of DS without compromising operational constraints. TSOs can use the developed coordination strategies to enhance system stability by dynamically interacting with DSOs in a structured and data-informed manner. FSPs, on the other hand, benefit from clear operational boundaries and confidence intervals that improve market participation strategies.
By bridging methodological innovation with real-world applicability, this research contributes to a more integrated, resilient, and decarbonized energy system. The open-source nature of the developed tools further supports knowledge transfer, training, and collaborative innovation between academia, industry, and regulatory stakeholders.

6.1. RECAP OF THE PROBLEM AND OBJECTIVES

The stated objective of this research is to "to develop approaches to estimate the steady-state and dynamic flexibility in distribution systems under limited observability to support secure TSO-DSO operation", which aims to address research questions Q1, Q2, Q3,

Q4, and Q5. This objective highlights the identified need to increase the coordination between TSOs and DSOs to utilize distributed flexibility. Toward realizing this need, algorithms need to consider the limited observability in distribution systems and the complexity in simulating and evaluating the operational stability of interconnected TS and DS systems. The developed approaches and algorithms in Chapters 3-5 address the research questions that comprise this research objective.

6.2. Q1: How can flexibility area estimation algorithms include reliability metrics with computational efficiency?

To answer this research question, a new FA algorithm was proposed in Chapter 3. To include reliability metrics, the proposed algorithm evaluates the feasibility of all possible flexibility combinations. This evaluation enables the algorithm to include a metric on the density of feasible combinations leading to each operating point. A higher density of feasible flexibility combinations indicates more options to achieve an operating condition using flexibility. More options indicate less dependency on specific FSP setpoint combinations, and safer conditions considering the DS constraints.

The proposed algorithm reconsiders simplifications of prior approaches that focused on exploring the range of aggregated flexibility. The proposed approach, to effectively evaluate the feasibility of all discrete flexibility combinations, also addressed the more fundamental question "how can the feasibility of each flexibility combination be evaluated with computational tractability?" Convolution operations can aggregate all discrete combinations between FSPs, and can represent the number of combinations leading to each FA point. Nevertheless, to evaluate the feasibility of each flexibility combination, the approach in Chapter 3 modifies the convolution operations, by excluding the summation part of the estimation and using tensors to store the flexibility combination impacts on network constraints.

6.3. Q2: How can flexibility area estimation algorithms generalize to meshed and radial network topologies, including disjoint and non-convex flexibility areas?

To answer this research question, the proposed FA estimation algorithm developed a convolution-based approach, without alternative aggregation techniques such as Minkowski sum, and without mathematical simplifications that linearize FSP flexibility in prior OPF-based FA approaches. The convolution operations work for disjoint and non-convex FAs, and the proposed FA algorithms' formulations do not depend to the DS topology (e.g., generalize to both radial and meshed networks). The algorithm proposed in Chapter 3 shows the approach and validates this generalization to meshed and radial DS topologies, and disjoint and non-convex FAs. The approach of Chapter 4 also extends the application of FA estimation algorithms to FAs with 2−PCCs. In answering this research

question, the below more fundamental questions were initially addressed:

1. How can the flexibility from multiple resources be aggregated, including non convexities? As indicated in Chapter 3 and Appendix B, the proposed approach recognized how convolution can aggregate flexibility from multiple resources, including non-convex shapes. This characteristic differs from prior approaches, which applied convex simplifications or the Minkowski sum for fast flexibility aggregation.

2. How can flexibility estimation algorithms include flexibility resources with a non-continuous set of flexibility setpoints? The proposed approach in Chapter 3 utilizes the Dirac function convolution property to effectively consider non-continuous flexibility setpoints.

6.4. Q3: How can flexibility area estimation incorporate the level of observability on distribution systems?

This question was first addressed by exploring the more fundamental question "What is the impact of limited DS observability on estimated FAs?". Initial results indicated that the limited observability in DSs can impact the FA feasibility levels, where operating conditions of unobserved DS components can be closer or more distant to the DS constraints, altering the set of feasible flexibility combinations [152].

To incorporate the uncertainty in DSs from limited observability, this research developed a BNN-based FA estimation approach in Chapter 4. In this approach, TSOs can select a safety percentage on flexibility feasibility, considering uncertainty levels due to limited DS observability. BNNs estimate the uncertainty levels for the measured DS conditions. The proposed approach uses these uncertainty levels and the safety percentage to add safety margins when evaluating the feasibility of flexibility combinations. The BNNs include epistemic and aleatoric uncertainty in their estimations.

6.5. Q4: What is the impact of controlling inverter-based resources on the overall system's dynamic stability?

Addressing this research question, Chapter 5 considered the existing literature, modeled GFM and GFL IBRs in DS, and examined various operating conditions between the TS and the DS in the Dutch Zealand region while simulating dynamic events such as TS line faults. The results indicated that the current saturation issue of GFM IBR can bring instability in low-inertia TSs after large disturbances such as line faults. Under low-interia TSs, large disturbances lead to more volatile dynamic responses. GFM IBRs rapidly adapt their output to maintain the network voltage and frequency with limited support from the TS. Therefore, GFM IBRs need larger headroom to avoid reaching current saturation. Therefore, controlling GFM IBR to increase their headroom before reaching saturation current in low inertia networks can reduce post-fault instability risks.

6.6. Q5: HOW CAN TRANSMISSION AND DISTRIBUTION SYSTEM OPERATORS COORDINATE TO CONTROL DISTRIBUTION SYSTEM RESOURCES IN REAL-TIME TO ENSURE STABILITY?

To control DS resources in real-time to ensure stability, a TSO-DSO coordination should consider the available measurements in DSs, the technical and computational feasibility of the solution, and data confidentiality. Addressing this research question, Chapter 5 considered computational burdens, TSO-DSO confidentiality, and existing platforms in the Netherlands, to enable real-time monitoring and controlling of DS IBRs. Through these considerations, Chapter 5 introduced a new approach for TSO-DSO coordination and showed its potential impact for instability risk minimization. In, the proposed coordination, the DSO develops a machine learning-based ADRR model representing the DS dynamic response to faults, considering the IBR outputs and DS operating condition. The TSO uses the ADRR model to simulate the dynamic response of the interconnected TS and DS to possible faults. If instability is detected, a GFM IBR active power setpoint optimizer finds the maximum GFM IBR setpoints that can provide enough headroom to avoid instability due to the probable events. If reducing the GFM IBR setpoints can minimize instability risks, the TSO informs the DSO of the new setpoints. The DSO validates the steady-state feasibility of these setpoints and then requests them from the IBR through RTI.

6.7. DISCUSSION AND FUTURE WORK

This research develops algorithms and ML-based approaches to address key issues in power system operation, on TSO-DSO coordination, and DS flexibility utilization. These approaches span from steady-state to dynamic simulations and propose algorithms spanning from convolution operations to LSTMs and BNNs. Nevertheless, power systems require further adaptation of operational methods to enable this transition to a green, more distributed power grid.

Future work will mainly be on developing, exploring, and utilizing new optimization-, and ML-based approaches to support the transition to green and decentralized energy systems. Considering the similarity between graphs and power system networks, future work includes further exploration of GNN approaches for power system applications. In addition, the potential for federated learning for cross-organizational collaboration will be explored. Finally, future work will continue exploring the utilization of probabilistic machine learning models to consider the uncertainty in power systems tasks.

ACKNOWLEDGEMENTS

Pursuing a PhD is a challenging journey that would not have been possible without the support, guidance, and encouragement of many people to whom I am deeply grateful.
First and foremost, I would like to express my sincere gratitude to my supervisors, Assoc. Prof. Jochen Cremer, Assoc. Prof. Jose Rueda Torres, and Prof. Mart van der Meijden, for their invaluable guidance, constructive feedback, and continuous support throughout my PhD. Their mentorship has been instrumental in shaping my academic development and research outlook.
I am also thankful to all my colleagues from the IEPG group at TU Delft, and the AI for Energy Lab. I am particularly thankful to my TU Delft colleagues, including Ali, Haiwei, Mert, Nanda, Ajay, Ensieh, Neda, Kutay, Ola, Stavros, Jochen, Alamin, Nikolina, Vetri, Shengren, Hongjin, Ioannis, Mojtaba, Mohammad, Shabnam, Wouter, Viktor, and Basel. I greatly appreciated the insightful discussions, productive collaborations, and the vibrant academic environment that made this journey both intellectually stimulating and personally rewarding. Furthermore, I also thank the MegaMind researchers and industrial partners for the discussions, sessions, and collaboration. I am thankful to TenneT partners Matija and Daria for the discussion, organizing events, and their support. I appreciated the collaboration and exchange of ideas with the MegaMind researchers, Brenda, Sascha, Andrey, Qiao, Haoyang, Eslam, Leander, Ivo, and Aliene.
My academic path has also been shaped by my earlier mentors. I am grateful to Prof. Charalambos Charalambous for his consistent mentorship, advice, and encouragement throughout my academic journey. His guidance helped me build a strong foundation in my academic work. I would also like to thank Assoc. Prof. Petros Aristidou for his support, constructive insights, and for fostering a strong sense of academic curiosity that has stayed with me throughout my career.
On a personal level, I am deeply indebted to my wife, Sofia Archontissa Lamprianidi, for her unwavering support and patience, and for making the bold decision to move to a different country with me during this journey. Her presence and encouragement have been a constant source of strength. I would also like to thank my brother, Michalis, for his continuous advice and support, and for sharing his own experiences. My sister, Antria, deserves heartfelt thanks for her optimism, warmth, and constant care. I am also grateful to my father, Chrysostomos, and my mother, Marigianna, for their endless support, thoughtful advice, encouragement, and the unique way that they always find to be present even across distances.
I am also grateful to my friends in the Netherlands, who enriched our everyday life and made this journey far more enjoyable. I want to thank Ioannis, Yulia, Diamanto, Mikael, Iakovos, Dianne, Emre, Ximena, Carlos, Chahil, Joe, Anthony, Kevin, Darragh, Demetris, Chyannie, Saeed, Sara, and Cemre for their friendship, and the many shared moments.

To all of you, thank you.

BIBLIOGRAPHY

[1] D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, "Exploring operational flexibility of active distribution networks with low observability", in *IEEE Belgrade PowerTech*, 2023. DOI: 10.1109/PowerTech55446.2023.10202841.

[2] L. I. Dulău, M. Abrudean, and D. Bică, "Distributed generation and virtual power plants", in *2014 49th International Universities Power Engineering Conference (UPEC)*, IEEE, 2014, pp. 1–5.

[3] S. Ghazi and K. Ip, "The effect of weather conditions on the efficiency of pv panels in the southeast of uk", *Renewable Energy*, vol. 69, pp. 50–59, 2014, ISSN: 0960–1481. DOI: https://doi.org/10.1016/j.renene.2014.03.018.

[4] J. K. Kaldellis, M. Kapsali, and K. A. Kavadias, "Temperature and wind speed impact on the efficiency of pv installations. experience obtained from outdoor measurements in greece", *Renewable Energy*, vol. 66, pp. 612–624, 2014, ISSN: 0960–1481. DOI: https://doi.org/10.1016/j.renene.2013.12.041.

[5] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation", *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012, ISSN: 0960–1481. DOI: https://doi.org/10.1016/j.renene.2011.05.033.

[6] M. S. Nazir, F. Alturise, S. Alshmrany, *et al.*, "Wind generation forecasting methods and proliferation of artificial neural network: A review of five years research trend", *Sustainability*, vol. 12, no. 9, p. 3778, 2020.

[7] F. Economics, "Review of flexibility platforms", Entso-E, Tech. Rep., 2021.

[8] P. system operation and control, "Tso-dso co-operation control centre tools requirements", CIGRE, Tech. Rep., 2021.

[9] "Report tso-dso challenges & opportunities for digital electricity system", ENTSO-E, Tech. Rep., Mar. 2025, Accessed: 2025-04-30. [Online]. Available: https://www.entsoe.eu/news/2025/03/06/report-tso-dso-challenges-opportunities-for-digital-electricity-system/.

[10] A. G. Givisiez, K. Petrou, and L. F. Ochoa, "A review on tso-dso coordination models and solution techniques", *Electric Power Systems Research*, vol. 189, p. 106 659, 2020.

[11] P. P. Barker and R. W. De Mello, "Determining the impact of distributed generation on power systems. i. radial distribution systems", in *2000 Power Engineering Society Summer Meeting (Cat. No. 00CH37134)*, IEEE, vol. 3, 2000, pp. 1645–1656.

[12] M. Begovic, A. Pregelj, A. Rohatgi, and D. Novosel, "Impact of renewable distributed generation on power systems", in *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society, vol. 3, 2001, pp. 2008–2008.

[13] A. Ishchenko, A. Jokic, J. Myrzik, and W. Kling, "Dynamic reduction of distribution networks with dispersed generation", in *2005 International Conference on Future Power Systems*, IEEE, 2005, 7–pp.

[14] U. Kuhar, M. Pantoš, G. Kosec, and A. Švigelj, "The impact of model and measurement uncertainties on a state estimation in three-phase distribution networks", *IEEE Transactions on Smart Grid*, vol. 10, no. 3, 2019.

[15] J. Vasiljevska and A. Marinopoulos, "The role of tso-dso cooperation towards the energy transition", 2019.

[16] P. Olivella-Rosell, P. Lloret-Gallego, Í. Munné-Collado, *et al.*, "Local flexibility market design for aggregators providing multiple flexibility services at distribution network level", *Energies*, vol. 11, no. 4, 2018, ISSN: 1996-1073. DOI: 10.3390/en11040822. [Online]. Available: https://www.mdpi.com/1996-1073/11/4/822.

[17] N. Hekmat, H. Cai, T. Zufferey, G. Hug, and P. Heer, "Data-driven demand-side flexibility quantification: Prediction and approximation of flexibility envelopes", *arXiv preprint arXiv:2110.12796*, 2021.

[18] M. Kalantar-Neyestanaki, F. Sossan, M. Bozorg, and R. Cherkaoui, "Characterizing the reserve provision capability area of active distribution networks: A linear robust optimization method", *IEEE Transactions on Smart Grid*, vol. 11, no. 3, 2019.

[19] J. Silva, J. Sumaili, R. J. Bessa, L. Seca, M. Matos, and V. Miranda, "The challenges of estimating the impact of distributed energy resources flexibility on the tso/dso boundary node operating points", *Computers & Operations Research*, vol. 96, pp. 294–304, 2018.

[20] D. M. Gonzalez, J. Hachenberger, J. Hinker, F. Rewald, *et al.*, "Determination of the time-dependent flexibility of active distribution networks to control their tso-dso interconnection power flow", in *Power Systems Computation Conference (PSCC)*, IEEE, 2018.

[21] M. Bolfek and T. Capuder, "An analysis of optimal power flow based formulations regarding dso-tso flexibility provision", *International Journal of Electrical Power & Energy Systems*, vol. 131, p. 106 935, 2021.

[22] F. Capitanescu, "Tso–dso interaction: Active distribution network power chart for tso ancillary services provision", *Electric Power Systems Research*, vol. 163, pp. 226–230, 2018.

[23] N. Savvopoulos, C. Y. Evrenosoglu, T. Konstantinou, T. Demiray, and N. Hatziargyriou, "Contribution of residential pv and bess to the operational flexibility at the tso-dso interface", in *International Conference on Smart Energy Systems and Technologies (SEST)*, IEEE, 2021.

[24] M. Heleno, R. Soares, J. Sumaili, R. J. Bessa, L. Seca, and M. A. Matos, "Estimation of the flexibility range in the transmission-distribution boundary", in *IEEE Eindhoven PowerTech*, IEEE, 2015.

[25] G. Prionistis, C. Vournas, and M. Vrakopoulou, "A fast method to approximate the flexibility region of an active distribution network in pq space", in *IEEE Belgrade PowerTech*, 2023. DOI: 10.1109/PowerTech55446.2023.10202983.

[26] A. Churkin, M. Sanchez-Lopez, M. I. Alizadeh, F. Capitanescu, *et al.*, "Impacts of distribution network reconfiguration on aggregated der flexibility", in *IEEE Belgrade PowerTech*, 2023. DOI: 10.1109/PowerTech55446.2023.10202791.

[27] B. Azimian, R. S. Biswas, S. Moshtagh, A. Pal, L. Tong, and G. Dasarathy, "State and topology estimation for unobservable distribution systems using deep neural networks", *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–14, 2022. DOI: 10.1109/TIM.2022.3167722.

[28] B. Habib, E. Isufi, W. v. Breda, A. Jongepier, and J. L. Cremer, "Deep statistical solver for distribution system state estimation", *IEEE Transactions on Power Systems*, 2023. DOI: 10.1109/TPWRS.2023.3290358.

[29] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation", *IEEE Transactions on Power Systems*, vol. 32, no. 5, 2017. DOI: 10.1109/TPWRS.2016.2632156.

[30] T. Chen, Y. Song, D. J. Hill, and A. Y. Lam, "Enhancing flexibility at the transmission-distribution interface with power flow routers", *IEEE Transactions on Power Systems*, vol. 37, no. 4, 2021.

[31] A. Sajadi, J. A. Rañola, R. W. Kenyon, B.-M. Hodge, and B. Mather, "Dynamics and stability of power systems with high shares of grid-following inverter-based resources: A tutorial", *IEEE Access*, vol. 11, pp. 29 591–29 613, 2023. DOI: 10.1109/ACCESS.2023.3260778.

[32] I. Green, U. G. I. C. G. LAMMERT, D. J. C. M. P. LD, *et al.*, "Modelling of inverter-based generation for power system dynamic studies", 2018.

[33] P. Wang, Z. Zhang, Q. Huang, X. Tang, and W.-J. Lee, "Robustness-improved method for measurement-based equivalent modeling of active distribution network", *IEEE Transactions on Industry Applications*, vol. 57, no. 3, 2021. DOI: 10.1109/TIA.2021.3057358.

[34] B. Barać, M. Krpan, and T. Capuder, "Aggregated model of a flexible multi-energy distribution system for coordinated tso-dso fast frequency services provision", in *International Conference on Smart Energy Systems and Technologies (SEST)*, 2023. DOI: 10.1109/SEST57387.2023.10257335.

[35] F. Conte, F. D'Agostino, and F. Silvestro, "Operational constrained nonlinear modeling and identification of active distribution networks", *Electric Power Systems Research*, vol. 168, pp. 92–104, 2019, ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2018.11.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037877961830381X.

[36] D. Chrysostomou, J. L. R. Torres, and J. L. Cremer, "Tensor convolution-based aggregated flexibility estimation in active distribution systems", *IEEE Transactions on Smart Grid*, 2024.

[37] B. Mohandes, M. S. El Moursi, N. Hatziargyriou, and S. El Khatib, "A review of power system flexibility with high penetration of renewables", *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3140–3155, 2019.

[38] J. Silva, J. Sumaili, R. J. Bessa, *et al.*, "Estimating the active and reactive power flexibility area at the tso-dso interface", *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4741–4750, 2018.

[39] E. Lannoye, D. Flynn, and M. O'Malley, "Evaluation of power system flexibility", *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 922–931, 2012. DOI: 10.1109/TPWRS.2011.2177280.

[40] H. Nosair and F. Bouffard, "Flexibility envelopes for power system operational planning", *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 800–809, 2015.

[41] W. D. Blizard *et al.*, "Multiset theory", *Notre Dame Journal of formal logic*, vol. 30, no. 1, 1989.

[42] J. Lygeros and F. Ramponi, "Lecture notes on linear system theory", *Automatic Control Laboratory, ETH Zurich*, 2010.

[43] A. Churkin, W. Kong, J. N. M. Gutierrez, P. Mancarella, and E. A. M. Cesena, "Characterizing power support from distribution networks via flexibility area segmentation", *arXiv preprint arXiv:2110.01086*, 2021.

[44] D. A. Contreras and K. Rudion, "Improved assessment of the flexibility range of distribution grids using linear optimization", in *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018, pp. 1–7.

[45] G. Chaspierre, "Reduced-order modelling of active distribution networks for large-disturbance simulations", Ph.D. dissertation, Universite de Liege, Liege, Belgique, 2020.

[46] J. V. Milanović and S. M. Zali, "Validation of equivalent dynamic model of active distribution network cell", *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2101–2110, 2012.

[47] M. J. Reno, R. J. Broderick, and S. Grijalva, "Formulating a simplified equivalent representation of distribution circuits for pv impact studies.", Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); Georgia ..., Tech. Rep., 2013.

[48] A. Jayawardena, L. G. Meegahapola, D. A. Robinson, and S. Perera, "Representation of a grid-tied microgrid as a reduced order entity for distribution system dynamics and stability studies", *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 591–600, 2015.

[49] M. Glavic, "Dynamic equivalents of active distribution networks: A short review", 2016.

[50] R. A. Ramos, A. P. Grilo-Pavani, A. B. Piardi, T. C. Fernandes, and M. E. Bento, "Method to build equivalent models of microgrids for rms dynamic simulation of power systems", *arXiv preprint arXiv:2110.06160*, 2021.

[51] N. Fulgêncio, B. Silva, J. Villar, *et al.*, "Flexibility hub's dynamic equivalent model: Improving the representation of the distribution grid for system planning", in *CIRED 2020 Berlin Workshop (CIRED 2020)*, IET, vol. 2020, 2020, pp. 657–660.

[52] C. Zheng, S. Wang, Y. Liu, *et al.*, "A novel equivalent model of active distribution networks based on lstm", *IEEE transactions on neural networks and learning systems*, vol. 30, no. 9, pp. 2611–2624, 2019.

[53] A. M. Azmy, I. Erlich, and P. Sowa, "Artificial neural network-based dynamic equivalents for distribution systems containing active sources", *IEE Proceedings-Generation, Transmission and Distribution*, vol. 151, no. 6, 2004.

[54] X. Shang, Z. Li, J. Zheng, and Q. Wu, "Equivalent modeling of active distribution network considering the spatial uncertainty of renewable energy resources", *International Journal of Electrical Power & Energy Systems*, vol. 112, 2019.

[55] D. Greene, P. Cunningham, and R. Mayer, "Unsupervised learning and clustering", in *Machine learning techniques for multimedia: Case studies on organization and retrieval*, Springer, 2008, pp. 51–90.

[56] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, "Unsupervised learning", in *An introduction to statistical learning: with applications in Python*, Springer, 2023, pp. 503–556.

[57] R. S. Sutton, A. G. Barto, *et al.*, "Reinforcement learning", *Journal of Cognitive Neuroscience*, vol. 11, no. 1, pp. 126–134, 1999.

[58] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement learning for selective key applications in power systems: Recent advances and future challenges", *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2935–2958, 2022. DOI: 10.1109/TSG.2022.3154718.

[59] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, "Linear regression", in *An introduction to statistical learning: With applications in python*, Springer, 2023, pp. 69–134.

[60] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013.

[61] L. Breiman, "Random forests", *Machine learning*, vol. 45, pp. 5–32, 2001.

[62] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system", in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[63] G. Ke, Q. Meng, T. Finley, *et al.*, "Lightgbm: A highly efficient gradient boosting decision tree", *Advances in neural information processing systems*, vol. 30, 2017.

[64] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: Unbiased boosting with categorical features", *Advances in neural information processing systems*, vol. 31, 2018.

[65] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backprop", in *Neural networks: Tricks of the trade*, Springer, 2002, pp. 9–50.

[66] C. M. Bishop, *Neural networks for pattern recognition*. Oxford university press, 1995.

[67] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects", *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.

[68] A. Graves and A. Graves, "Long short-term memory", *Supervised sequence labelling with recurrent neural networks*, pp. 37–45, 2012.

[69] F. M. Salem and F. M. Salem, "Gated rnn: The gated recurrent unit (gru) rnn", *Recurrent neural networks: from simple to gated architectures*, pp. 85–100, 2022.

[70] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review", *Computational Social Networks*, vol. 6, no. 1, pp. 1–23, 2019.

[71] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks", in *International Conference on Learning Representations*, 2018.

[72] E. Jiawei, Y. Zhang, S. Yang, H. Wang, X. Xia, and X. Xu, "Graphsage++: Weighted multi-scale gnn for graph representation learning", *Neural Processing Letters*, vol. 56, no. 1, p. 24, 2024.

[73] H. Zhou, S. Zhang, J. Peng, *et al.*, "Informer: Beyond efficient transformer for long sequence time-series forecasting", in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, 2021, pp. 11 106–11 115.

[74] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale", in *International Conference on Learning Representations*.

[75] J. Lee, Y. Lee, J. Kim, A. Kosiorek, S. Choi, and Y. W. Teh, "Set transformer: A framework for attention-based permutation-invariant neural networks", in *International conference on machine learning*, PMLR, 2019, pp. 3744–3753.

[76] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning", in *international conference on machine learning*, PMLR, 2016.

[77] C. Zhang, J. Bütepage, H. Kjellström, and S. Mandt, "Advances in variational inference", *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, 2018.

[78] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles", *Advances in neural information processing systems*, vol. 30, 2017.

[79] P. Gupta, M. Gasse, E. Khalil, P. Mudigonda, A. Lodi, and Y. Bengio, "Hybrid models for learning to branch", *Advances in neural information processing systems*, vol. 33, pp. 18 087–18 097, 2020.

[80] M. Zhang, M.-Á. Fernández-Torres, and G. Camps-Valls, "Hybrid recurrent neural network for drought monitoring", in *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*, 2022.

[81] N. Rodríguez Pérez, J. M. Domingo, G. L. López, *et al.*, "Ict architectures for tso-dso coordination and data exchange: A european perspective", *IEEE Transactions on Smart Grid*, vol. 14, no. 2, 2023. DOI: 10.1109/TSG.2022.3206092.

[82] N. Savvopoulos and N. Hatziargyriou, "An effective method to estimate the aggregated flexibility at distribution level", *IEEE Access*, vol. 11, 2023. DOI: 10.1109/ACCESS.2023.3262730.

[83] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants", *IEEE Transactions on Power Systems*, vol. 33, no. 2, 2018.

[84] S. Nowak, Y. C. Chen, and L. Wang, "Measurement-based optimal der dispatch with a recursively estimated sensitivity model", *IEEE Transactions on Power Systems*, vol. 35, no. 6, 2020.

[85] A. Churkin, W. Kong, J. N. M. Gutierrez, P. Mancarella, *et al.*, "Assessing distribution network flexibility via reliability-based pq area segmentation", in *IEEE Belgrade PowerTech*, IEEE, 2023.

[86] W. Dai, J. Jian, and S. Guo, "Fast characterization of nonlinear feasible region based on deep neural network association mining", *Energy Reports*, 2023.

[87] X. Pan, M. Chen, T. Zhao, and S. H. Low, "Deepopf: A feasibility-optimized deep neural network approach for ac optimal power flow problems", *IEEE Systems Journal*, 2023.

[88] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, and H. Yu, "Data-driven optimal power flow: A physics-informed machine learning approach", *IEEE Transactions on Power Systems*, 2021.

[89] X. Hu, H. Hu, S. Verma, and Z.-L. Zhang, "Physics-guided deep neural networks for power flow analysis", *IEEE Transactions on Power Systems*, 2021.

[90] A. Yaniv, P. Kumar, and Y. Beck, "Towards adoption of gnns for power flow applications in distribution systems", *Electric Power Systems Research*, 2023, ISSN: 0378-7796.

[91] S. Riaz and P. Mancarella, "Modelling and characterisation of flexibility from distributed energy resources", *IEEE transactions on power systems*, vol. 37, no. 1, pp. 38–50, 2021.

[92] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. USA: California Technical Publishing, 1997, ISBN: 0966017633.

[93] S. Poudel, Z. Ni, and W. Sun, "Electrical distance approach for searching vulnerable branches during contingencies", *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3373–3382, 2016.

[94] J. Klaus, M. Blacher, and J. Giesen, "Compiling tensor expressions into einsum", in *International Conference on Computational Science*, Springer, 2023.

[95] S. Repo, M. Attar, S. Repo, *et al.*, "Holistic view of needs and requirements of multi-market flexibility trading", in *CIRED 2020 Berlin Workshop (CIRED 2020)*, IET, vol. 2020, 2020.

[96] I. V. Oseledets, "Tensor-train decomposition", *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[97] E. Bisong and E. Bisong, "Google colaboratory", *Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners*, pp. 59–64, 2019.

[98] M. Q. Tran, "Voltage monitoring and control in the active distribution networks", 2024.

[99] E. Goan and C. Fookes, "Bayesian neural networks: An introduction and survey", *Case Studies in Applied Bayesian Data Science: CIRM Jean-Morlet Chair, Fall 2018*, 2020.

[100] A. G. Wilson and P. Izmailov, "Bayesian deep learning and a probabilistic perspective of generalization", *Advances in neural information processing systems*, vol. 33, 2020.

[101] N. Savvopoulos, N. Hatziargyriou, and H. Laaksonen, "A holistic approach to the efficient estimation of operational flexibility from distributed resources", *IEEE Open Access Journal of Power and Energy*, 2024.

[102] S. Ge, Z. Xu, H. Liu, C. Gu, and F. Li, "Flexibility evaluation of active distribution networks considering probabilistic characteristics of uncertain variables", *IET Generation, Transmission & Distribution*, vol. 13, no. 14, 2019.

[103] N. Majumdar, P. Kengkat, R. Yermekbayev, and L. Hofmann, "Reliability parameterised distribution grid flexibility aggregation considering renewable uncertainties", in *2023 58th International Universities Power Engineering Conference (UPEC)*, IEEE, 2023.

[104] R. J. De Groot, J. Morren, and J. G. Slootweg, "Closed-ring operation of medium voltage distribution grids: Theory meets practice", in *23rd International Conference and Exhibition on Electricity Distribution (CIRED)*, 2015.

[105] J. A. D. Massignan, J. B. A. London, M. Bessani, C. D. Maciel, R. Z. Fannucchi, and V. Miranda, "Bayesian inference approach for information fusion in distribution system state estimation", *IEEE Transactions on Smart Grid*, vol. 13, no. 1, 2022.

[106] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "State estimation for unobservable distribution systems via deep neural networks", in *2018 IEEE Power and Energy Society General Meeting (PESGM)*, 2018.

[107] P. A. Pegoraro, A. Angioni, M. Pau, *et al.*, "Bayesian approach for distribution system state estimation with non-gaussian uncertainty models", *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, 2017.

[108] M. Vanin, T. Van Acker, R. D'hulst, and D. Van Hertem, "Exact modeling of non-gaussian measurement uncertainty in distribution system state estimation", *IEEE Transactions on Instrumentation and Measurement*, vol. 72, 2023.

[109] K. R. Mestav, J. Luengo-Rozas, and L. Tong, "Bayesian state estimation for unobservable distribution systems via deep learning", *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4910–4920, 2019. DOI: 10.1109/TPWRS.2019.2919157.

[110] D. Chrysostomou, J. L. R. Torres, and J. L. Cremer, "Tensor convolution-based aggregated flexibility estimation in active distribution systems", *IEEE Transactions on Smart Grid*, 2024. DOI: 10.1109/TSG.2024.3453667.

[111] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?", *Advances in neural information processing systems*, vol. 30, 2017.

[112] S. K. Lind, Z. Xiong, P.-E. Forssén, and V. Krüger, "Uncertainty quantification metrics for deep regression", *Pattern Recognition Letters*, vol. 186, 2024.

[113] Y. Lu, Y. He, G. Wang, C. Fang, and Z. Sun, "Multi-target neurallss regression with learned confidence space", in *2024 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2024.

[114] L. O. Bentsen, N. D. Warakagoda, R. Stenbro, and P. Engelstad, "Relative evaluation of probabilistic methods for spatio-temporal wind forecasting", *Journal of Cleaner Production*, vol. 434, 2024.

[115] N. Tagasovska and D. Lopez-Paz, "Single-model uncertainties for deep learning", *Advances in neural information processing systems*, vol. 32, 2019.

[116] P. Weng, Y. Tian, Y. Liu, and Y. Zheng, "Time-series generative adversarial networks for flood forecasting", *Journal of Hydrology*, vol. 622, 2023, ISSN: 0022-1694.

[117] Q. Zhu, Y. Xu, Q. Lin, Z. Ming, and K. C. Tan, "Clustering-based short-term wind speed interval prediction with multi-objective ensemble learning", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 9, no. 1, 2025.

[118] J. Wang, S. Wang, B. Zeng, and H. Lu, "A novel ensemble probabilistic forecasting system for uncertainty in wind speed", *Applied Energy*, vol. 313, 2022, ISSN: 0306-2619.

[119] D. Chrysostomou, J. L. R. Torres, and J. L. Cremer, "Machine learning-based method to support tso-dso adaptive coordination with active power management for instability prevention", *International Journal of Electrical Power & Energy Systems*, vol. 173, p. 111 353, 2025.

[120] G. Chaspierre, G. Denis, P. Panciatici, and T. Van Cutsem, "An active distribution network equivalent derived from large-disturbance simulations with uncertainty", *IEEE Transactions on Smart Grid*, vol. 11, no. 6, 2020.

[121] A. Vicente-Pastor, J. Nieto-Martin, D. W. Bunn, and A. Laur, "Evaluation of flexibility markets for retailer–dso–tso coordination", *IEEE Transactions on Power Systems*, vol. 34, no. 3, 2018.

[122] S. I. Vagropoulos, P. N. Biskas, and A. G. Bakirtzis, "Market-based tso-dso coordination for enhanced flexibility services provision", *Electric Power Systems Research*, vol. 208, 2022.

[123] J. L. Cremer and G. Strbac, "A machine-learning based probabilistic perspective on dynamic security assessment", *International Journal of Electrical Power & Energy Systems*, vol. 128, 2021.

[124] P. Kundur, J. Paserba, V. Ajjarapu, *et al.*, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions", *IEEE Transactions on Power Systems*, vol. 19, no. 3, 2004.

[125] A. Quedan, W. Wang, D. Ramasubramanian, E. Farantatos, and S. Asgarpoor, "Behavior of a high inverter-based resources distribution network with different participation ratios of grid-forming and grid-following inverters", in *North American Power Symposium (NAPS)*, IEEE, 2021.

[126] D. Pattabiraman, R. H. Lasseter., and T. M. Jahns, "Comparison of grid following and grid forming control for a high inverter penetration power system", in *IEEE Power & Energy Society General Meeting*, 2018. DOI: 10.1109/PESGM.2018.8586162.

[127] H. Xin, L. Huang, L. Zhang, Z. Wang, and J. Hu, "Synchronous instability mechanism of p-f droop-controlled voltage source converter caused by current saturation", *IEEE Transactions on Power Systems*, vol. 31, no. 6, 2016. DOI: 10.1109/TPWRS.2016.2521325.

[128] S. Alshahrani, K. Khan, M. Abido, and M. Khalid, "Grid-forming converter and stability aspects of renewable-based low-inertia power networks: Modern trends and challenges", *Arabian Journal for Science and Engineering*, vol. 49, no. 5, 2024.

[129] P. Makolo, R. Zamora, and T.-T. Lie, "The role of inertia for grid flexibility under high penetration of variable renewables - a review of challenges and solutions", *Renewable and Sustainable Energy Reviews*, vol. 147, 2021, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2021.111223.

[130] ENTSO-E, *Inertia and rate of change of frequency (rocof)*, 2020. [Online]. Available: https://www.fingrid.fi/en/electricity-market-information/.

[131] M. Laasonen. "Inertia of the nordic power system". (), [Online]. Available: https://www.fingrid.fi/en/electricity-market-information/InertiaofNordicpowersystem/. (accessed: 29.04.2024).

[132] S. Dalhues, Y. Zhou, O. Pohl, *et al.*, "Research and practice of flexibility in distribution systems: A review", *CSEE Journal of Power and Energy Systems*, vol. 5, no. 3, 2019.

[133] G. Mitrentsis and H. Lens, "A gaussian process framework for the probabilistic dynamic modeling of active distribution networks using exogenous variables", *Electric Power Systems Research*, vol. 211, 2022, ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2022.108403.

[134] E. O. Kontis, T. A. Papadopoulos, M. H. Syed, E. Guillo-Sansano, G. M. Burt, and G. K. Papagiannis, "Artificial-intelligence method for the derivation of generic aggregated dynamic equivalent models", *IEEE Transactions on Power Systems*, vol. 34, no. 4, 2019.

[135] E. O. Kontis, A. R. del Nozal, S. C. Dimoulias, and J. M. Mauricio, "Dynamic equivalent model of active distribution networks providing frequency-related ancillary services to the transmission system", *Applied Energy*, vol. 367, 2024.

[136] J. Ungerland, N. Poshiya, W. Biener, and H. Lens, "A voltage sensitivity based equivalent for active distribution networks containing grid forming converters", *IEEE Transactions on Smart Grid*, vol. 14, no. 4, 2022.

[137] D. B. Rathnayake, M. Akrami, C. Phurailatpam, *et al.*, "Grid forming inverter modeling, control, and applications", *IEEE Access*, vol. 9, 2021. DOI: 10.1109/ACCESS.2021.3104617.

[138] I. Chychykina, Z. A. Styczynski, C. O. Heyde, and R. Krebs, "Power system instability prevention and remedial measures with online dynamic security assessment", in *2015 IEEE Eindhoven PowerTech*, 2015. DOI: 10.1109/PTC.2015.7232303.

[139] T. Liu, Y. Liu, J. Liu, *et al.*, "A bayesian learning based scheme for online dynamic security assessment and preventive control", *IEEE Transactions on Power Systems*, vol. 35, no. 5, 2020. DOI: 10.1109/TPWRS.2020.2983477.

[140] C. Liu, K. Sun, Z. H. Rather, *et al.*, "A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees", *IEEE Transactions on Power Systems*, vol. 29, no. 2, 2014. DOI: 10.1109/TPWRS.2013.2283064.

[141] T. S. WG, "Realtime Interface System Operator - DER Dutch implementation of RfG interface requirements Technical Specification Document", Netbeheer Nederland, Tech. Rep., Feb. 2024.

[142] R. Bründlinger, G. Arnold, N. Schäfer, T. Schaupp, G. Graditi, and G. Adinolfi, "Implementation of the european network code on requirements for generators on the european national level current status-trends and challenges", in *8th Solar Integration Workshop*, 2018.

[143] H. Bevrani, T. Kato, T. Ise, and K. Inoue, *Grid Connected Converters: Modeling, Stability and Control*. Elsevier, 2022.

[144] J. Matevosyan, B. Badrzadeh, T. Prevost, *et al.*, "Grid-forming inverters: Are they the key for high renewable penetration?", *IEEE Power and Energy Magazine*, vol. 17, Nov. 2019.

[145] L. Rezai, F. Pöschke, M. Andrejewski, H. Schulte, and J. Fortmann, "Evaluating flicker damping capabilities of wind turbine inverters with grid following and grid forming controls applying the proposed iec 61400-21-4", in *21st Wind & Solar Integration Workshop*, IET, 2022.

[146] G. A. Susto, A. Cenedese, and M. Terzi, "Time-series classification methods: Review and applications to power systems data", *Big data application in power systems*, 2018.

[147] G. Rietveld, A. Jongepier, J. van Seters, *et al.*, "Application of pmus for monitoring a 50 kv distribution grid", in *Proceedings of the 23rd International Conference on Electricity Distribution (CIRED)*, 2015.

[148] N. Veerakumar, D. Ćetenović, K. Kongurai, M. Popov, *et al.*, "Pmu-based real-time distribution system state estimation considering anomaly detection, discrimination and identification", *International Journal of Electrical Power & Energy Systems*, vol. 148, 2023.

[149] Y. Li, J. M. Guerrero, J. Yang, Y. Guan, G. Ma, and J. Feng, "Dynamic equivalent modeling for black-box microgrids under multi-operating-point by using lstm", *CSEE Journal of Power and Energy Systems*, vol. 10, no. 2, pp. 639–648, 2022.

[150] B. Choi, *ARMA model identification*. Springer Science & Business Media, 2012.

[151] B. Espinosa Apráez and M. Noorman, "Regulating ai in the 'twin transitions': Significance and shortcomings of the ai act in the digitalised electricity sector", *Review of European, Comparative & International Environmental Law*, vol. 33, no. 3, 2024.

[152] D. Chrysostomou, J. L. R. Torres, and J. L. Cremer, "Exploring operational flexibility of active distribution networks with low observability", *2023 IEEE Belgrade PowerTech*, 2023.

[153] D. Chrysostomou, J. L. R. Torres, and J. L. Cremer, "Tensorconvolutionplus: A python package for distribution system flexibility area estimation", *SoftwareX*, vol. 31, p. 102 241, 2025, ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2025.102241. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711025002080.

[154] M. Troncia, S. Ruggeri, G. G. Soma, *et al.*, "Strategic decision-making support for distribution system planning with flexibility alternatives", *Sustainable Energy, Grids and Networks*, vol. 35, p. 101 138, 2023, ISSN: 2352-4677. DOI: https://doi.org/10.1016/j.segan.2023.101138.

[155] Y. Weng, J. Xie, P. Wang, and H. D. Nguyen, "Asymmetrically reciprocal effects and congestion management in tso-dso coordination through feasibility regularizer", *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1948–1962, 2023. DOI: 10.1109/TPWRS.2022.3193052.

[156] H. Bakhtiari, M. R. Hesamzadeh, and D. W. Bunn, "Tso-dso operational coordination using a look-ahead multi-interval framework", *IEEE Transactions on Power Systems*, vol. 38, no. 5, pp. 4221–4239, 2023. DOI: 10.1109/TPWRS.2022.3219581.

[157] N. Majumdar, P. Kengkat, R. Yermekbayev, and L. Hofmann, "Reliability parameterised distribution grid flexibility aggregation considering renewable uncertainties", in *2023 58th International Universities Power Engineering Conference (UPEC)*, 2023, pp. 1–6. DOI: 10.1109/UPEC57427.2023.10294524.

[158] H. Hui, M. Bao, Y. Ding, J. Yan, and Y. Song, "Probabilistic integrated flexible regions of multi-energy industrial parks: Conceptualization and characterization", *Applied Energy*, vol. 349, p. 121 521, 2023.

[159] M. Kalantar-Neyestanaki and R. Cherkaoui, "Grid-cognizant tso and dso coordination framework for active and reactive power flexibility exchange: The swiss case study", *Electric Power Systems Research*, vol. 235, p. 110 747, 2024, ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2024.110747.

[160] U. Cali, M. Kuzlu, M. Pipattanasomporn, J. Kempf, and L. Bai, "Introduction to the digitalization of power systems and markets", *Digitalization of Power Markets and Systems Using Energy Informatics*, 2021.

[161] R. Fan, T. Yin, K. Yang, J. Lian, and J. Buckheit, "New data-driven approach to bridging power system protection gaps with deep learning", *Electric Power Systems Research*, vol. 208, p. 107 863, 2022.

[162] T. Ahmad, R. Madonski, D. Zhang, C. Huang, and A. Mujeeb, "Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm", *Renewable and Sustainable Energy Reviews*, vol. 160, p. 112 128, 2022.

[163] M. L. Di Silvestre, S. Favuzza, E. Riva Sanseverino, and G. Zizzo, "How decarbonization, digitalization and decentralization are changing key power infrastructures", *Renewable and Sustainable Energy Reviews*, vol. 93, 2018, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2018.05.068.

[164] L. Thurner, A. Scheidler, F. Schäfer, *et al.*, "Pandapower—an open-source python tool for convenient modeling, analysis, and optimization of electric power systems", *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510–6521, 2018.

[165] F. Milano, "An open source power system analysis toolbox", *IEEE Transactions on Power systems*, vol. 20, no. 3, 2005.

[166] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education", *IEEE Transactions on Power Systems*, vol. 26, no. 1, 2011. DOI: 10.1109/TPWRS.2010.2051168.

[167] Y. Xiang, P. Salemink, B. Stoeller, N. Bharambe, and W. van Westering, "Power grid model: A high-performance distribution grid calculation library", in *27th International Conference on Electricity Distribution (CIRED 2023)*, vol. 2023, 2023. DOI: 10.1049/icp.2023.0633.

[168] J. Johnston, R. Henriquez-Auba, B. Maluenda, and M. Fripp, "Switch 2.0: A modern platform for planning high-renewable power systems", *SoftwareX*, vol. 10, 2019, ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2019.100251.

[169] M. Mirz, S. Vogel, G. Reinke, and A. Monti, "Dpsim—a dynamic phasor real-time simulator for power systems", *SoftwareX*, vol. 10, 2019, ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2019.100253.

[170] A. Plietzsch, R. Kogler, S. Auer, *et al.*, "Powerdynamics.jl—an experimentally validated open-source package for the dynamical analysis of power grids", *SoftwareX*, vol. 17, 2022, ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2021.100861.

[171] A. Ramos, E. F. Alvarez, and S. Lumbreras, "Opentepes: Open-source transmission and generation expansion planning", *SoftwareX*, vol. 18, 2022.

[172] G. Brandl, "Sphinx documentation", *URL http://sphinx-doc. org/sphinx. pdf*, 2010.

[173] G. K. Papazoglou, A. A. Forouli, E. A. Bakirtzis, P. N. Biskas, and A. G. Bakirtzis, "Estimating the feasible operating region of active distribution networks using the genetic algorithm", in *2023 IEEE PES GTD International Conference and Exposition (GTD)*, IEEE, 2023, pp. 182–187.

[174] A. Churkin, W. Kong, P. Mancarella, and E. A. M. Ceseña, "Quantifying phase unbalance and coordination impacts on distribution network flexibility", *arXiv preprint arXiv:2408.06516*, 2024.

A

TensorConvolutionPlus: A python package for distribution system flexibility area estimation

Power system operators need new, efficient operational tools to use the flexibility of distributed resources and deal with the challenges of highly uncertain and variable power systems. Transmission system operators can consider the available flexibility in distribution systems (DSs) without breaching the DS constraints through flexibility areas. However, there is an absence of open-source packages for flexibility area estimation. This chapter introduces TensorConvolutionPlus, a user-friendly Python-based package for flexibility area estimation. The main features of TensorConvolutionPlus include estimating flexibility areas using the TensorConvolution+ algorithm, the power flow-based algorithm, an exhaustive PF-based algorithm, and an optimal power flow-based algorithm. Additional features include adapting flexibility area estimations from different operating conditions and including flexibility service providers offering discrete setpoints of flexibility. The TensorConvolutionPlus package facilitates a broader adaptation of flexibility estimation algorithms by system operators and power system researchers.

Parts of this chapter have been published in: D. Chrysostomou, J. L. R. Torres and J. L. Cremer, "TensorConvolutionPlus: A python package for distribution system flexibility area estimation", SoftwareX, 2025. DOI: https://doi.org/10.1016/j.softx.2025.102241, [153].
Parts of this chapter have been published in: D. Chrysostomou, J. L. R. Torres and J. L. Cremer, "TensorConvolutionPlus: A python package for distribution system flexibility area estimation", ArXiv. DOI: https://doi.org/10.48550/arXiv.2501.06976

A

Metadata

Tab. A.1 describes the proposed package's code metadata.

Nr.	Code metadata description	Metadata
C1	Current code version	v0.1.1
C2	Permanent link to code/repository used for this code version	https://github.com/Demetris-Ch/TensorConvolutionFlexibility
C3	Permanent link to Reproducible Capsule	https://codeocean.com/capsule/4767295/tree/v1
C4	Legal Code License	CC-BY
C5	Code versioning system used	git
C6	Software code languages, tools, and services used	Python
C7	Compilation requirements, operating environments & dependencies	Python≥ 3.10, matplotlib≥ 3.8.2, networkx≥ 3.1, numpy≥ 1.24.3, pandapower≥ 2.13.1, pandas≥ 1.5.3, scikit-learn≥ 1.3.0, scipy≥ 1.11.2, seaborn≥ 0.13.2, tntorch≥ 1.1.1, torch≥ 2.0.1, tqdm≥ 4.66.1
C8	Link to developer documentation/manual	https://demetris-ch.github.io/TensorConvolutionFlexibility/
C9	Support email for questions	D.Chrysostomou@tudelft.nl

Table A.1: Code metadata.

A.1. Motivation and significance

Power systems encounter an operational transition as renewable energy sources (RES) penetration rises, and the conventional generation output decreases. This operational transition includes coordinating transmission system operators (TSOs) and distribution system operators (DSOs). RES are mainly connected to distribution systems (DSs) and have high variability and uncertainty, challenging the TSOs and DSOs who need to maintain their system balance. However, RES and active users in DSs can also offer flexibility to contribute to the reduction of these challenges. This flexibility corresponds to the RES or active users changing their generation or consumption setpoints to support the system operators. The RES and active users that offer flexibility constitute the flexibility service providers (FSPs). Therefore, TSOs and DSOs need operational tools that can efficiently allow communicating and using FSP flexibility [81][154].

TSO-DSO coordination approaches can be categorized in TSO-managed, DSO-managed, or TSO-DSO hybrid models [10][155], with recent TSO-DSO coordination approaches also developing multi-interval TSO-DSO coordination [156]. The proposed package is focused on DSO-managed coordination, where the DSO validates the feasibility and aggregates the FSP flexibility to inform the TSO of the available, feasible FSP services. The DSO can validate and aggregate the FSP flexibility using Flexibility areas (FAs) [10]. FAs

are areas in the active (P) and reactive (Q) power plane, illustrating which setpoints TSOs can achieve at a TSO-DSO interconnection node when utilizing feasible flexibility from the DSs.

FA estimation approaches mainly apply power flows (PF) or optimal power flows (OPF) [18]–[23], [25], [26], [30], [82], [152] to explore the limits of the offered flexibility in the PQ space. PF-based algorithms are simple and consistent but slow, whereas OPF-based algorithms are faster but can have convergence issues. A recently proposed FA estimation algorithm, TensorConvolution+ [110], explores the limits and the density of feasible flexibility shift combinations to reach each FA setpoint. TensorConvolution+ applies convolution and tensor operations to combine flexibility shifts and evaluate their feasibility for the system's technical constraints. Additional functionalities of the TensorConvolution+ approach include storing tensors from prior estimations and adapting FAs for altered operating conditions (OCs). Forecast errors in distribution systems can provide uncertainty in day ahead FA estimation, with recent approaches exploring chance constraint [30], robust optimization [101][18], or probabilistic [20][157][102][158] algorithms. The proposed package implements deterministic FA estimation. Fast FA estimation algorithms can potentially be used in close to real-time, reducing the impact of forecast errors [22]. Existing FA estimation approaches include case studies with real-world systems and data. For example [38] used data from the French DSO whereas [159] used the Swiss TS. The proposed software currently supports systems in the pandapower format. Thus, users should import their systems in the pandapower format, or use the existing pandapower systems.

The operational transition and data availability in power systems provided opportunities for digitalizing power systems. This digitalization corresponds to more intelligent, effective, green power grid operations [160]. Data-driven approaches for power system operations are emerging with works on dynamic [161] and static simulations [162][28], from protection gaps [161], to dynamic security assessment [123] and probabilistic approaches [162]. The main drivers for change in power systems are decarbonization, digitalization, and decentralization, with flexibility as a key for decarbonization [163]. The digitalization of power systems resulted in the emergence of open-source tools. Power systems open-source tools include the PandaPower [164] in Python, PSAT [165] in Matlab and GNU/Octave, MatPower [166] in Matlab. More recent tools with increased efficiency include [167] in C++. As highlighted by [164], software developed in languages with open-source licenses, such as Python, C++, and Julia, can be used as stand-alone or extended with other libraries. These advantages of open-source libraries drove researchers to design more specialized power system-related packages such as [168]–[171]. FA estimation is an emerging field in power engineering that can improve the power system stability and utilization of flexibility from decentralized resources. However, currently, there are no open-source FA estimation packages. An open-source package for FA estimation can accelerate the adoption of FAs by power system operators and attract more researchers to this emerging field.

The developed Python-based package for FA estimation focuses on the TensorConvolution+ algorithm [110] but also includes a traditional PF-based algorithm, an exhaustive PF-based algorithm, and an OPF-based algorithm.

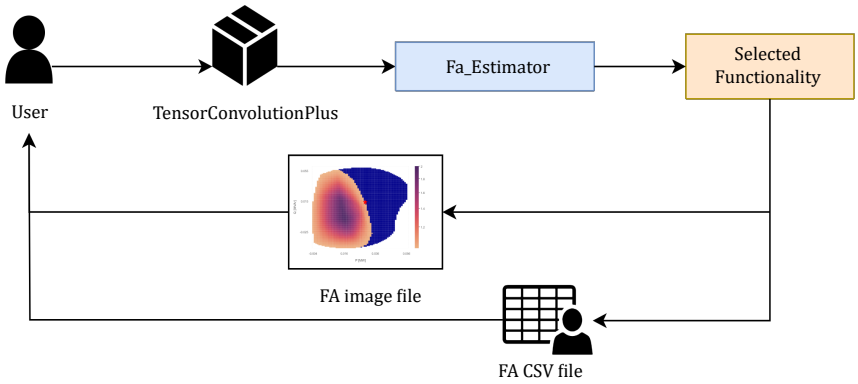


Figure A.1: TensorConvolutionPlus package usage through the script ( ) FA_Estimator and its main functionalities ( ).

A.2. Software description

The software framework is implemented in Python. The package can be installed from the Python package index (PyPi). The code implementation is available on GitHub. The documentation for the package's main functions, classes, supporting functions, and case studies is available online and was built using the Sphinx library [172].
The seven main software functionalities are two PF-based algorithms, one OPF-based algorithm, and four versions of the TensorConvolution+ algorithm. Fig.A.1 illustrates the usage of the package functionalities. The user calls the *FA_Estimator* script of the TensorConvolutionPlus package and selects one of the main functionalities. The selected algorithm functionality estimates the FA and stores locally:

1. The FA image in a portable document format (PDF) file.
2. The FA results in a comma-separated values (CSV) file.
3. A text file with the simulation information on duration and algorithm-specific details.

The *tcp_plus_save_tensors* also includes additional files from the FA results. The user inputs depend on the functionality.

A.2.1. Software architecture

The proposed software architecture intends to allow efficient modification and expansion of specific sub-processes of the FA estimation problem. Tab.A.2 highlights the roles of the Python scripts implementing the package functionalities. The *json_reader* script checks if each input is within the acceptable options to avoid erroneous results. If *json_reader* detects an unacceptable input, the FA estimation does not begin, and the user is informed about the input causing the issue and the acceptable values. Users can identify which scripts and functionalities to modify or expand to fulfill additional needs. For example, to add or modify the plotting functions of the package, one would modify the *plotting* script. For more complex modifications, such as adding different sampling

techniques for the PF-based algorithms, the *data_sampler* and *json_reader* would be the only scripts requiring modification. The user could modify the *sample_from_rng* function in *data_sampler* to sample FSP shifts using a new distribution when the new distribution keyword is selected. The *json_reader* modification corresponds to adding new acceptable options, the new sampling distribution keyword in *self.distribution* in the function *tester* of the class *SettingReader*. An input outside the acceptable options in the *json_reader* stops the process before the simulation to avoid erroneous or untested results. This architecture also allows potential future expansions to new FA estimation algorithms, where an additional script can be created and integrated with the *json_reader* and *FA_Estimator* scripts without impacting other processes. For new FA estimation algorithms, the *json_reader* would need to create any new variables for the algorithms in the *__init__* of the class *SettingReader*, and add the acceptable options for each variable in the *tester* function. The new algorithms should be created as a new functionality in the *FA_Estimator* script. Depending on the needs of the new algorithms, the user can call sampling functions from the *data_sampler*, plotting functions from the *plotting* script, or other generic functions from the *utils* script.

Script	**Role**
FA_Estimator	Package main script which includes the main functionalities.
json_reader	(i) Read input settings and create a SettingReader object with the algorithm parameters. (ii) Validate that the inputs are within the acceptable options.
data_sampler	Sample flexibility shifts from flexibility providers.
scenario_setup	Update network and SettingReader object based on the algorithm input parameters.
opf	Perform the OPF-based FA estimation algorithm.
monte_carlo	Perform the PF-based FA estimation algorithms.
conv_simulations	Perform the TensorConvolution+ algorithm functionalities.
utils	Provide generic functions to the other scripts.
plotting	Generate figures of resulting FAs.

Table A.2: Package script roles.

The package's GitHub repository includes the Python scripts under the "src/TensorConvolutionPlus" directory. The package dependencies include *pandapower* to perform PF and OPF operations, *PyTorch* for tensor operations, *SciPy* for convolution operations, *tntorch* and *scikit-learn* for additional TensorConvolution+ subprocesses. *NumPy* and *pandas* are used for data storage, processing, and sampling, *tqdm* illustrates the FA estimation progress, and *matplotlib* and *seaborn* generate the figures.

A.2.2. SOFTWARE FUNCTIONALITIES

The *FA_Estimator* script includes the main functionalities as in Fig.A.2. The *monte_carlo_pf* and *exhaustive_pf* functions apply PF-based FA estimation algorithms. The *opf* function applies the OPF-based FA estimation. The *tc_plus, tc_plus_merge, tc_plus_save_tensors,*

A

tc_plus_adapt functions perform different versions of the TensorConvolution+ algorithm. The common inputs for all main functionalities are the network pandapower object (*net*), the network name (*net_name*), indices of load FSPs (*fsp_load_indices*), indices of distributed generation FSPs (*fso_dg_indices*), scenario type for initial topology and OCs (*scenario_type*), and system constraints for maximum component loading [%] (*max_curr_per*), maximum voltage [p.u.] (*max_volt_pu*), and minimum voltage [p.u.] (*min_volt_pu*). All functionality inputs are optional. However, to estimate FAs, at least one distributed generation or load FSP is required. The remaining scripts, at the right of Fig.A.2, provide functions and sub-processes to implement the main functionalities.

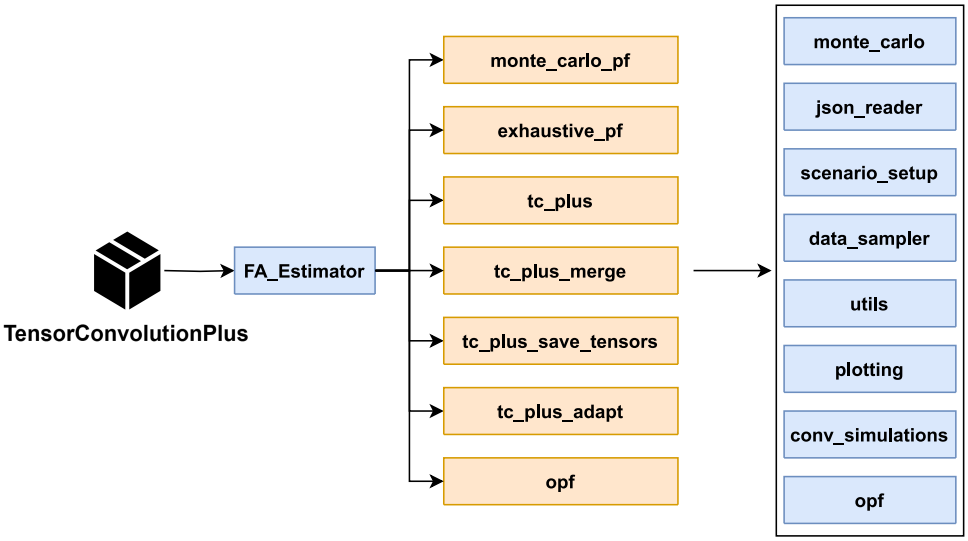


Figure A.2: Package main functions ( ) relationship (⟶) with python scripts ( ).

A.3. Implementation and empirical results

The main building blocks for the implemented FA estimation algorithms are (i) initializing network and FA estimation settings, (ii) performing simulations for FSP flexibility shifts on the network, (iii) processing the simulation results, and (iv) plotting and storing the simulation results. All functions have similar block (i), the Alg.3. The plotting functions differ between the functionalities.

A.3.1. PF-based functions

The PF-based functions differ from the FSP flexibility shift sampling functions. Therefore, the main difference between the PF-based functions is:

- *monte_carlo_pf* uses a probability *distribution* (input) to obtain *no_samples* (input) of flexibility shift combinations.
- *exhaustive_pf* uses the increments *dp, dq* (inputs) for P and Q to sample all possible discretized flexibility shift combinations.

Algorithm 3 Initialize network and FA estimation settings.

Require: *fsp_load_indices* and/or *fsp_dg_indices*,
initialize *SettingReader* object with estimation settings,
check if SettingReader has acceptable values,
if *net* is None **then**,
net ← pandapower network with name=*net_name*,
change *net* topology and OC for *scenario_type*,
end if
return *SettingReader* with *net*.

Alg.4 illustrates the algorithm for both the PF-based functions after the samples are obtained. The *samples* array includes the FSP shift combination samples. PCC is the point of common coupling between the TSO and the DSO.

Algorithm 4 PF-based FA estimation.

Require: *samples*, *SettingReader*, *net*,
init_net ← *net*,
for *sample* ∈ *samples* **do**,
apply *sample* on *net*,
run PF on *net*,
if *net* OC are within system constraints **then**,
store *sample* index and PCC P, Q as feasible,
else
store *sample* index and PCC P, Q as not-feasible,
end if
net ← *init_net*,
end for
store FA PDF, CSV, and text file.

EMPIRICAL RESULTS

The PF-based functions can illustrate consistent performance under various network structures and FSP combinations. However, as FSPs increase, the performance deteriorates. The Monte-Carlo-based algorithm could require large *no_samples* to capture the margins of the flexibility area. The exhaustive PF-based algorithm can become intractable for more than 3 FSPs and small *dp, dq*.

A.3.2. OPF-BASED FUNCTION

The OPF-based algorithm applies four multi-objective optimizations (MOO). These optimizations aim to identify the maximum feasible active (P_{PCC}) and reactive power (Q_{PCC}) at the PCC achieved using the available flexibility as:

1. $\max(\alpha P_{PCC} + (1-\alpha)Q_{PCC})$,
2. $\max(-\alpha P_{PCC} + (1-\alpha)Q_{PCC})$,

3. $\max(\alpha P_{PCC} + (\alpha - 1)Q_{PCC})$,

4. $\max(-\alpha P_{PCC} + (\alpha - 1)Q_{PCC})$.

The variable $\alpha \in [0,1]$ provides a plane in which the active and reactive power shifts are combined. Therefore, the algorithm iteratively changes α in steps provided through the additional input the *opf_step*. For example, an *opf_step*= 0.1 results in 11 iterations per MOO, thus 44 OPFs to estimate the FA.

Empirical results

The OPF-based function has convergence issues for different network structures. The OPF-based function can converge for the radial CIGRE MV network when ignoring transformer loading limitations but might not converge in other networks, e.g., the Oberrhein network. These issues are due to the OPFs performed for each MOO failing to converge to identify an optimal solution within the constraints and network settings.

A.3.3. TensorConvolution+ functions

The TensorConvolution+ functions correspond to the algorithm proposed in [110]. The function *tc_plus* corresponds to the generic approach of the algorithm, whereas the rest accommodate specific use cases.
TensorConvolution+ initially creates samples of all flexibility shifts for each FSP with increments *dp, dq* (inputs) for active and reactive power, respectively. The samples do not include combinations of FSPs. Thus, the number of samples increases linearly with the addition of FSPs. The *flex_shape* input characterizes the boundaries of each FSP flexibility. Currently, the FSP shapes can be:

1. *Smax*: The FSP output apparent power cannot exceed its maximum apparent power, resulting in a semi-oval flexibility shape.

2. *PQmax*: The FSP active and reactive power outputs cannot exceed the maximum apparent power, resulting in a rectangular flexibility shape.

Using these samples and the outputs of Alg.3, the function *tc_plus* performs Alg.5 to estimate and plot the FA. The TensorConvolution+ algorithm applies convolutions to combine the flexibility shifts between FSPs. However, as FAs consider the network constraint limitations, the FA algorithm needs to first filter the feasible and non-feasible combinations. Thus, for sensitive network components (close to network constraints), TensorConvolution+ modifies the convolution operation to avoid the summation step and stores the resulting impacts from combined FSP shifts to tensors. The first two dimensions of the tensors correspond to the PCC active and reactive powers, whereas the additional dimensions include the impact of a specific flexibility combination on a sensitive network component. Following these tensor operations, TensorConvolution+ filters out the FSP combinations resulting in non-feasible conditions and sums all but the first two tensor dimensions, as the convolution summation step.[110] provides further illustrative and mathematical analysis of TensorConvolution+. The main differences between *tc_plus*, *tc_plus_merge*, *tc_plus_save_tensors*, and *tc_plus_adapt*:

Algorithm 5 *tc_plus* FA estimation.

Require: *samples*, *SettingReader*, *net*,
init_net ← *net*,
Ω^{γ} ← set of *net* components,
for *sample* ∈ *samples* **do**,
run PF on *sample* and record impact on network components,
end for
Ω_{sm}^{FSPs} ← set of FSPs with capacity smaller than dp, dq,
impacts ← the FSP impacts on each $\gamma \in \Omega^{\gamma}$,
uFA ← the unconstrained FA using convolutions on all $FSP \in \Omega^{FSP}$,
Ω_{γ}^{FSP} ← set of FSPs that impact γ more than the sensitivity thresholds,
Ω^{γ} ← remove all γ from Ω^{γ} that cannot reach the system constraints from the maximum FSP impacts,
for $\gamma \in \Omega^{\gamma}$ **do**,
Ξ_{γ} ← apply tensor-convolution for all feasible Ω_{γ}^{FSP} combinations,
A_{γ} ← sum Ξ_{γ} in all dimensions except the first 2,
Υ_{γ} ← apply convolution between A_{γ} and the $\Omega^{FSP} \setminus \Omega_{\gamma}^{FSP}$ shifts,
end for
if $\Omega^{\gamma} = \emptyset$ **then**,
$FA \leftarrow uFA$,
else
FA ← the element-wise minimum between $\Upsilon_{\gamma} \forall \gamma \in \Omega^{\gamma}$,
end if
normalize FA,
get axes of result and create a result data frame,
if $\Omega_{sm}^{FSPs} \neq \emptyset$ **then**,
FA_i ← bilinear interpolation on FA to increase its size,
FA_{sm} ← convolute FA_i with the FSPs in Ω_{sm}^{FSPs},
normalize FA_{sm},
end if
store FA PDF, CSV, and text file.

- *tc_plus_save_tensors* stores extracted information and sensitivity tensors locally. This functionality reduces the tensors' memory requirements using tensor train decomposition (TTD). This reduction causes delays and is therefore excluded from the *tc_plus* function.

- *tc_plus_adapt* does not sample flexibility shifts nor estimates network component sensitives, as it adapts from the FA estimated in previous simulations for the same FSP offers.

- *tc_plus_merge* is useful when memory limitations do not allow estimating FAs with the *tc_plus* function. The *tc_plus_merge* function estimates the electrical distance between all FSPs. When a network component is sensitive to more than *max_FSPs* (input), this function merges the flexibility between the two electrically closest

A

components iteratively until the network component is sensitive to *max_FSPs*.

Empirical results

The TensorConvolution+ functions perform computationally better in GPUs, where tensor operations can be faster. Simulations in different network topologies showcased consistent performance with TensorConvolution+.
TensorConvolution+ can have memory issues and terminate the simulation for networks with multiple components close to the system constraints, small *dp, dq*, and increased FSPs. GPUs with higher VRAM reduce these limitations. When memory issues persist, *tc_plus_merge* can mitigate these issues but could reduce the estimation accuracy.

A.4. Illustrative examples

All examples were performed using the A100 GPU in Google Colab [97]. To use the package, the user can perform two steps. The first step is installing the package through pip as:

```
pip install TensorConvolutionPlus
```

The second step is importing the package's *FA_Estimator* in a Python script as:

```
from TensorConvolutionPlus import FA_Estimator as TCP
```

The user can use any main function from Fig.A.2 using the imported *TCP*. The following subsections showcase the main functions of the package after the above steps.

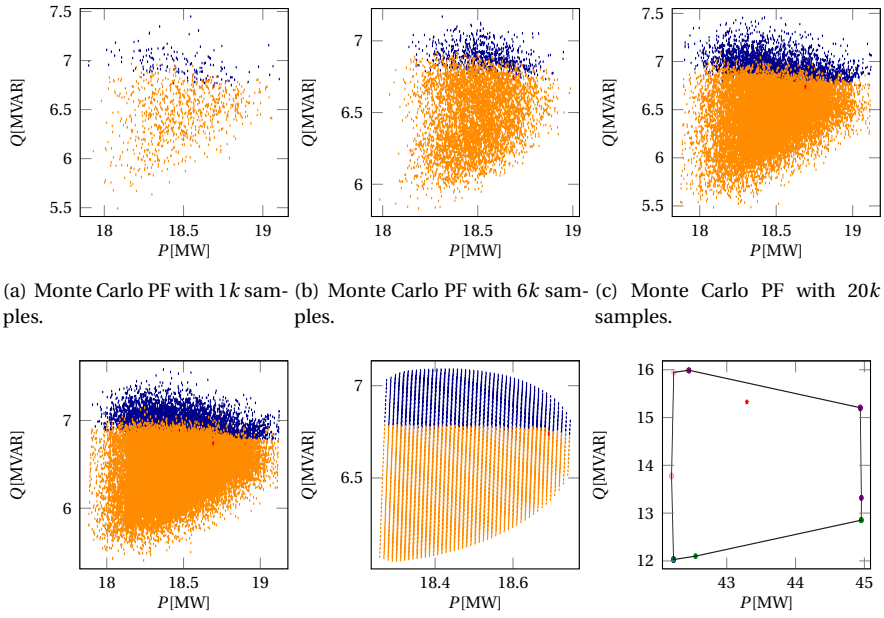
A.4.1. PF and OPF Functionalities

This section includes examples using the Monte Carlo PF, exhaustive PF, and OPF functionalities. These examples used the Python script code:

```
TCP.monte_carlo_pf(net_name='MV Oberrhein0', no_samples=1000,
    fsp_load_indices=[1, 2, 3], fsp_dg_indices=[1, 2, 3], distribution=
    'Uniform')

TCP.monte_carlo_pf(net_name='MV Oberrhein0', no_samples=6000,
    fsp_load_indices=[1, 2, 3], fsp_dg_indices=[1, 2, 3], distribution=
    'Uniform')

TCP.monte_carlo_pf(net_name='MV Oberrhein0', no_samples=20000,
    fsp_load_indices=[1, 2, 3], fsp_dg_indices=[1, 2, 3], distribution=
    'Uniform')

TCP.monte_carlo_pf(net_name='MV Oberrhein0', no_samples=40000,
    fsp_load_indices=[1, 2, 3], fsp_dg_indices=[1, 2, 3], distribution=
    'Uniform')

TCP.exhaustive_pf(net_name='MV Oberrhein0', dp=0.01, dq=0.02,
    fsp_load_indices=[5], fsp_dg_indices=[5])

TCP.opf(net_name='CIGRE MV', opf_step=0.1, fsp_load_indices=[1, 4, 9],
    fsp_dg_indices=[8])
```

Fig.A.3 illustrates the expected output FA for each line respectively. In terms of computational burden, the simulations required 55 s for Fig.A.3(a), 5 minutes and 5 s for

Fig.A.3(b), 18 minutes for Fig.A.3(c), 36 minutes for Fig.A.3(d), 36 minutes and 18s for Fig.A.3(e), and 33.7s for Fig.A.3(f). Fig.A.3(e) performed 43121 power flows and Fig.A.3(f) executed 44 OPFs. The Monte Carlo-based results showcase clearer FA margins for 20K and 40k samples than lower values. However, the number of samples for clearer margins can also depend on the number of FSPs. Monte Carlo-based functions can be better than the exhaustive PF-based function in exploring FA margins for scenarios with more FSPs. Lowering the resolution for the exhaustive approach for producing clear FA margins can be intractable as FSPs increase.



(a) Monte Carlo PF with 1k samples.

(b) Monte Carlo PF with 6k samples.

(c) Monte Carlo PF with 20k samples.

(d) Monte Carlo PF with 40k samples.

(e) Exhaustive PF

(f) OPF

Figure A.3: PF-based and OPF-based FA estimations.

A.4.2. TENSORCONVOLUTION+

This section illustrates examples using the TensorConvolution+ FA estimation functionality, using the Python lines:

```
TCP.tc_plus(net_name='MV Oberrhein0', fsp_load_indices=[1, 2, 3], dp
    =0.05, dq=0.1, fsp_dg_indices=[1, 2, 3])

TCP.tc_plus(net_name='MV Oberrhein0', fsp_load_indices=[1, 2, 3], dp
    =0.075, dq=0.15, fsp_dg_indices=[1, 2, 3])

TCP.tc_plus_merge(net_name='MV Oberrhein0', fsp_load_indices=[1, 2, 3],
     dp=0.025, dq=0.05, fsp_dg_indices=[1, 2, 3], max_fsps=5)

```

```
TCP.tc_plus(net_name='MV Oberrhein0', fsp_load_indices=[1, 2], dp=0.05,
    dq=0.1, fsp_dg_indices=[1, 2, 3])

TCP.tc_plus(net_name='MV Oberrhein0', fsp_load_indices=[1, 2], dp
    =0.025, dq=0.05, fsp_dg_indices=[1, 2, 3])

TCP.tc_plus(net_name='CIGRE MV', fsp_load_indices=[3, 4, 5], dp=0.05,
    dq=0.1, fsp_dg_indices=[8], non_linear_fsps=[8])
```

Fig.A.4 shows the expected output FAs from the above lines respectively. Fig.A.4(a) required 13.1s whereas the second line reduces the resolution, with Fig.A.4(b) requiring 7.6s. Increasing the resolution from the first line to $dp = 0.025, dq = 0.5$ exceeded the A100 GPU memory, stoping the simulation. Thus, running the *tc_plus_merge* function merged FSPs when more than 5 FSPs impacted a network component and estimated the FA of Fig.A.4(c) for this higher resolution in 37.9s. The fourth line reduced the number of FSPs to 5 from Fig.A.4(a) with the same resolution, resulting in Fig.A.4(d) in 9s. Increasing the resolution of Fig.A.4(d) to $dp = 0.025, dq = 0.5$ did not cause memory issues for the GPU, resulting in Fig.A.4(e) in 26.5s. The last line includes an FSP offering discrete setpoints of flexibility. The input *non_linear_fsps* specifies which of the FSPs referenced in the *fsp_dg_indices* can only offer 2 setpoints; current output or full output reduction. The duration for Fig.A.4(f) was 17.8s.

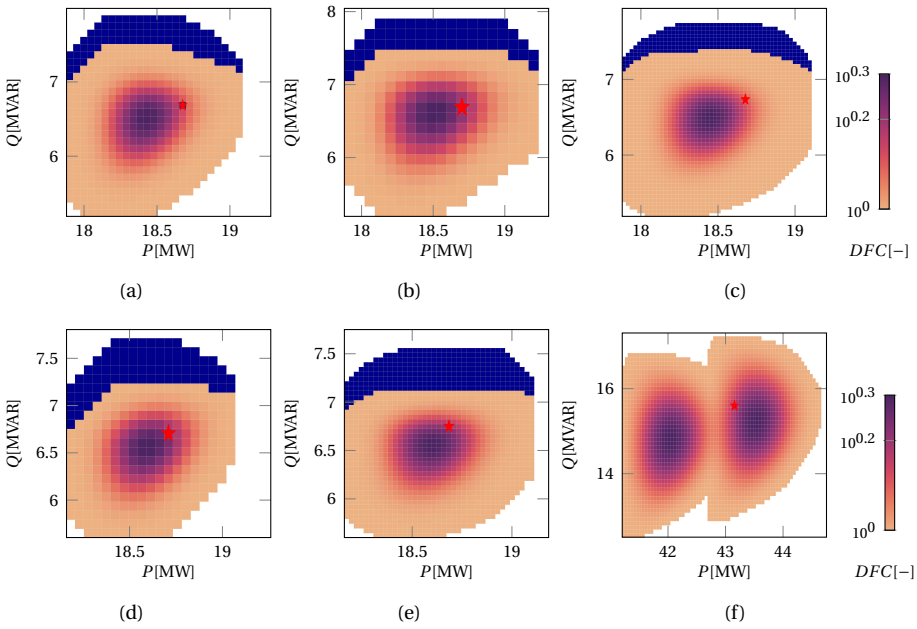


Figure A.4: TensorConvolution+ algorithm examples with all linear FSPs in (a)-(e), merging of FSPs in (c) and one non-linear FSP in (f).

A.4.3. TENSORCONVOLUTION+ ADAPT

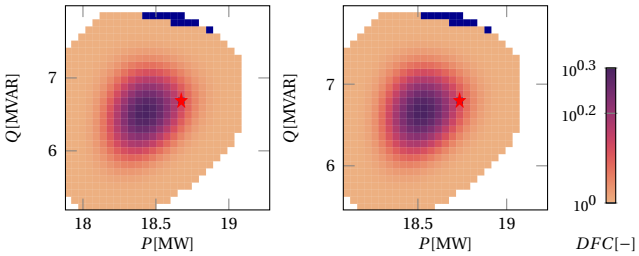
To adapt FA estimations using estimations from expected or prior operating conditions, TensorConvolution+ requires storing the relevant information from these prior FA estimations locally. Therefore, the package's user should first call the *tc_plus_save_tensors* function to store the information. This function performs TTD to reduce the space required to store tensors but requires more extensive computational time than *tc_plus* to execute the additional TTD and storing operations. After *tc_plus_save_tensors* is executed, then *tc_plus_adapt* can use the stored information to estimate FAs for altered related OCs if the network topology and FSPs are consistent. Bellow, an example script storing the information, altering the operating conditions, and adapting the FA for the new operating conditions:

```
# Step1: Define the consistent FSPs for the storing and adapting
    functions
fsp_load_indices = [1, 2, 3]
fsp_dg_indices = [1, 2, 3]
# Step 2: Estimate the FA and store the relevant information for
    adaptation
TCP.tc_plus_save_tensors(net_name='MV Oberrhein0', fsp_load_indices=
    fsp_load_indices, dp=0.05, dq=0.1, fsp_dg_indices=fsp_dg_indices)
# Step 3: Modify the network operating conditions
net, net_tmp = pn.mv_oberrhein(separation_by_sub=True)
net.load['sn_mva'] = list(net.load['p_mw'].pow(2).add(net.load['q_mvar'
    ].pow(2)).pow(0.5))
net.load['scaling'] = [1 for i in range(len(net.load))]
net.sgen['scaling'] = [1 for i in range(len(net.sgen))]
net.switch['closed'] = [True for i in range(len(net.switch))]
# Step 4: Fix the network structure
net = fix_net(net) # This function is included in the documentation (C8
     of Tab.1)
# Step 5: Sample a new operating condition with randomness
rng = np.random.RandomState(212)
net, rng = rand_resample(net, fsp_load_indices, fsp_dg_indices, rng,
    0.05, 0.01, 0.05, 0.01) # This function is included in the
    documentation (C8 of Tab.1)
# Step 6: Adapt the FA using the locally stored information
TCP.tc_plus_adapt(net=net, fsp_load_indices=fsp_load_indices,
    fsp_dg_indices=fsp_dg_indices)
# Step 7: Estimate the FA without adapting to compare with the above-
    adapted result
TCP.tc_plus(net=net, fsp_load_indices=fsp_load_indices, fsp_dg_indices=
    fsp_dg_indices, dp=0.05, dq=0.1)
```

The expected output FA from the storing function is the same as in Fig.A.4(a). However, this function also stores:

1. TTD results for 20 network components with total size $241MB$.
2. FA axes values with total size $2KB$.
3. Matrix of the unconstrained convolution results with total size $4KB$.
4. Dictionary with FSP impacts $382KB$.
5. Dictionary of impactful FSPs per network component $4KB$.

The storing function required 61s. Fig.A.5 illustrates the expected output for the adapted FA and the FA without adaptation, i.e., not using the stored information. The FAs of Fig.A.5 have a high resemblance. The GPU needed 1.4s for the adapted FA of Fig.A.5(a) and 10.4s for the FA of Fig.A.5(b).



(a) Adapted FA. (b) FA witouth adaptation.

Figure A.5: TensorConvolution+ example in adapted FA estimation.

A.5. Impact

TensorConvolutionPlus is the first open-source package for FA estimation. Users can execute the package functionalities with ease. The package can estimate FAs only with two lines of Python code, importing the package and calling the selected functionality. Existing power systems-related libraries do not offer FA estimation functionalities, but rather power flows and optimal power flows. Other FA frameworks have been proposed to make use of such power flows. However, to the best of the authors knowledge, this proposed package is the first that focuses directly on the FA estimation algorithms and is ready to use without requiring the implementation of FA algorithms. Existing works have released code implementations for FA estimation [152][173][174], but no dedicated FA package has been available. This package release enhances FA estimation accessibility, reusability, and ease of integration, providing a standardized and user-friendly solution. Nevertheless, the structure also allows using networks developed by the user. The software design strengthens the potential for further expansion, improvement, and adoption of FA estimation methods. The package structure enables users to expand FA estimation subprocesses without modifying other subprocesses. For example, possible expansions for the PF-based functions include new shapes from the flexibility resources or new sampling distributions for the Monte Carlo-based function. These expansions should not impact Alg.4 but only the scripts *data_sampler* for the new sampling shapes or distributions, *json_reader* to accept the new input keywords, and *FA_Estimator* to obtain the new inputs and call the new *data_sampler* functions.

The developed package includes different FA estimation approaches, allowing users to select and identify the best-performing approach for their tasks. Nevertheless, this package focuses on the TensorConvolution+ [110] algorithm, which can require more complex implementation compared to OPF-based and PF-based algorithms. Through this package, researchers will be able to familiarize themselves with the FA estimation topic

and the TensorConvolution+ algorithm and further advance the field of FA estimation. Similarly, power system operators can use this package directly for their networks and case studies, improving the potential of adopting FAs in their operations. Recent publications explore the application of FA for improved TSO-DSO coordination frameworks and more efficient real-time grid operation [159][10]. The proposed package can further facilitate these applications by alleviating the burden of reimplementing the FA estimation algorithms. Existing power system workflows in Python can integrate with the proposed package if the systems are in a pandapower format. Power system workflows in other languages or software such as DigSilent or Matlab would first require integration with Python before using TensorConvolutionPlus.

A.5.1. LIMITATIONS

The main limitation of the released package is the absence of unit testing for future package versions. Individual limitations from the main functionalities include the GPU memory overflow for TensorConvolution+ and convergence failures for the OPF-based functionality. For TensorConvolution+, the memory overflow can be reduced by using the *tc_plus_merge* functionality. Another mitigation can extend the increments dp, dq, resulting in lower resolution FAs. For the OPF-based convergence failures, simplified OPF implementations might mitigate the issues, but were not tested in the present package release.

A.6. CONCLUSIONS

Power system digitalization and developing open-source power system specialized tools are significant for intelligent and effective power grid operations. In the absence of open-source tools for FA estimation, the developed package can improve the reachability and adoption of TensorConvolution+ and FA estimation algorithms in academia and industry. With a user-friendly structure, the package allows straightforward installation and execution of FA estimation.

The package documentation showcases example usages and details on the scripts and their functions. The package structure diversifies between FA sub-processes. This diversification allows users to better understand and expand the FA estimation algorithms in specific sub-processes, such as the FSP shapes, without impacting the remaining sub-processes. This research is part of the MegaMind project which involves Dutch system operators and grid companies. The package will be pursued for integration with relevant industrial partners in future developments. Unit testing implementation will also be pursued in future work to facilitate consistency in later package versions. Future package expansions include implementing FA functionalities using alternative power system tools and libraries, and comparing the performance to the present functionalities. This package will be maintained with updates reviewed annually. Pull requests are welcome and will be tested and reviewed by the authors before merging. Improvement on the GPU requirements of TensorConvolution+ and the OPF implementations will also be pursued in future developments.

B

Proof: Exploring all Feasible Combinations

The authors use a proof through induction to show that the convolution between any k FSPs will result in the total number of shift combinations between these k FSPs leading to the point $\Delta p^o, \Delta q^o$. Let's assume 2 FSPs, $\hat{\imath}$ and $\tilde{\imath}$. Due to (3.15), the inner part of (3.14) can be described as:

$$F_{\tilde{\imath}}(\Delta p, \Delta q) F_{\hat{\imath}}(\Delta p^o - \Delta p, \Delta q^o - \Delta q) = \begin{cases} 1, & \text{if } [\Delta p^o - \Delta p, \Delta q^o - \Delta q]^T \epsilon \Omega^S_{\hat{\imath}}, [\Delta p, \Delta q]^T, \epsilon \Omega^S_{\tilde{\imath}}, \\ 0, & \text{otherwise.} \end{cases} \tag{B.1}$$

Therefore, each counted point is from shifts existing in the FSP $\hat{\imath}$ and $\tilde{\imath}$ capabilities. From (B.1), all points of $\Delta p, \Delta q$ that are counted, result in $\Delta p_{\hat{\imath}} + \Delta p_{\tilde{\imath}} = \Delta p^o$, $\Delta q_{\hat{\imath}} + \Delta q_{\tilde{\imath}} = \Delta q^o$. Thus, counted shifts only lead to $\Delta p^o, \Delta q^o$. As the summations in convolution are $\sum_{\Delta p=-\infty}^{\infty} \sum_{\Delta q=-\infty}^{\infty}$, all possible combinations of $\Delta p, \Delta q$ are explored. Hence, the convolution counts all available shifts from $\hat{\imath}, \tilde{\imath} \in \Omega^{FSP}$, that lead to the OP $p^o + \Delta p^o, q^o + \Delta q^o$. Thus, the result of (B.1) for all s^o, for $\mathcal{K} = \hat{\imath}, \tilde{\imath}$ can be expressed through:

$$\tilde{F}_{\mathcal{K}}(s^o) = \begin{cases} m_{\mathcal{K}}(s^o), & \text{if } s^o \in \Omega^S_{\mathcal{K}}, \\ 0, & \text{otherwise,} \end{cases} \tag{B.2}$$

$$\Omega^S_{\mathcal{K}} = \{s_{\mathcal{K}} | s_{\mathcal{K}} = \sum_{s \in \pi} s, \forall \pi \in \Omega^S_{\hat{\imath}} \times \Omega^S_{\tilde{\imath}}\}. \tag{B.3}$$

$m_{\mathcal{K}}(s^o) \in \mathbf{N}$ is the number of combinations of $\mathcal{K}$ for s^o.

Parts of this chapter have been published in: D. Chrysostomou, J. L. R. Torres and J. L. Cremer, "Tensor Convolution-Based Aggregated Flexibility Estimation in Active Distribution Systems," in IEEE Transactions on Smart Grid, vol. 16, no. 1, pp. 87-100, Jan. 2025, doi: 10.1109/TSG.2024.3453667 [36].

B

Extending to more FSPs, let the convolution of $\mathcal{K} = 1,..,k$ FSPs be described by an extended version of (B.2), the relationship (3.16). Then, the convolution of $\tilde{F}_{\mathcal{K}}$ and F_{k+1} is:

$$(\tilde{F}_{\mathcal{K}} * F_{k+1})(\Delta p^o, \Delta q^o) = \sum_{\Delta p=-\infty}^{\infty} \sum_{\Delta q=-\infty}^{\infty} F_{k+1}(\Delta p, \Delta q) F_{\mathcal{K}}(\Delta p^o - \Delta p, \Delta q^o - \Delta q). \tag{B.4}$$

The inner part of (B.4), expressed through (B.2) and (3.15) is:

$$F_{k+1}(\Delta p, \Delta q) F_{\mathcal{K}}(\Delta p^o - \Delta p, \Delta q^o - \Delta q) = \begin{cases} m_{\mathcal{K}}([\Delta p, \Delta q]^T), & \text{if } [\Delta p^o - \Delta p, \Delta q^o - \Delta q]^T \\ & \in \Omega^S_{k+1}, [\Delta p, \Delta q]^T \in \Omega^S_{\mathcal{K}}, \\ 0, & \text{otherwise.} \end{cases} \tag{B.5}$$

Each non-zero point is obtained from combinations within the FSP capabilities. From $\sum_{\Delta p=-\infty}^{\infty} \sum_{\Delta q=-\infty}^{\infty}$, all possible combinations from the FSPs are explored. Therefore, if any combination reaching $\Delta p^o, \Delta q^o$ exists, it will be counted. Furthermore, the summation of combinations from a subset of the FSPs ($m_{\mathcal{K}}([\Delta p, \Delta q]^T)$), will ensure that if $\Delta p, \Delta q$ can be used to reach $s^o = [\Delta p^o, \Delta q^o]^T$, then the total number of feasible combinations is accounted. Therefore, the convolution result of $\mathcal{K}+1 = 1, ..., k+1$ FSPs is described by:

$$\tilde{F}_{\mathcal{K}+1}(s^o) = \begin{cases} m_{\mathcal{K}+1}(s^o), & \text{if } s^o \in \Omega^S_{\mathcal{K}+1}, \\ 0, & \text{otherwise,} \end{cases} \tag{B.6}$$

$$\Omega^C_{\mathcal{K}+1} = \Omega^S_1 \times \Omega^S_2 \times ... \times \Omega^S_{k+1},$$
$$\Omega^S_{\mathcal{K}+1} = \{s_{\mathcal{K}+1} | s_{\mathcal{K}+1} = \sum_{s \in \pi} s, \forall \pi \in \Omega^C_{\mathcal{K}+1}\} \subset \Omega^{S^o}. \tag{B.7}$$

The above result is similar to (B.2), (3.16). Therefore, the convolution between any number of FSPs will result in all the possible shift combinations for each reachable shift.

C

BNN Comparative Analysis

The comparative analysis included 1 training, 1 validation, and 3 test sets for each of the 4 networks in Fig.3.10. All datasets followed residential load, PV, and wind turbine (where available) hourly profiles and assumed a 40% increase in generation capacity and 20% in load consumption as [28]. For each hour, the datasets sampled different operating conditions for the load consumption, PV, and wind turbine generations. The CIGRE dataset also included topological changes between all switch combinations in Fig.3.10, and 3 binary features on the status of each switch. The test sets varied in measurement and pseudo-measurement noises (and correlations in CIGRE). The CIGRE network also included an RTI-measured bus (generation with capacity > 1MVA) with noise double from the respective measurement noise. TabC.1 includes the information for each dataset, σ_v^m is the measurement voltage standard deviation. The BNN model settings were:

1. VI: A variational inference model with 4 dense layers, 1 dense variational layer with multivariate normal posterior distribution, with $Tanh(\cdot)$ activation function, and $Z = 100$.

2. MCD: A Monte Carlo dropout model with 4 dense layers, 0.1 dropout, with $ReLU(\cdot)$ activation function, and $Z = 100$.

3. DE: A deep ensemble model with 20 FNNs ($Z = 20$) of 3 dense layers, with $ReLU(\cdot)$ activation function.

All models were trained with a 0.0001 learning rate. VI and MCD were trained for 500 epochs, and each DE FNN was trained for 250 epochs. MCD and DE had 512 hidden units for all layers. For VI, the hidden units were 256, 128, 128, 32 for the dense, 32 for the variational layer in CIGRE, and 256, 128, 128, 64 for the dense, and 64 for the variational layer in OB0, OB1, OB.
For each network, the models were trained once in the corresponding "Tr." dataset of Tab.C.1. Each test set was used to evaluate the models trained on the same network. The validation sets were used to store the model parameters with the best performance

C

Table C.1: Dataset Settings.

Name	**Usage**	**Net.**	**Size** -	ψ^{ld} %	ψ^{pv} %	$\sigma^{m}_{\eta}, \sigma^{m}_{l}$ %	σ^{m}_{v} %	σ^{pm} %
Tr.C	train	CIGRE	$33k$	60	70	2	1	20
V.C	validate	CIGRE	$7k$	60	70	2	1	20
Te1.C	test	CIGRE	$7k$	60	70	1	0.5	10
Te2.C	test	CIGRE	$7k$	60	70	2	1	20
Te3.C	test	CIGRE	$7k$	60	70	5	3	50
Te4.C	test	CIGRE	$7k$	30	35	1	0.5	10
Te5.C	test	CIGRE	$7k$	30	35	2	1	20
Te6.C	test	CIGRE	$7k$	30	35	5	3	50
Tr.OB0	train	OB0	$16k$	60	70	2	1	20
V.OB0	validate	OB0	$3k$	60	70	2	1	20
Te1.OB0	test	OB0	$3k$	60	70	1	0.5	10
Te2.OB0	test	OB0	$3k$	60	70	2	1	20
Te3.OB0	test	OB0	$3k$	60	70	5	3	50
Tr.OB1	train	OB1	$15k$	60	70	2	1	20
V.OB1	validate	OB1	$3k$	60	70	2	1	20
Te1.OB1	test	OB1	$3k$	60	70	1	0.5	10
Te2.OB1	test	OB1	$3k$	60	70	2	1	20
Te3.OB1	test	OB1	$3k$	60	70	5	3	50
Tr.OB	train	OB	$15k$	60	70	2	1	20
V.OB	validate	OB	$3k$	60	70	2	1	20
Te1.OB	test	OB	$3k$	60	70	1	0.5	10
Te2.OB	test	OB	$3k$	60	70	2	1	20
Te3.OB	test	OB	$3k$	60	70	5	3	50

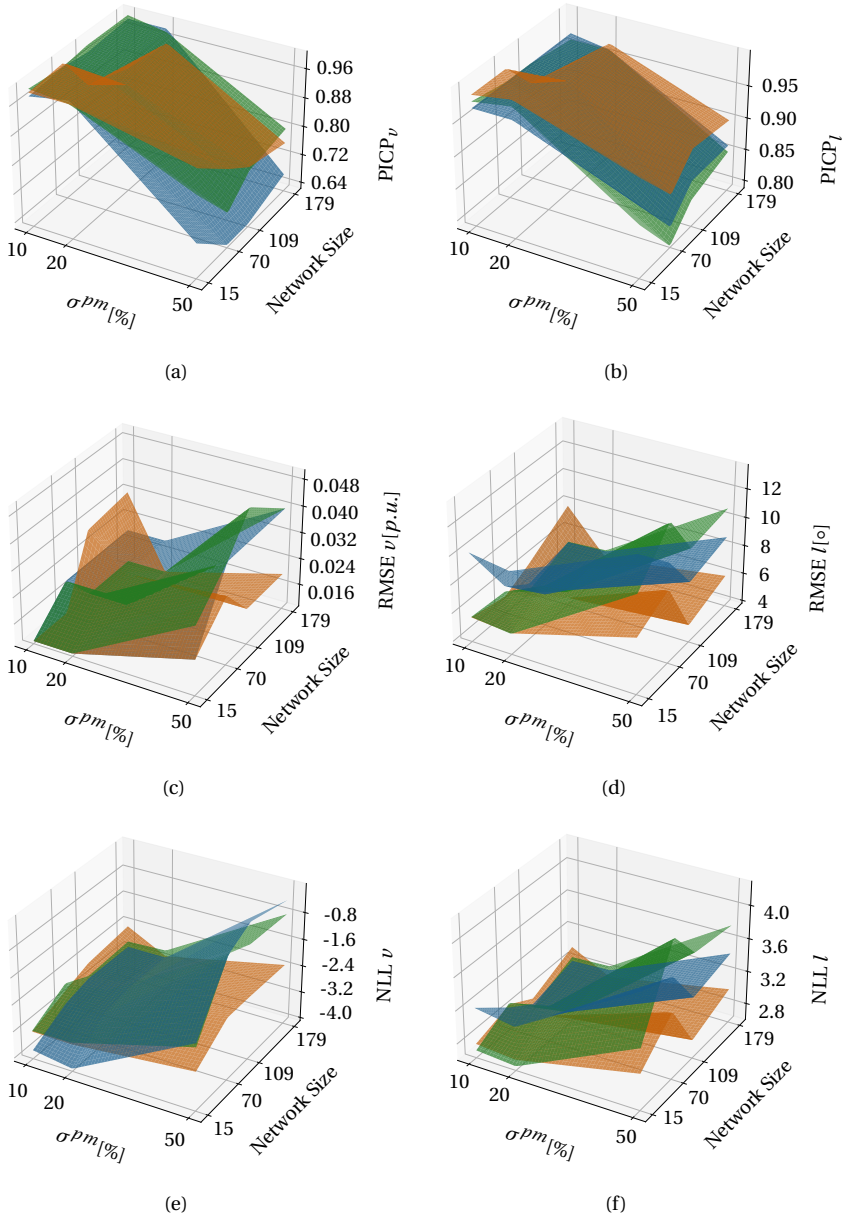


Figure C.1: Comparison of VI (▬), MCD (▬) and DE (▬) models under varying pseudo-measurement (and measurement) noise and network size. The network size corresponds to the number of buses.

Table C.2: Model Computational Aspects.

Net.	Training Duration [hours]			Prediction Duration [s]		
	VI	**MCD**	**DE**	**VI**	**MCD**	**DE**
Cigre	1.8	1	8.1	0.28	0.09	0.01
OB0	10.6	0.8	6.2	0.4	0.12	0.04
OB1	16.6	0.7	7.7	0.36	0.16	0.04
OB	14.7	1.2	9.8	0.46	0.3	0.03

during training for each network. Fig.C.1 illustrates the model performances for each metric under the varying noise levels and network sizes.
Considering $PICP_{95}$, all models were slightly underconfident $(0.95-1)$ in the low and medium noise scenarios, excluding the MCD which was overconfident in the low noise v for OB1 (0.89) and OB (0.85), and slightly overconfident for OB1 (0.92). In all models, a lower correlation in the test sets reduced $PICP_{95}$ but maintained its levels between $0.91-1$ for the low and medium noise. All models were impacted by the high noise scenarios, with MCD maintaining the best performance in most. Considering the voltage RMSE metric, all models showed fluctuations for the different test sets. MCD showed deterioration for lower and higher noise in the larger networks, which can indicate fitting to the noise uncertainty. However, MCD also showed a consistently better RMSE than VI and DE for higher noise levels. Considering the loading RMSE, the VI was highly impacted by the topological changes within the CIGRE datasets, whereas MCD and DE maintained a similar RMSE for all topologies and networks. The MCD mainly outperformed the DE and VI in loading RMSE and was less impacted by the test set noise variations. The NLL metric showed similar levels for all low and medium noise scenarios in all networks and models. However, MCD maintained almost consistent performance in all scenarios with minor noise impacts compared to VI and DE. The VI model was negatively impacted by the different DS topologies in the training dataset and reduced PV and load correlations in the test sets. The MCD and DE did not show significant impacts from the DS topologies. The reduced test set PV and load correlations did not significantly impact the MCD and DE, except for a slight improvement for the l $PICP_{95}$ and a slight deterioration for the l NLL. The results indicate that the MCD model is the most consistent among the 3 BNNs, especially with increased noise levels. Nevertheless, all BNN models showed low RMSEs, NLL, and close to 0.95 $PICP_{95}$.
Tab.C.2 shows the training and prediction durations per model and network using the Google Colab CPU. The number of input features, neurons, and outputs highly impacted the VI training time, with $\times 6-9$ larger durations for OB0, OB1, OB than CIGRE. MCD required significantly less time than DE, as MCD only trained one BNN, whereas DE trained 20 BNNs. The prediction durations were small for all networks, with DE being the fastest. The difference between MCD and DE prediction duration was potentially caused by the larger Z and the MCD having more layers (and parameters).

D

Machine Learning-based Instability Prevention Coordination Case Study Details

D.1. IBR Classifier Dataset

The IBR classifier dataset includes 93 load events from simulations with initially stable operating conditions. The load events were randomly sampled between the DS loads, with active and reactive power steps uniformly sampled between $[-15, 15]$[%] and minimum absolute active power change at $0.5[MW]$. H varied between 1–20$[s]$. The DER outputs and load consumption followed a typical summer profile in the Netherlands, with added random uniform noise with ± 25[%] bounds.

D.2. ADRR Model Dataset

The sampled parameters between simulations were the DER outputs, the DS load consumption, the external grid inertia, the faulted lines, and the fault durations. To increase the convergence of simulations, multiple simulations were performed with reduced noise, fixed power factor, or excluding profiles. H varied between 1–20$[s]$. The faulted lines were randomly sampled from the 9 options in Fig.5.4. The fault duration before clearance was randomly sampled between 0.1–0.25$[s]$ as [120] from a uniform distribution. The ADRR dataset had 1545 simulations, of which 284 were highly instable, and the rest are the regression data. The ADRR high instability classifier dataset includes all 1545 simulation data. The min-max normalization technique was applied to the regressor dataset features.

D.3. CASE STUDY SETTINGS

The case study of Sec.5.4.3 uses the test system, including the DS, excluding the ADRR. The risk example assumed $C^{\iota}(e) \approx 400k$[€] due to outage, $C^{\eta}(e) \approx 100$[€/%], and the fault e happening once every 2 years. The total durations for the setpoint optimizer were recorded using an i7-1185G7 CPU.

For the case study of Sec.5.4.4, the proposed ADRR regression model had 50 RNNs with 3 LSTM layers with 32 neurons, a dense layer with 64 neurons, and an FNN with 3 hidden layers with 256, 128, 128 neurons. In fine-tuning, the Tanhshrink activation function for the FNN showed better performance than common alternatives such as Tanh or ReLu. The baselines for the ADRR model are:

(i) An FNN with the sequential and non-sequential input features, four hidden layers with 1024, 512, 512, 256 neurons, and Tanhshrink activation function. The sequential features also included the 4 previous values at each timestep ($\mu = 4$).

(ii) 100 RNNs with 2 LSTM layers of 64 neurons with the $x(t)$ and repeated z features as inputs, followed by a 2-hidden-layer FNN of 128, 64 neurons with Tanhshrink activation.

(iii) The proposed structure, excluding the ADRR classifier. This model studies the impact from the extreme values observed in highly instable samples in ADRR.

(iv) The proposed structure, excluding the non-sequential features. This model studies the impact of the non-sequential features in ADRR.

All models had $\mu = 4$, a learning rate of 0.0003, and 300 epochs, excluding (i), which was trained for 3000 epochs. The stored models were the ones with the minimum validation error during training. All models had the same training, validation, and test sets, excluding (iii), which split the complete dataset into 60[%], 20[%], 20[%] training, validation, and test sets. The training durations for ADRR regression and classifier models were recorded using an L4 GPU.

For the case study in Sec.5.4.5 the classifier training process added a normal noise with 5[%] standard deviation on $y(t)$ to represent $\hat{y}(t)$ of x_c as in (5.19), accounting for the ADRR regression model. The proposed FB model has one hidden layer with 64 neurons and a Tanh activation function. The best-performing DB model used 5 nearest neighbors.

BIOGRAPHY

Demetris Chrysostomou (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Cyprus, Nicosia, Cyprus, in 2018, and the M.Sc. degree in robotics, systems and control from the ETH Zurich, Zurich, Switzerland, in 2021. He is currently a Ph.D. Candidate at the Technische Universiteit Delft, Delft, The Netherlands. He previously worked as a Research Assistant at the PSM laboratory of the University of Cyprus and as a Research and Development Intern at ABB Corporate Research. He is a member of the AI energy lab and the IEPG group of the EEMCS department of TU Delft. His research is part of the MEGAMIND program. His research interests are machine learning, control, and optimization applications for power systems operation.

List of Publications

PhD Journal Papers

4. **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"Selection for Flexibility Areas using Probabilistic Machine Learning Under Measurement Uncertainty"*, Under Review, 2025.

3. **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"Machine Learning-based Method to Support TSO-DSO Adaptive Coordination with Active Power Management for Instability Prevention"*, in International Journal of Electrical Power & Energy Systems, vol. 173, 2025.

2. **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"TensorConvolutionPlus: A Python package for distribution system flexibility area estimation"*, in SoftwareX, vol. 31, 2025.

1. **D. Chrysostomou**, J. L. R. Torres and J. L. Cremer, *"Tensor Convolution-Based Aggregated Flexibility Estimation in Active Distribution Systems,"*, in IEEE Transactions on Smart Grid, vol. 16, no. 1, Jan. 2025.

Other Journal Papers

3. M. Chrysostomou, N. Christofides, **D. Chrysostomou**, *"A Novel Machine Learning-Based Load-Adaptive Power Supply System for Improved Energy Efficiency in Datacenters"*, in IEEE Access. Dec. 2021.

2. U. Markovic, **D. Chrysostomou**, P. Aristidou, & G. Hug, *"Impact of inverter-based generation on islanding detection schemes in distribution networks"*, in Electric Power Systems Research, 190, 106610, 2021.

1. **D. Chrysostomou**, A. Dimitriou, N. D. Kokkinos, & c. A. Charalambous, *"Short-term electromagnetic interference on a buried gas pipeline caused by critical fault events of a wind park: A realistic case study"*, in IEEE Transactions on Industry Applications, 56(2), 2020.

Conference Paper

D. Chrysostomou, J. L. Rueda Torres, and J. L. Cremer, *"Exploring operational flexibility of active distribution networks with low observability"*, in IEEE Belgrade PowerTech, 2023.

Project Deliverables

3. *"Stability performance of future integrated DSO-TSO systems"* Report, 2025.

2. *"Future scenarios and research boundaries"* Report, 2023.

1. Generic power system model in Digsilent PowerFactory, for simulation of the dynamic performance of the interconnected future-like DSO-TSO power networks, 2023.

MSc Thesis Supervision

A. Neagu, *"Data-Driven Adaptive Dynamic Equivalent of Active Distribution Networks."*, 2023.

Teaching Assistant

MSc course *"Intelligent Electrical Power Grids."*, Electrical Engineering, Mathematics and Computer Science Faculty, TU Delft, 2022 & 2023.

TU Delft