Faculty for Electrical Engineering and Information Technology
Institute for Renewable and Sustainable Energy Systems

Future Energy System Dispatch and Control – Development of Decentralized Control Algorithms

Anna Elisabeth Kellerer

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Paolo Lugli, Ph. D.

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Thomas Hamacher
2. Univ.-Prof. Dr. Claudia Klüppelberg

Die Dissertation wurde am 11.11.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 12.09.2016 angenommen.

Abstract

Electricity generation is undergoing fundamental changes at the moment. In the past, electricity was centrally supplied by large plants connected to High Voltage (HV), high capacity transmission grids and owned by a few utilities. Small decentralized generation units are now taking over. These units are typically connected in the Low Voltage (LV) distribution grid and are owned by a multitude of parties. This transition of the power sector has important consequences for various power dispatch problems. It changes the number of units to dispatch, the units' owner structure and the previously unidirectional active power flow in distribution grids.

This thesis proposes two computational methods for solving such power dispatch problems arising in the grids of the future. The first method solves the Economic Dispatch (ED) problem in a distributed fashion that is inspired by Graphical Model (GM) methods. The second approach targets a technical challenge connected with distributed feed-in that guarantees voltage quality with a multitude of discretely controllable grid devices. It utilizes the complexity reduction in hand with the numeral reduction of discrete variables and is inspired by dynamic programming. The commonality of both methods is the 'divide and conquer' trick: the idea is to split a large, complex problem into small, simple subproblems, to solve them separately while coordinating them with the help of functions, and to thereby obtain the optimal overall solution.

The number of units to dispatch in future grids pushes central Economic Dispatch (ED) approaches, typically based on Mixed Integer Linear Programming (MILP), to their limits. Moreover, the required central knowledge of all component models conflicts with the privacy requirements in competitive, multi-owner markets. Decentralized dispatch approaches can meet such requirements but have to operate with minimal communication effort and guarantee the optimality of the results. For this purpose, a classical machine-learning Message Passing (MP) approach is adapted in two ways, for discrete and continuous decision variables. The discrete variable method can straightforwardly be deduced from belief propagation. The method gives an intuitive introduction and shows the applicability of the approach to distribution grid dispatch problems. The concept is then extended to the problem of continuous variable dispatch by exploiting the properties of the optimality conditions for our specific setup and, similar to approximated inferences, applying functional approximation techniques. The result is an ED method that can handle very different decision scales in a distributed fashion, with little communication and near-optimal results.

A technical challenge resulting from widespread generation units in the distribution grids is that active power flows may alter their direction and lead to dynamic, previously atypical voltage increases and drops. Reactive power feed-in, the deployment of capacitor banks and of On-Load Tap Changer (OLTC) are measures currently under discussion to resolve these issues. However, their control leads to a full Alternating Current (AC) Optimal Power Flow (OPF) with partially discrete variables. To avoid the drastic complexity increase arising from the use of discrete variables in classic, central OPF approaches, a method based on dynamic programming is developed. The overall problem is divided into problems of a single discrete decision variable, and these subproblems are coordinated by functions, similar to the messages in the dispatch approach. This method leads in theory to globally optimal dispatch results; a practical implementation with compact messages of one dimension can guarantee optimality if at least the reactive power feed-in of Distributed Energy Resources (DER) is pre-set.

Zusammenfassung

Aktuell ist die elektrische Energieerzeugung starken Veränderungen unterworfen. In der Vergangenheit wurde die elektrische Energie zentral von wenigen großen Kraftwerken, die an das Hochspannungsübertragungsnetz angeschlossen sind und wenigen Energieversorgern gehören, erzeugt. Jetzt wird mehr und mehr elektrische Energie dezentral von kleinen Einheiten, die im Verteilnetz angeschlossen sind und eine Vielzahl von Eignern haben, erzeugt. Diese Entwicklung hat einige Konsequenzen für die Steuerung im Verteilnetz, wie den Anstieg der Zahl zu kontrollierender Einheiten, ihre Eigentumsstruktur und die wechselnde Flussrichtung der Wirkleistung im Verteilnetz.

Diese Arbeit stellt zwei Methoden vor, die speziell diese Neuerungen berücksichtigen. Die Erste ist inspiriert von Lösungsansätzen für probabilistische grafische Modelle und löst das wirtschaftliche Einsatzplanungsproblem lokal. Die Zweite baut auf die Prinzipien der dynamischen Optimierung auf und bestimmt die optimale Regelung einer Vielzahl interferierender, diskret steuerbarer elektrischer Anlagen um die Spannungsqualität sicherzustellen. Gemeinsam ist beiden Methoden der 'teile und herrsche'-Ansatz: Ein großes komplexes Problem wird in kleine einfache Teilprobleme zerlegt und diese Teilprobleme über Funktionen koordiniert, um die übergreifende Lösung zu finden.

Die große Anzahl der zu steuernden Einheiten bringt viele der derzeitigen zentralen Einsatzplanungsalgorithmen, die häufig auf gemischt ganzzahliger linearer Optimierung basieren, an ihre Grenzen. Darüber hinaus steht das hierfür zentral benötigte Wissen im Konflikt mit der Wissenshoheit der unterschiedlichen Eigner. Dezentrale Ansätze umgehen diese Probleme, jedoch stellt die benötigte Kommunikation und Sicherstellung der Optimalität neue Herausforderungen dar. Um den Herausforderungen zu begegnen, wird in dieser Arbeit der nachrichtenbasierte Ansatz des maschinellen Lernens auf zwei Weisen weiterentwickelt, zum einen für diskrete Variablen zum anderen für kontinuierliche Variablen. Der diskrete Ansatz ist eine direkte Adaption des 'belief propagation'-Schemas. Er zeigt die Anwendbarkeit des Ansatzes für die Einsatzplanung und illustriert die Methodik. In diesem diskreten Ansatz ist es komplex, eine angemessene Diskretisierung zu finden, wenn die Maßstäbe der Einsatzplanungsvariablen sich stark unterscheiden. Der kontinuierliche Variablen Ansatz baut die Methodik dann über die Optimalitätsbedingungen und Anwendung einer Funktionsapproximation aus. Diese Schritte ergeben einen dezentralen, wirtschaftlichen Einsatzplanungsansatz, der mit stark unterschiedlichen Maßstäben umgehen kann, minimale Kommunikation gewährleistet und annähernd optimale Ergebnisse.

Eine andere technische Herausforderung, die aus einer Vielzahl von dezentralen und erneuerbaren Erzeugern erwächst, ist das Alternieren des Wirkleistungsflusses, der zu schwer vorhersagbaren, kurzfristigen Spannungshüben und senken führt. Diese Spannungsänderungen können mit Blindleistungseinspeisung, Kondensatorbänken und regelbaren Ortsnetztransformatoren kontrolliert werden. Ein Teil der Anlagen können jedoch nur diskret gesteuert werden, somit erweitert sich das komplexe optimale Leistungsflussproblem um diskrete Variablen. Um diesen Komplexitätsanstieg zu begegnen, stellt diese Arbeit einen dynamischen Optimierungsansatz vor, der das zentrale Problem in Probleme mit einer diskreten Variable zerlegt. Dieser Ansatz führt in der Theorie zu optimalen Ergebnissen. Eine erste Implementierung mit eindimensionalen Funktionen zeigt, dass die Optimalität garantiert werden kann, solange die Blindleistungseinspeisung der dezentralen Erzeuger gesetzt ist.

Acknowledgements

First of all, I like to thank my colleagues, my friends and my family for all the not in detail mentionable things – thematic discussions, personal advices, or simply a smile – that in sum made this thesis possible.

Prof. Dr. Claudia Klüppelberg encouraged me to do a PhD and aroused enduring research curiosity. I am very thankful for that and to have her as second reviewer, which closes the cycle from the initial idea of doing a PhD to the final thesis.

In between, Prof. Dr. Hamacher took the guidance white chary force, scientific freedom and inspiring ideas. I would like to thank him for the possibility to work on this topic and for his continuous parental support over the last years.

A special thanks goes to my Siemens supervisor Dr. Florian Steinke for the fruitful advisor-student relationship. This work was substantially formed by our regular white board sessions, creative idea Ping-Pong and discussions into the very detail.

Over the past three and a half years, I've been privileged to work with wonderful colleagues at Siemens and at Technische Universität München. Without whom this experience would not have been what it was.

As part of my Siemens Corporate Technology research group, I've had the pleasure to be entertained by surreal lunch discussions – a cats sense of balance in outer space, traveling from London to Paris without energy or the moon influence on climbing performance. It was not directly constructive but a valuable rest –Thank you.

I like to thank my colleagues at Lehrstuhl für Energiewirtschaft und Anwendungstechnik und Lehrstuhl Erneuerbare und Nachhaltige Energiesysteme for the excellent office-asylum in a warm and open atmosphere. I enjoyed been with pears, the unconventional afternoon re-feed pull-ups and 'running with sport'.

Moreover, I am thankful for the comments on this work by colleagues and friends: Annika Will, Florian Sänger, Franziska Aigner, Gerhard Aigner, Johannes Honold, Magdalena Dorfner, Maria Barthmes, Maria Recker, Philipp Kuhn, Rita Dornmair, Stephen Starck, Thomas Deetjen, and the incomparable Dominik Fischer.

In the ups and downs of the last years, I realized what an incredible family I have. Thank you all for the company especially my little nephew Alexander for his charming smile that makes you forget about all drawbacks and endless hospital visits. I am so proud of his strength and tenacity of life.

This last lines and the remainder of this thesis are dedicated to my boyfriend Maximilian: For the backing you have been for me in the last seven years, for the the admirable calmness you meet my tautness during the last months and for simply being how you are – the further most wonderful men I ever met.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The German word ‘Energiewende’ is an established term in the energy community worldwide[1] for the ambitious transition towards a “secure, environmentally friendly and economically successful energy future”[27]. ‘Energiewende’ combines a complex conglomerate of diverse goals and drivers that interfere with each other. The phase out of nuclear generation by 2022 in Germany counteracts the primary motivation to reduce the carbon dioxide emissions, due to the increased reliance on gas and coal. To account for that, renewable sources and more efficient energy use are planned as a counterbalance.

One of the main measures to implement this transition has been the renewable energy act, the ‘Erneuerbare Energien Gesetz’, in 2000, which was amended in 2004, 2009, 2012 and 2014 [25]. The ‘Erneuerbare Energien Gesetz’ guarantees a connection to the grid and a fixed feed-in tariff for newly installed renewable generation units. As a result of this incentive, the number of installed biomass driven plants, wind turbines and Photovoltaic (PV) panels has significantly increased in Germany [24], depicted in Figure 1.1.

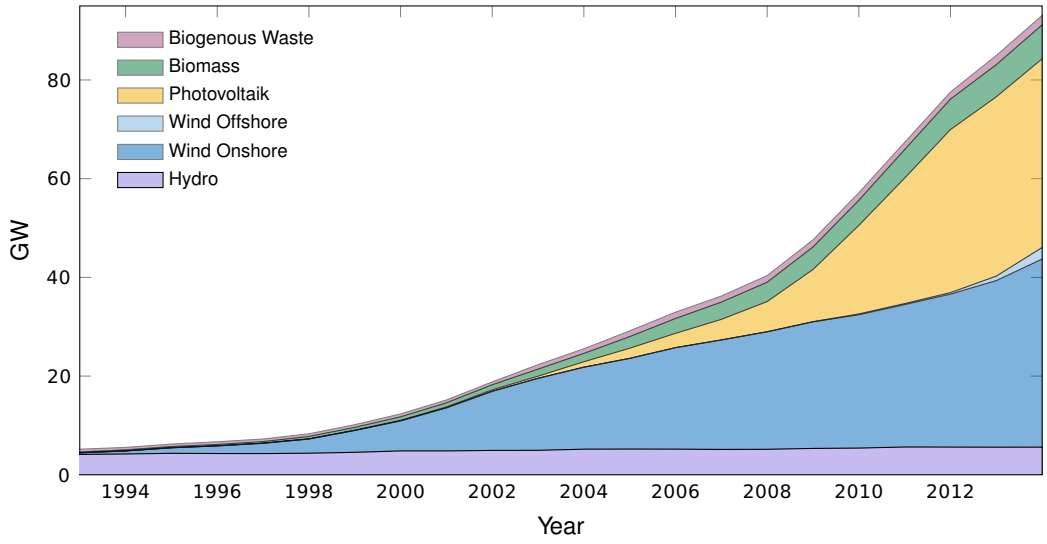


Figure 1.1: Development of the installed capacity for renewable energy generation units in Germany between 1993 and 2014 [24].

[1] An example is the introduction of the electricity conference at Austin, TX [112].

With intermittent sources, e.g. wind and sun, oversupply alternates with undersupply, which needs to be balanced. Demand side management, e.g. heaters [116, 60], Electrical Vehicles (EVs) [80] and deferrable industrial processes [75], can offer flexibility for this balancing and might add further Distributed Energy Resources (DER) in addition to the generation. All this results in a steadily growing number of DER that have consequences for dispatch and voltage control in electrical grids. These developments ensue the algorithmic requirements for such controls. In the following two sections, these requirements are discussed, first for Economic Dispatch (ED) and second for voltage control.

Planning horizons for conventional plants span over several years. Within this timeframe the conventional plant component model is incorporated into the dispatch program. Some DER eliminate these long planning horizons: PVs require a few months of planning, small-scale batteries even less, and EVs are simply plugged in. This is a comparably short term notice for adaptation and future dispatch approaches have to account for that.

The fixed 'Erneuerbare Energien Gesetz' feed-in tariff has made and still makes the payback of investments in renewables predictable. This relatively low market risk has attracted small investors. For institutional investors and utilities, the rate of return for these renewable units was not satisfactory, especially in the beginning. They left the field to small investors, whose participation in the market was additionally enabled by the lower investment magnitudes for DER. Municipalities, local manufacturers, cooperatives down to individuals add to the traditional few utilities and diversify the owner structure. This diversity hinders freely sharing component models for control applications. Previously, these models were centrally gathered within one of the few utilities and used for dispatch.

Subsequent to the first boom, scaling effects have reduced the price for renewable units [13, 135]. This has made renewables economically attractive, not just for environmentally aware groups but also from a pure investment perspective. Countries like the United States of America and China are now becoming leaders in capacity expansions of renewables [93]. This demand for renewable technologies is pushing technological progress and price reductions, which indicates a lasting growth in renewable units worldwide. In addition to these renewable units, a significant number of demand side management units might need to be integrated. Thus, dispatch approaches need to be scalable.

Electricity demand and production need to be balanced, since electricity cannot be stored directly. Instead of a few centrally placed plants with communication infrastructure, a large number of DER have to be coordinated with demand. This requires communication between these units spread in the distribution grid. However, communication infrastructure in distribution grids is almost non-existent or very limited [34] at the moment. The lack of bandwidth and the high number of DER to schedule precludes extensive communication, e.g. to a central intelligence. At the same time, multiple minimal controls are insufficient due to the required frequent communication and its delay, even when approached in a decentralized manner. As time is a determining factor for dispatch, not just a minimal but an appropriately complex control is needed.

In Germany, 73% of the renewables are installed in Low Voltage (LV) and Medium Voltage (MV) level [26], where the former goes up to 1kV and the latter ranges from 1kV to

less than 110kV [28] [2]. Both constitute the voltage levels in distribution grids. In contrast, large conventional plants are almost exclusively connected to High Voltage (HV), greater than or equal 110 kV [28], and high capacity transmission grids. Tighter distribution line capacities pose active operation restrictions on resources in highly DER penetrated grids. Moreover, existing electrical equipment, e.g. transformers, may overload because of the additional flow. An operation of DER with respect to those restrictions significantly reduces the required distribution grid extensions [3]. Thus, to circumvent costly grid extensions the operation of all DER has to be aligned with the grid limits.

The transition from a few central plants to a vast number of installed DER in the distribution grid necessitates dispatch approaches, which are scalable, quickly adaptable and which incorporate grid limits, preserve privacy, and guarantee optimal results with minimal communication effort. Central optimization approaches conflict with at least three of these requirements – scalability, quick adaptability and preserve privacy. The idiosyncrasy of central approaches, like the central Mixed Integer Linear Programming (MILP) approach in [31], for white box component models contradicts the privacy. Their integration hinders a quick adoption for new components. Moreover, the scaling of central approaches is insufficient for an extremely high number of units. This gives rise to the question whether there exists a decentralized dispatch approach that can handle these challenges: this question is addressed in the first part of this thesis.

Moreover, the distribution grid infrastructure was built and operated to facilitate voltage limits for prospective demands. Instead of central demand supply by controllable conventional units, today intermittent renewables spread in the distribution grid contribute a significant share of this demand or oversupply it. This development forces an infrastructure designed for unidirectional active power flow to deal with a bidirectional flow [44, 80].

Normally, transformers set a high voltage at the point of connection, which was consecutively lowered by the demands along the branches, so that the last customer was supplied with an acceptably high voltage. In the past, the transformer setting was manually adjusted at most few times a year, according to seasonal load patterns. Now, significant feed-in by DER alters with demand, either raising or lowering the voltage magnitude to the limits of the voltage band within a day. This gives rise to transformers that automatically gradually adjust their tap according to the current situation, i.e. On-Load Tap Changer (OLTC) [3]. Transformers in one distribution grid interfere with each other leading to a complex problem of multiple, discrete tap position decisions.

Voltage fluctuations occur not only at the point of connection but also within the grid. A feed-in dominated branch can in total even-out a demand dominated one and still locally affect the limits. In the absence of direct measures on the voltage, e.g. transformers, here reactive power can be utilized to ensure the voltage band [80]. Reactive power can be used in two ways, either to reduce the voltage drop when its flow is reduced or to increase the voltage drop when its flow is increased [119]. Some DER allow for continuous reactive power feed-in or capacitor banks can gradually supply reactive power [38]. Similarly to

[2] This definition differs slightly from the IEEE definition [43], where the LV is less than 1kV, MV greater than or equal 1kV and less than 100kV and High Voltage (HV) is greater than or equal 100kV and less than 230kV.

the transformers, these reactive power sources depend on each other as well as on the transformer settings.

In addition to the voltage limits, the voltage has an influence on active power losses that contributes value to its control. Line losses are minimal for a homogenous high voltage profile since these losses are determined by the current flows [119]. Hence, for loss reduction the voltages should be as high as possible within its limits. That can be concurrently obtained by the previously described measures.

Thus, simultaneously with the dispatch of DER, the electrical equipment needs to be optimally controlled to facilitate a distributed and renewable production. This problem, to satisfy voltage limits at minimal losses, is called Volt-VAr Control (VVC) problem due to the two influence factors voltage, Volt, and reactive power, VAr. OLTC and capacitor banks contribute additional flexibility to ensure voltage limits and reduce losses but have a discrete control nature. Moreover, these discrete controllable units interfere with each other via the grid topology, which complicates their control. Optimal control of these discrete controllable devices is the question addressed in the second part of this thesis.

1.1 Contributions

Grid topology implied problem decomposition is the idea that inspired the methods in this thesis. The decomposition into subproblems, efficient subproblem solving and coordination by functions are the methods for the proposed algorithms – 'divide and conquer'. This concept is implemented in two fashions, two Graphical Model (GM) inspired methods and one dynamic programming related method.

The decentralized, cost optimal dispatch algorithms for radial distribution grids are based on probabilistic GM methods. Therefore, the dispatch problem is reformulated as an optimization problem on a graph. For this reformulation the algorithms capitalize the dependence structure given by the grid topology to split the global problem into small local/nodal problems. Then, these local/nodal problems are coordinated with transportation cost functions/messages such that the local solution is equivalent to the global solution. The key for this Economic Dispatch (ED) application is the energy conservation at each node; it splits the global energy conservation constraint into equivalent nodal balances based on transport equations. The transportation cost functions/messages enrich the node with a full knowledge of their influence. Moreover, these approaches ensure the new requirements: scalable, quickly adaptable, incorporate grid limits and preserve privacy.

The first implementation of this idea is based on discrete variables and has been introduced by the author in [73]. Here the coordinating variables are assumed to take values in a discrete set. Hence the communicated transportation cost function/message is given as a vector of the set size. This allows for arbitrary DER cost function shapes and non-convex feasibility sets. For rising, non-convex cost functions and on-off ruled feasibility sets this special purpose solver outperforms the MILP solver CPLEX.

However, for DER and lines of significantly different scales an equally fine discretization becomes computationally intractable. To overcome this, a continuous variable approach is formed. This generalization requires the restriction to convex cost functions and feasible

sets. In this convex framework a novel efficient message update mechanism is developed, based on local optimality conditions. Here only five parameters fully characterize the transportation cost function/messages. Thus, the communication effort of this non-iterative approach is smaller by magnitudes than other decentralized approaches, e.g. Alternating Direction Method of Multipliers (ADMM) [79].

The optimal VVC problem with a number of discrete controllable variables, i.e. OLTC or capacitor banks, is addressed by an especially adapted dynamic programming approach. This dynamic programming procedure splits the complex VVC problem into subproblems of one discrete variable and coordinates these by pseudo generator functions. The pseudo generator functions include the optimal subproblem solutions of interrelated problems by cost of production/cost-to-go, reactive and active power flow. These functions are effectively obtained by utilizing the monotonous step nature of the optimal discrete variables. Thereby, this approach finds the optimal solution in contrast to heuristics, like hill climbing [6, 57]. This is of especial value in future active distribution grids where multiple tap changers and capacitor banks interfere with each other.

1.2 Overview

This thesis is structured in four main chapters, namely Background, Chapter 2, Economic Dispatch (ED) as Graph Problem, Chapter 3, Dynamic Programming Volt-VAr Control (VVC), Chapter 4, and Outlook, Chapter 5.

Chapter 2 provides a formal description of the problems examined in this thesis, offers a fundamental understanding of the methodology as well as an introduction to the relevant literature. The ED problem is the starting point, Section 2.1, extended by physical power flows to Optimal Power Flow (OPF), Section 2.2, and further to the VVC problem, Section 2.3. The problem formulations are followed by a brief introduction to the basic methodologies, GM methods Section 2.4 and dynamic programming Section 2.5.

In Chapter 3 the GM inspired methods to solve the ED problem are detailed. First in Section 3.1 the problem is reformulated as a graph problem, Section 3.2 derives the GM inspired message based solution strategy and Section 3.3 discusses the implications of this Message Passing (MP) strategy for ED. Second the implementation and experiment results for the Discrete Message Passing (DMP) and Continuous Message Passing (CMP) approach are stated in Section 3.4 and Section 3.5.

Chapter 4 proposes the dynamic programming solution to the VVC problem. The optimality principle for the VVC subproblems in dynamic programming approach is developed in Section 4.1 as well as a motivation for its structure. Then the three steps, namely obtaining subproblems Section 4.2.1, solving these subproblems Section 4.2.2 and re-obtaining the overall solution Section 4.2.3, for an implementation are detailed in Section 4.2. Finally in Section 4.3 the approach is tested in a simplified distribution grid Section 4.3.1 and some realistic distribution grids Section 4.3.2.

Chapter 5 sets the requirements for the derived theoretic methods in a practical context.

1.3 Notation

Let $\mathbb{R}$ be the real numbers, $\mathbb{C}$ complex numbers and j the imaginary unit. The complex conjugate transposed for a variable x, vector $\mathbf{x}$ and matrix $\mathbf{X}$ is denoted by $\cdot^H$ and the componentwise multiplication by $\odot$. The variable x upper and lower limit is given by $\overline{x}$ and $\underline{x}$; the vector $\mathbf{x}$ componentwise upper/lower limit is given by $\overline{\mathbf{x}}$/$\underline{\mathbf{x}}$. $|\mathcal{X}|$ denotes the cardinality of the set $\mathcal{X}$; $|\mathbf{x}|$ is the vector of the componentwise absolute values. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \leq \mathbf{y}$ denotes the componentwise inequality and for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $\mathbf{x} \leq \mathbf{y}$ is used to represent $Re(\mathbf{x}) \leq Re(\mathbf{y})$ and $Im(\mathbf{x}) \leq Im(\mathbf{y})$. The diagonal operator $diag(\mathbf{X})$ yields the main diagonal elements of the matrix $\mathbf{X}$ as vector.

A graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is given by its set of vertices $\mathcal{V}$ and edges $\mathcal{E}$. An element $v \in \mathcal{V}$ is called a vertex/node and $e \in \mathcal{E}$ is called an edge. Edges can be either directed, $e = (vw)$, or undirected, $e = \{vw\}$. A graph consisting of (un-)directed edges is called (un-)directed. Two vertices are called adjacent if they have a common edge. All adjacent vertices to v form the neighborhood $\mathcal{N}(v)$ of v, each $w \in \mathcal{N}(v)$ is a neighbor of v. The subgraph $\mathcal{G}[\mathcal{V}_l]$ is induced by a subset of nodes $\mathcal{V}_l \subset \mathcal{V}$ and consists of vertices in $\mathcal{V}_l$ and all edges in $\mathcal{E}$ connecting two such vertices.

A sequence of unique adjacent vertices is called walk. If all vertices in a walk are distinct, it is called a path. A walk with the same start and end vertex and distinct vertices in between is called a cycle. Graphs are categorized as cyclic/loopy if they contain cycles or acyclic/radial if not. The graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is called complete if each pair of vertices in $\mathcal{V}$ is connected by an edge in $\mathcal{E}$. A clique C is a complete subgraph $\mathcal{G}[\mathrm{C}]$ of $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The clique C is maximal if there exists no complete subgraph greater than and containing C. The clique size is the maximal number of vertices in a clique in $\mathcal{G}(\mathcal{V}, \mathcal{E})$.

In this thesis all electrical grid graphs $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are considered to be undirected and the introduced graph notation directly transfer to these graphs. By convention, the grid graph's root node is the highest voltage level leaf or transformer. The parent set $\mathcal{P}(v)$ denotes all upstream adjacent nodes, i.e. nodes closer to the root node, and children $\mathcal{C}(v)$ all downstream adjacent nodes, i.e. nodes further apart from the root node. Power flow analyze in these grids are carried out in a per-unit system that is relative to nominal voltage $\hat{u}_{v,n}$ and apparent power s_n. Then, the v node voltage u_v is defined as a fraction of its reference voltage level $\hat{u}_{v,n} \in \mathbb{R}$, i.e. the absolute complex voltage is $u_v \hat{u}_{v,n}$; the same is true for s_v. The base values for impedance $s_n/{\hat{u}_{v,n}}^2$ and admittance ${\hat{u}_{v,n}}^2/s_n$ follow naturally. The ground node's reference voltage level is defined to $u_{0,n} = 0$. In this notation, the all voltages are relative to their level and thus ease the notation at transformers, e.g. Section 2.3.

Chapter 2

Background

This chapter provides an overview of the applications and solution methodologies and introduces the notation and related literature.

The starting point is the overall welfare macro perspective of the Economic Dispatch (ED), Section 2.1 and see red frame in Figure 2.1. The problem here is to minimize the production cost minus the utility of consumption, while the active power flow, if under consideration, is determined by lossless transport. This perspective is refined by the physical power flow to Optimal Power Flow (OPF) in Section 2.2, blue frame in Figure 2.1, and further refined by the consideration of the voltage and reactive power to the Volt-VAr Control (VVC) problem in Section 2.3, green frame in Figure 2.1.

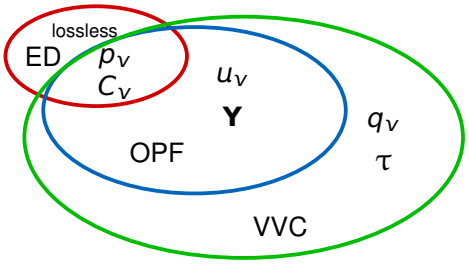


Figure 2.1: Overlaps between variables and objectives of the ED, OPF and VVC problem.

Finally the two applied solution methodologies are briefly introduced, Graphical Model (GM) in Section 2.4 and dynamic programming in Section 2.5. The message based GM approach is a fine cliques-wise decomposition method, Figure 2.2 in red, that coordinates the subproblems by functional messages. Dynamic programming in general takes a coarser decomposition, Figure 2.2 in blue, but utilizes also functions for coordination.

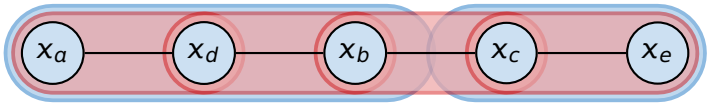


Figure 2.2: Examplary illustration of the decomposition granularity difference between the GM, red, and the dynamic programming, blue, approach.

2.1 Economic Dispatch

The Economic Dispatch (ED) problem is to schedule a number of electricity generation or consumption units such that the overall cost, i.e. the summed cost of all production minus the utility of all consumption, is minimized under the conservation of active power [138, 15]. Power lines and generators can pose additional limits on the feasible flows, generation and consumption. In the following the general ED problem is formulated and extended to cover those constraints. This problem overlaps with the Unit Commitment (UC) problem, i.e. the problem to determine which unit should run given the units' feasible set $\mathcal{P}_v$, when focus is turned on the non-convex units' feasible set $\mathcal{P}_v$. Likewise, it overlaps with minimum cost flow problems, i.e. the problem to find the flows in a graph with minimal costs, when the focus is placed on the feasible transports $\mathcal{P}_{vw}$. Both foci are discussed below. Finally, the transport model and its grounds in radial distribution grids are analyzed.

For electricity generation and consumption units $\mathcal{V}$, let p_v denote the net active power injection at node $v \in \mathcal{V}$, where a positive sign means consumption and a negative sign production. The associated function $C_v(p_v)$ is the cost or negative utility function, which is defined on the feasibility set $\mathcal{P}_v$. Then, the one time step ED problem reads as

$$\min_{p_v \in \mathcal{P}_v} \quad \sum_{v \in \mathcal{V}} C_v(p_v), \tag{2.1}$$

$$s.t. \quad \sum_{v \in \mathcal{V}} p_v = 0. \tag{2.2}$$

If no constraints apply, i.e. $\mathcal{P}_v$=$\mathbb{R}$, it holds that an optimal assignment $\mathbf{p} = (p_v)_{v \in \mathcal{V}}$ of (2.1) subject to (2.2) satisfies that all marginal costs for generation equal marginal gains for consumption, i.e. $\partial C_v(p_v)/\partial p_v = \mu$ for all $v \in \mathcal{V}$. Then, neither producers nor consumers improve the overall wealth by altering their active power injection p_v, due to the constraint-enforced (2.2) counter action[1]. When imposing convex feasibility intervals $\mathcal{P}_v$, this only needs to hold for the last infinitesimal production and consumption unit. Since active constraints, i.e. constraints satisfied with equality, enforce higher marginal costs at the lower interval limit and costs smaller at the upper interval limit[2]. This parity can be obtained graphically by observing the intersection of the marginal utility curve and the marginal costs curve per unit. This is known as the microeconomic supply-and-demand model, which holds in roughly competitive markets [108].

In energy markets the merit order curve is an approach similar to the marginal supply curve. A step function where the steps' heights are fuel cost times heat rate of the generation units sorted from the lowest to the highest cost and the step lengths are the summed respective capacities [113], i.e. assuming linear generator cost functions. It decides which units should run from a pure fuel cost perspective, where only the last unit's output needs to be adjusted for non-convex feasibility sets $\mathcal{P}_v$. For more complex

[1] This property can also be derived from first KKT condition $\{\partial C_v(p_v)/\partial p_v\}_{v \in \mathcal{V}} + \mathbf{e}\mu = 0$ [128, 16], where $\mathbf{e}$ is a vector of ones in dimension $\mathbf{p}$, or via Lagrange multipliers, i.e. $\tilde{C}(\mathbf{p}) = \sum_{v \in \mathcal{V}} C_v(p_v) + \lambda\left(\sum_{v \in \mathcal{V}} p_v\right)$, claim $\partial\tilde{C}/\partial p_v = 0$ and $\partial\tilde{C}/\partial\lambda = 0$ [22].

[2] Here, the KKT condition reads as $\{\partial C_v(p_v)/\partial p_v\}_{v \in \mathcal{V}} + \mathbf{1}\mu + e\overline{\lambda} - e\underline{\lambda} = 0$ with $\overline{\lambda}, \underline{\lambda} \in \mathbb{R}^+$ [128, 16].

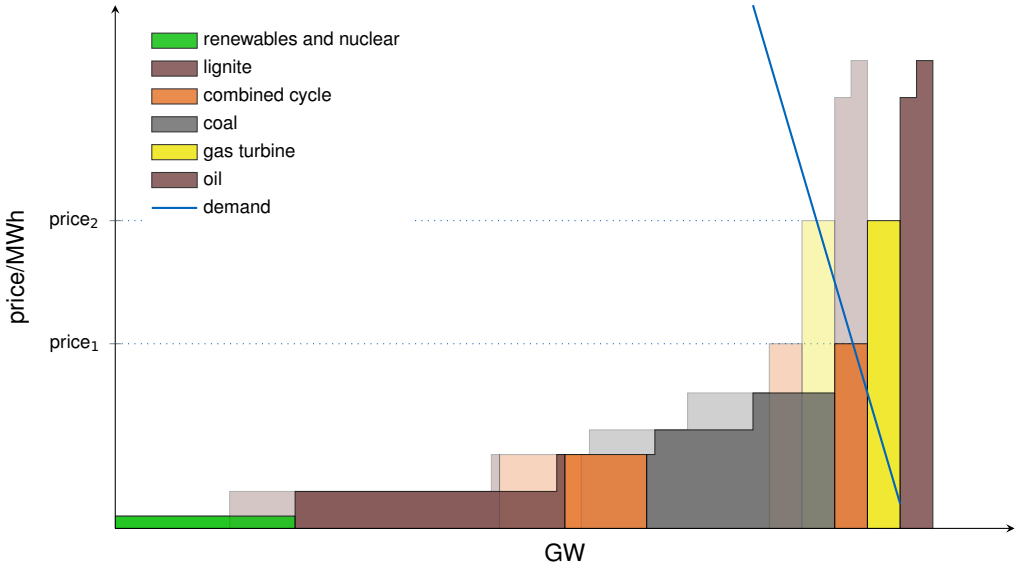


Figure 2.3: Merit order curves with and without renewable energy. The translucent staircase represents the supply curve without contributions by the renewable sources and the more covering staircase represents the supply curve with contributions by renewables; the blue line shows the demand. The merit order effect is the gap between the market clearing price in the pure conventional power market ($price_1$) and the clearing price in the market with renewables ($price_2$).

constellations, e.g. starting costs, ramp rates or transmission constraints, this approach cannot be straightforwardly applied. Nevertheless, heuristics adjusting for those additional requirements can still lead to feasible and especially quickly calculable dispatch results [39], but with limited theoretical guarantees. Similar holds for non-linear generation cost functions.

Moreover, the merit order effect which describes the electricity price reduction because of the feed-in of renewables was created using this approach [121]. Since some renewable generation driving fuels are free of charge, e.g. wind or sun, they are added to the lower end of the merit order and shift the conventional merit order by their current, intermittent production, and thus reduce the price at parity, see Figure 2.3 [3].

Since for most generators the feasible interval, with minimum $\underline{p}_v$ and maximum $\overline{p}_v$, generation does not include 0 (if switched on) the overarching feasible set

$$\mathcal{P}_v = \{0\} \cup [\underline{p}_v, \overline{p}_v]$$

is non-convex. Determining the set of operating generation units over a number of time steps minimizing production costs, while satisfying demand, is the UC problem. Here the complexity arises from the non-convex feasibility set [36]. A variety of methods are applied to handle these problems; a detailed review can be found in [105]. Similar to the merit order approach, heuristics, like priority list [85, 120] and expert systems [98],

[3] The merit order curve is abstracted from German merit order curve per technology in [113].

can be calculated quickly but lack qualitative guarantees, while a variety of optimization approaches, e.g. dynamic programming [104] and Mixed Integer Linear Programming (MILP) [40, 31, 64], also supply qualitative measures for the results' quality. This thesis focuses on the MILP method for ED problems. In the MILP approach an additional binary variable b_v curtails the production limits to zero in the off status in order to reflect non-convex $\mathcal{P}_v$, i.e. $b_v \underline{p}_v \leq p_v \leq b_v \overline{p}_v$.

This formulation allows for moderate model complexity, e.g. to include start-up costs [31] or more detailed thermo-physical ruled costs [64], while providing good solver performance control and scalability for medium-sized problems. For example, a solution can be found using the branch-and-bound method, which builds a tree of additional constraints in order to obtain an integral solution from the enlarged, relaxed problem, i.e. the problem without integrality constraints. This approach can guarantee a solution within a predefined accuracy [137]. Moreover, additional constraints added in line with the integrality allow for additional performance improvements [64]. These additional constraints/cuts make some non-integral solutions infeasible but let all integral solution feasible, analog to the idea of the cutting plan approach [118]. For these reasons, MILP has evolved to the dominant method applied in commercial dispatch systems since its first implementation at a utility by Dillon et al. [40].

Transport restrictions are becoming more and more prevalent when evolving from large generation units placed in the transmission grid, with high line capacities, to a multitude of Distributed Energy Resources (DER) spread in the distribution grid, with small line capacities. When turning the focus to the transports p_{vw} with capacities $\mathcal{P}_{vw}$ and defining the injection via the nodal balance

$$p_v = \sum_{w \in \mathcal{N}(v)} p_{vw},$$

the optimization problem (2.1) subject to (2.2) can be reformulated as minimum cost flow problem. A minimum cost flow problem aims at minimizing transport cost subject to transport constraints and given demands.

In this framework the nodal costs C_v, $v \in \mathcal{V}$, need to be transformed into flow costs, therefore an additional node 0 is added and edges are drawn from this node to each initial node $v \in \mathcal{V}$. These edges are assigned a flow capacity of $\mathcal{P}_{0v} = \mathcal{P}_v$ and costs $C_{v0}(p_{0v}) = C_v(p_v)$. Then the flows $p_{0v} = -p_{v0}$ are equal to the p_v, and the p_v in the initial formulation drains to the reference node 0 leading to injection $p_v = 0$ for all $v \in \mathcal{V} \cup \{0\}$. Moreover, $p_0 = 0$ ensures constraint (2.2). Hence the minimum cost flow problem, with $\mathcal{N}(0) = \mathcal{V}$, reads as

$$\begin{aligned} \min_{p_{vw} \in \mathcal{P}_{vw}} \quad & \sum_{v \in \mathcal{V}} C_{v0}(p_{v0}), \\ s.t. \quad & \sum_{w \in \mathcal{N}(v) \cup \{0\}} p_{vw} = 0, \ \forall v \in \mathcal{V} \cup \{0\}. \end{aligned}$$

Combinatory algorithms to solve the minimum cost flow problem for convex quadratic costs in polynomial time date back to the 1980s, e.g. [95, 96]. From then onwards different

algorithms have been proposed, which can be categorized in cycle-cancelling, successive shortest path and primal-dual out-of-kilter algorithms. Many of them combine a sequence of shortest path calculations, i.e. finding paths with minimal costs [7]. For example, the successive shortest path algorithm initially sets the minimal feasible flow p_{vw} on all edges. Then, the flows are enlarged along the shortest paths from positive, source

$$\sum_{w \in \mathcal{N}(v) \cup \{0\}} p_{vw} > 0,$$

to negative, sink

$$\sum_{w \in \mathcal{N}(v) \cup \{0\}} p_{vw} < 0,$$

imbalanced nodes until all nodal imbalances are resolved. These algorithms are still of research interest. Modifying these approaches results in powerful polynomial time algorithms for solving continuous variable minimum cost flow algorithm [129]. Even though computationally efficient, minimum cost flow algorithms require convex feasibility sets and global knowledge, e.g. aggregated at additional node 0.

In this setup with known injection $p_v = 0$, $v \in \mathcal{V} \cup \{0\}$, global coordinating node 0 and piece-wise linear cost, Message Passing (MP) which is introduced in Section 2.4 has been proven to converge to an optimal solution [50][4]. Similar guarantees for general loopy models are still undergoing research.

Gradually aggregating the balances of areas at area aggregator nodes, instead of at a single node 0, relates the minimum cost flow formulation to local master approaches. These methods, e.g. [125], received growing interest in recent years, since they handle a multitude of different operators well, respect transport restrictions and partly secure information privacy. Other approaches meeting this requirements are market-based, e.g. SO-EASY [94] and PowerMatcher [76]. Here intelligent local agents bid on a central marketplace whose market clearing then determines the dispatch result. While this approach is extremely scalable and suits a competitive environment well, it does not achieve a globally optimal dispatch result, even if the local agent models are very detailed and are solved in an optimal way, e.g. by employing dynamic programming approaches [136].

A final remark concerning the models used, all problem formulations have been represented in the transport model, ignoring for the moment Alternating Current (AC) flows, which will be discussed in Section 2.2, and the DC-approximation [138]. The DC-approximation draws the flows closer to the actual physical flows, since a significant flow on only a section of a cycle in the grid which can result from transport models requires two different angles θ_v at one node and is precluded by the angle θ_v uniqueness. But for distribution grids, which are radial in nature, this inconsistency does not occur and the angles θ_v, $v \in \mathcal{V}$, can be directly recalculated from the transport flows. A qualitatively similar result holds for AC flows in trees, enabling the calculation of the full AC flow by the branch flow equations [12], called DistFlow. Accordingly, the root node voltage in combination with real and reactive feed-in at all nodes fully determines the AC power flow.

[4] The nomenclature belief propagation for the method used in [50] is equal to MP.

2.2 Optimal Power Flow

The Optimal Power Flow (OPF) problem combines the ED, Section 2.1, with the physical, Alternating Current (AC) power flow. Here the variables are the complex voltages

$$u_v = \hat{u}_v \exp j\theta_v$$

at node $v \in \mathcal{V}$ composed of the voltage magnitude $\hat{u}_v$ and voltage angle θ_v. Thus, the OPF problem is to determine the complex voltage vector $\mathbf{u}$ such that the implied electricity generation and consumption maximize the overall welfare subject to the limits above.

In the following, all power flows are calculated utilizing the π-model. The current flow along the line $\{vw\} \in \mathcal{E}$ is fixed by the complex voltage difference $(u_v - u_w)$ of its defining nodes and the opposition of the line, admittance y_{vw}[5], to

$$i_{vw} = y_{vw}(u_v - u_w).$$

Additionally, each node has a flow with opposition y_{v0} to the reference/ground node 0 with complex voltage $u_0 = 0$. Those current flows determine by Kirchhoff's current law [83, 89] the nodal currents as follows

$$i_v = y_{v0}u_v + \sum_{w \in \mathcal{N}(v)} y_{vw}(u_v - u_w),\ v \in \mathcal{V}. \tag{2.3}$$

In matrix vector notation, (2.3) reads as $\mathbf{i} = \mathbf{Y}\,\mathbf{u}$ with admittance matrix $\mathbf{Y}$, given by its vw-entries below

$$(\mathbf{Y})_{vw} = \begin{cases} y_{v0} + \sum\limits_{\{vk\} \in \mathcal{E}} y_{vk} & v = w, \\ -y_{vw} & v \neq w,\ \{vw\} \in \mathcal{E}, \\ 0 & else. \end{cases}^{6} \tag{2.4}$$

The complex powers are then given by the complex voltage u_v times the complex conjugated current ${i_v}^H$ to

$$s_v = u_v {i_v}^H = p_v + j\,q_v,\ v \in \mathcal{V}, \tag{2.5}$$

here p_v is the active power and q_v the reactive power. In matrix vector notation these read as

$$\mathbf{s} = \mathbf{u} \odot \mathbf{i}^H = \mathbf{u} \odot (\mathbf{Y}\,\mathbf{u})^H, \tag{2.6}$$

with component-wise multiplication $\odot$ and complex conjugate transpose $\cdot^H$.

Then the real power feasible set $\mathcal{P}_v$ needs the reactive power feasible set $\mathcal{Q}_v$ to fully characterize the node v's feasibility sets. For convex $\mathcal{P}_v$ and $\mathcal{Q}_v$ this reads as

$$\underline{\mathbf{s}} \leq \mathbf{u} \odot (\mathbf{Y}\,\mathbf{u})^H \leq \overline{\mathbf{s}}, \tag{2.7}$$

but due to the quadratic nature in $\mathbf{u}$ the feasible set in $\mathbf{u}$ is not convex. Moreover, the voltage magnitude band

$$\underline{\hat{\mathbf{u}}} \leq |\mathbf{u}| \leq \overline{\hat{\mathbf{u}}}, \tag{2.8}$$

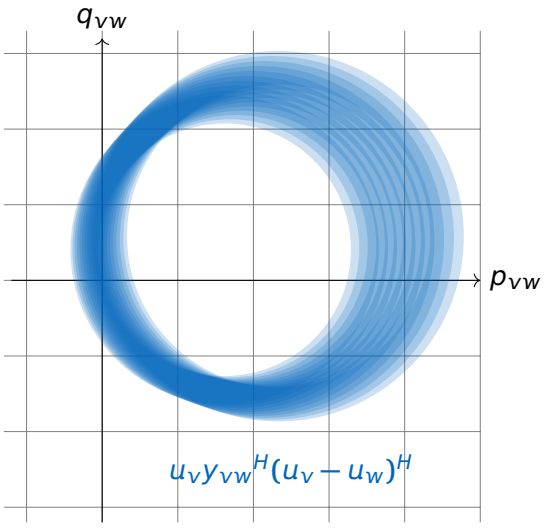


Figure 2.4: Non-convex feasible flow, blue, on a line implied by the voltage limits (2.8).

adds another non-convexity, as shown in Figure 2.4.

Then the feasible voltage set is given by

$$\mathcal{U} = \{\mathbf{u} \in \mathbb{C}^{|\mathcal{V}|} : \underline{\hat{\mathbf{u}}} \leq |\mathbf{u}| \leq \overline{\hat{\mathbf{u}}} \wedge \underline{\mathbf{s}} \leq \mathbf{u} \circ (\mathbf{Y}\,\mathbf{u})^H \leq \overline{\mathbf{s}}\}. \tag{2.9}$$

A visualization of the non-convex $\mathcal{U}$ for a loopy, three-bus grid can be found in [21].

The OPF problem with respect to the feasible set (2.9) is given by

$$\min_{\mathbf{u} \in \mathcal{U}} \quad \sum_{v \in \mathcal{V}} C_v(p_v), \tag{2.10}$$

$$s.t. \quad \mathbf{s} = \mathbf{u} \circ (\mathbf{Y}\,\mathbf{u})^H. \tag{2.11}$$

Note that the active power p_v is set by the complex power s_v, which are determined by the constraints (2.11), via definition (2.5). This problem is NP hard despite the cost function C_v, due to the non-convex feasible voltage set $\mathcal{U}$.

In order to overcome this issue, a variety of convex relaxations of $\mathcal{U}$ have been proposed. The research focus [83, 124, 84, 82, 89] has usually been on the bus injection model, stated in (2.11). The complex power equation (2.11) and constraints (2.9) in the bus injection model can be equivalently rewritten as

$$\mathbf{s} = diag(\mathbf{u}\mathbf{u}^H \mathbf{Y}^H). \tag{2.12}$$

Here, the convexification relies on a change of variables from the voltages $\mathbf{u}$ to the matrix $\mathbf{W} = \mathbf{u}\mathbf{u}^H \succeq 0$ and relaxation of the resulting rank constraint, i.e. $\text{rank}\{\mathbf{W}\} = 1$, as proposed in [9, 83]. Neccesary and sufficent condition, when this relaxed problem solution is optimal are discussed in [84], e.g. if by accident the optimal matrix $\mathbf{W}$ has rank one.

[5] The admittance in this thesis is assumed to be in the per unit system, see Section 1.3.

[6] This thesis is going to be in the directional arrows framework resulting in this notation, while the opposite signs result for counter rotated directions, e.g. [103, Chapter 13].

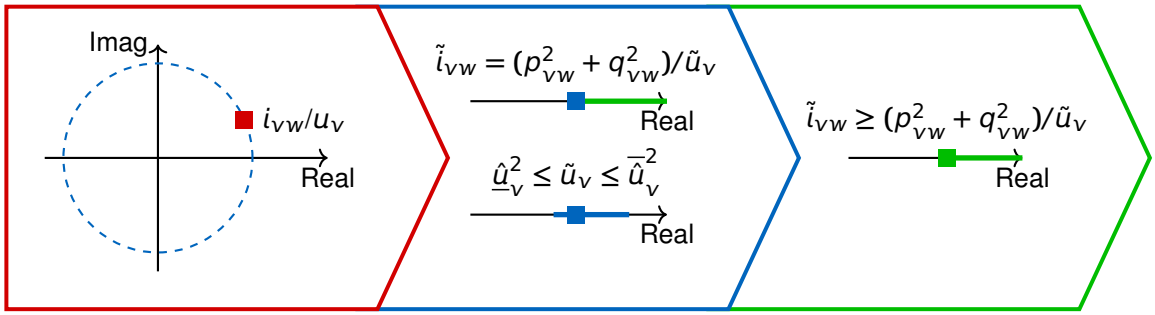


Figure 2.5: Two convexification steps, angle elimination (blue) and relaxation of the equality constraint (green), for the branch flow formulation.

In contrast to this relatively abstract procedure, an intuitive graphically explainable, equivalent convexification [21] can be derived from the branch flow formulation [45, 46]. This is based on quadratic approximation ideas as introduced in [53, 68, 69]. Instead of the directly obtained bus injections s_v by equation (2.11), flows

$$s_{vw} = \hat{u}_v {y_{vw}}^H (\hat{u}_v - \hat{u}_w e^{-j\theta_{vw}}),\ \{vw\} \in \mathcal{E}, \tag{2.13}$$

define the bus injections in the branch flow formulation [45, 46] to

$$s_v = \sum_{\{vw\}\in\mathcal{E}} s_{vw} - \sum_{\{wv\}\in\mathcal{E}} (s_{wv} - z_{vw}|i_{vw}|^2) + y_{v0}^H |u_v|^2,\ v \in \mathcal{V}, \tag{2.14}$$

where $z_{vw} = 1/y_{vw}$ is the impedance of the power line $\{vw\}$. Two steps turn this formulation into a convex problem, illustrated in Figure 2.5, first eliminating the angles $\tilde{i}_{vw} = |i_{vw}|^2$ and $\tilde{u}_v = |u_v|^2$, and then relaxing the non-convex equality constraint

$$\tilde{i}_{vw} = (p_{vw}^2 + q_{vw}^2)/\tilde{u}_v \qquad \text{to} \qquad \tilde{i}_{vw} \geq (p_{vw}^2 + q_{vw}^2)/\tilde{u}_v.$$

An optimum of the resulting second order cone programming is a solution to the OPF if the cost function is monotone rising in $\tilde{i}_{vw}$, since then the minimum will enforce the equality, and the angles can be reconstructed from $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{i}}$ [45]. Similar conditions, e.g. linear separability requiring non-binding p_v and q_v constraints at a minimum of one node for each line, guarantee exactness of convexified OPF solution in radial grids [90]. For meshed grids tunable phase-shifters need to be applied to obtain equivalent results [90, 46]. Nevertheless, these criteria are to some extend artificial as they do not arise from a straightforward implementation of real grids, e.g. [86].

Moreover, this second order cone programming is equivalent to the second order cone programming [20] arising from the bus injection semidefinite programming by an edgewise semidefiniteness inspection of $\mathbf{W}$ [124]. Similar approaches that exploit the sparse topology of electrical grids can be employed to circumvent a prohibitively large matrix in semidefinite programming. In [82], the problem is decomposed in line with the maximal cliques of a triangulated grid, i.e. the initial grid with additional edges such that

all circles are cordless [7]. Then [82] shows that the run-time is linear with respect to the number of such cliques. This concept is an analog to the Junction-Tree algorithm in the GM's community, see e.g. [130, 77].

Several other methods to decompose the OPF problem have been proposed recently. The robust augmented Lagrangian method of [74] coordinates the OPF computations in different electrical regions via Lagrange multipliers. The Alternating Direction Method of Multipliers (ADMM) proposes an elegant formulation for convex OPFs and can easily be implemented in a decentralized fashion via exchanging primal and dual variables between the electrical network nodes [79].

2.3 Optimal Volt-VAr Control

The Volt-VAr Control (VVC) problem in active grids adds an additional variable dimension, i.e. discrete variables τ for a number of tap changing transformers and capacitor banks, to the OPF problem in Section 2.2. Instead of an a priori given admittance matrix $\mathbf{Y}$, the discrete variables τ modify the admittance matrix $\mathbf{Y}(\tau)$. The optimal VVC problem determines the complex voltages $\mathbf{u}$ and discrete variables τ, such that the implied electricity generation and consumption maximize the overall welfare, minimize losses and guarantee regulatory bounds.

Transformers connect different circuits by electromagnetic induction. The AC flow of one circuit through a winding induces a current flow in the other circuit windings. The resulting current flow depends on the turn ratio, magnetizing current and non-load losses [103]. For now, neglecting magnetizing current and non-load losses, the transformer model in per unit system simplifies [8] to

$$\begin{pmatrix} i_v \\ i_w \end{pmatrix} = \begin{pmatrix} |\tau_{vw}|^2 y_{vw} & -\tau_{vw}{}^H y_{vw} \\ -\tau_{vw} y_{vw} & y_{vw} \end{pmatrix} \begin{pmatrix} u_v \\ u_w \end{pmatrix},$$

where τ_{vw} is the turn ratio. The neglected losses can be partially re-integrated by adding appropriate shunts y_{v0} and y_{w0}. This transformer model is common for most computational implementations [52].

On-Load Tap Changers (OLTCs) are transformers able to adjust their ratio under load. They have become more popular in recent years [47, 51, 92, 78, 59], since they offer a counter action for intraday voltage rises and drops caused by pattern variations from high feed-in to load in distribution grids. According to the current practice in Germany, the distribution system operator should ensure the voltage limits for two case scenarios and extend the distribution grid accordingly if the grid fails to fulfill the voltage limits [2, 1]. In this framework, [59] showed that OLTCs have the capability to ensure voltage limits within an existing infrastructure and thus are an alternative to grid extension. The investment considerations of network extension versus OLTC installation are examined in [92], with regard to financial, environmental, technological and social dimensions. Moreover, the adjustment to the current situation enables an operation in higher voltage regimes that

[7] A circle is cordless if there are exists no edges between nodes in a cycle that are not successors.

lead to lower losses. In the following, OLTCs are going to be modeled by a discrete set of feasible turn ratios $\mathcal{T}_{vw}$, denoted as taps $\tau_{vw} \in \mathcal{T}_{vw}$, in the transformer model.

Another way voltages can be modified and transmission losses reduced are capacitor banks. A capacitor is charged and discharged by AC flows, which draws reactive power. Capacitor banks can simply be seen as additional switchable shunt parts. Thus, they are modeled [8] by scaling y_{v0} according to the feasible switch status $\tau_{v0} \in \mathcal{T}_{v0}$ as follows

$$i_v = \tau_{v0}^2 y_{v0} u_v.$$

By drawing reactive power, the capacitor reduces the voltage drop along a line and thereby losses [119].

Let $\mathcal{E}^0 = \mathcal{E} \cup \{0\} \times \mathcal{V}$ be the edges $\mathcal{E}$ accomplished by reference/ground node 0 and nodal objective function $c_v(s_v, u_v)$ by the per unit complex voltage u_v. The nodal objective function c_v aggregates the production cost or negative utility and penalties for voltage band $[\underline{\hat{u}}_v, \overline{\hat{u}}_v]$ violations. A per unit admittance y_{vw} attributes to each edge $\{vw\} \in \mathcal{E}^0$ either standing for the line admittance if $v, w \in \mathcal{V}$ or shunt admittance if either v or w equals 0.

Transformers and capacitor banks form a discrete variable subset $\mathcal{D} \subset \mathcal{E}^0$. Each $\{vw\} \in \mathcal{D}$ is equipped with a turn ratio or capacitance variable τ_{vw} out of its attainable discrete position set $\mathcal{T}_{vw}$, where τ_{vw} is 1 if $u_v^n < u_w^n$ and otherwise any value in $\mathcal{T}_{vw}$. The admittance matrix $\mathbf{Y}(\tau) \in \mathbb{C}^{|\mathcal{V}|\times|\mathcal{V}|}$ depending on the discrete variables τ is given by its vw-entries, $v, w \in V$, as follows

$$\mathbf{Y}(\tau)_{vw} = \begin{cases} \sum\limits_{\{vu\}\in\mathcal{E}^0\setminus\mathcal{D}} y_{vu} + \sum\limits_{\{vk\}\in\mathcal{D}} \tau_{vk}^2 y_{vk} & v = w, \\ -y_{vw} & v \neq w,\ \{vw\} \notin \mathcal{D}, \\ -\tau_{vw}\tau_{wv}^H y_{vw} & v \neq w,\ \{vw\} \in \mathcal{D}, \\ 0 & else. \end{cases} \tag{2.15}$$

In this formulation, the complex powers are given by $\mathbf{s} = \mathbf{u} \odot (\mathbf{Y}(\tau)\,\mathbf{u})^H$. Thus, the active overall grid optimization problem reads as

$$\min_{\mathbf{u}\in\mathbb{C}^{|\mathcal{V}|},\tau\in\mathcal{T}} \quad \sum_{v\in\mathcal{V}} c_v(s_v, u_v), \tag{2.16}$$

$$s.t. \quad \mathbf{s} = \mathbf{u} \odot (\mathbf{Y}(\tau)\,\mathbf{u})^H, \tag{2.17}$$

$$s_v \in [\underline{s}_v, \overline{s}_v],\ v \in V. \tag{2.18}$$

Here the mix of continuous, i.e. $\mathbf{u} \in \mathbb{C}^{|\mathcal{V}|}$, and discrete, i.e. $\tau \in \mathcal{T}$, variables adds an additional complexity to the NP hard OPF problem (2.10) subject to (2.11).

In [10], this complexity was addressed by solving a relaxed problem, i.e. continuously modeled discrete capacitor bank variables and fixed transformer tap $\tau = 1$, and by re-obtaining the best initial solution by randomized permutations around the relaxed solution towards the closest discrete variables. This approach relies on the assumption

that transformer taps are close to their nominal value and line impedances are small [11]. A MILP formulation was proposed in [19] and applied in [111]. It is based on the MILP formulation for distribution grid configuration, i.e. to find the optimal switch constellation in distribution grids, in [18] for radial distribution grids. This formulation relies on the assumption of a fixed p_v and a small deviation of u_v from the slack node voltage. The temporal dynamic programming approach in [91] sets the tap position to the minimal objective cost one of the three possible taps – current, next lower and higher – for the next time step and approaches the capacitor banks analogously. This is systematic similar to the hill climbing approach in [6, 57], which alters tap positions until the objective function is no longer improved.

Automated control concepts [47, 51, 38, 37, 42, 80], i.e. concepts based on state estimations of the underlying grid based on a limited number of measurements, are another research focus. The concept proposed in [47] relies on a single measurement and adjusts for distributed generation along one feeder by an invariant DER correction, a promising method despite invariant adjustment [51]. Another source of influence on the voltage is reactive power feed-in $Q(\hat{u})$ by DER based on locally measured voltage magnitudes $\hat{u}$, e.g. evaluated in [38, 37] for different control approaches. A hybrid approach is taken by [42, 80], where automated control concepts for a single OLTC and multiple DER are analyzed with regard to the voltage quality improvements and engineering costs. The inference of a number of OLTC in series at different voltage levels was firstly addressed in [29].

2.4 Probabilistic Graphical Models

A probabilistic GM describes a family of multivariate probability distributions that share a common (conditional) independence structure via a graph. In this graph nodes are random variables and edges show the interrelation of two random variables. This graphical representation is used to develop efficient graph-based algorithms for various statistical inference computations, such as testing independence, marginalization or finding the most likely variable assignment. A detailed introduction to the theory and its applications is given in [77].

GMs are divided into undirected GMs and directed GMs according to the underlying graph edge structure, detailed in Section 1.3. This thesis focuses on undirected GMs since electrical grids relate to undirected GMs. For these undirected GMs, the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ represents all probability distributions $P(x_V)$ over random variables x_v, $v \in \mathcal{V}$, that factor as

$$P(x_V) = \frac{1}{Z} \prod_{C \in \mathbf{C}} \phi_C(x_C). \tag{2.19}$$

Here, x_A for set A denotes the set of all random variables x_v, $v \in A$. Moreover, $\mathbf{C}$ is the set of all cliques C, i.e. fully connected subsets, and the factors ϕ_C are non-negative potential functions over the variables x_C. Z denotes the necessary normalization factor to ensure the probability distributions owned property to sum/integrate to unity.

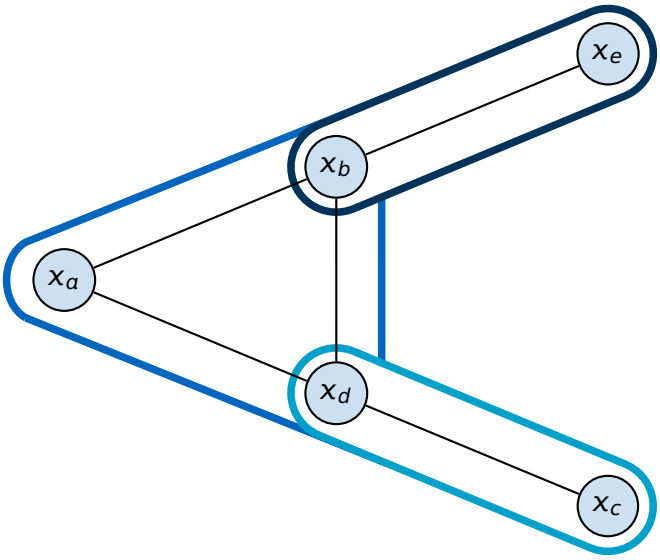


Figure 2.6: An undirected GM with five nodes $x_a, \ldots, x_e$ and three maximal cliques, $\{x_a, x_b, x_d\}$, $\{x_b, x_c\}$ and $\{x_d, x_e\}$.

Now focus on one of these computations that is the closest to the adaptation proposed in Chapter 3. Finding the maximum probability assignment, i.e. most likely variable assignment, means determining

$$x_V{}^* = \underset{x_V}{\operatorname{argmax}}\ P(x_V). \tag{2.20}$$

Factorization in line with the graph structure comes into play to reduce the complexity of this task. The independence structure related commutative property leads to decomposition into clique subproblems. A sample GM depicted in Figure 2.6 is used to illustrate this concept, before a general approach has been given.

The example GM depicted in Figure 2.6 contains three maximal cliques, namely $\{x_a, x_b, x_d\}$, $\{x_b, x_c\}$ and $\{x_d, x_e\}$. Computing the maximum probability can be reformulated to

$$\begin{aligned} \max_{x_a,\ldots,x_e} \quad & \frac{1}{Z}\phi_{\{abd\}}(x_a, x_b, x_d)\phi_{\{bc\}}(x_b, x_c)\phi_{\{ed\}}(x_e, x_d) \\ = \quad & \max_{x_a,x_b,x_d} \frac{1}{Z}\phi_{\{abd\}}(x_a, x_b, x_d) \underbrace{\left(\max_{x_c} \phi_{\{bc\}}(x_b, x_c)\right)}_{=m_{\{cb\}\{abd\}}(x_b)} \underbrace{\left(\max_{x_e} \phi_{\{ed\}}(x_e, x_d)\right)}_{=m_{\{ed\}\{abd\}}(x_d)}. \end{aligned} \tag{2.21}$$

The two variables x_c/x_e are solely dependent on x_b/x_d and thus its factor maximization can be computed independently of the other variables when conditioned on x_b/x_d. These conditional optimization subproblem results can then be integrated into the overall problem by function-valued messages $m_{\{cb\}\{abd\}}(x_b)$, $m_{\{ed\}\{abd\}}(x_d)$. This shows that the maximization problem over five variables can be split into two optimizations over two variables and one over three by exploiting the commutative property of conditional independence.

Figure 2.7: Computational complexity of the maximum probability assignment for the GM, Figure 2.6, with binary variables $x_a, \ldots, x_e$ by full enumeration of the first line (2.21) grey square and reformulated second line (2.21) blue squares.

For this example Figure 2.6 and binary variables $x_a, \ldots, x_e$, full enumeration of the formula (2.21) in the first line requires 2^5 computations, equivalent to the grey square area in Figure 2.7, while the factorization into subproblems (second line) leads to $2 \cdot 2^2 + 2^3$ computations, equivalent to blue squares in Figure 2.7. In general the reduction is exponential in the problem size for discrete variables x_v and qualitatively similar results hold for continuous variables. Thus, this factorization potentially amounts to a tremendous reduction in computation time – especially if this trick is repeated for larger graphs.

The example above shows the fundamental idea behind the general algorithm, known as MP, to solve the maximum probability assignment in tree-structured clique graphs. In the first stage, the messages $m_{ij}(x_{\mathsf{S}_{ij}})$ from clique C_i to clique C_j are computed for every pair of overlapping cliques C_i and C_j, i.e. cliques where the separator set $\mathsf{S}_{ij} = \mathsf{C}_i \cap \mathsf{C}_j$ is not empty. For each separator set S_{ij} the message is computed as

$$m_{ji}(x_{\mathsf{S}_{ij}}) = \max_{x_{\mathsf{C}_j \setminus \mathsf{S}_{ij}}} \phi_{\mathsf{C}_j}(x_{\mathsf{C}_j}) \prod_{k \neq i} m_{kj}(x_{\mathsf{S}_{kj}}). \tag{2.22}$$

The optimal values for variables contained in one of the separator sets S_{ij} are then given by

$$x^*_{\mathsf{S}_{ij}} = \operatorname*{argmax}_{x_{\mathsf{S}_{ij}}} m_{ij}(x_{\mathsf{S}_{ij}}) m_{ji}(x_{\mathsf{S}_{ij}}). \tag{2.23}$$

Computing the optimal values for variables not contained in one of the S_{ij} is described in [77, Chapter 10].

The MP starts at a leaf clique, where the maximization problem (2.22) reduces to a maximization over one potential function ϕ_{C} only. Then continues to calculate the message for separator sets S_{ij} where all messages $m_{ki}(x_{\mathsf{S}_{ki}})$, $k \neq j$, from all neighboring

cliques except the target clique are known until all messages are known. The idea how a reordering in line with the commutative property can be computationally derived by messages was proposed by [107]. The formulation backing was based on GMs without loops since then a feasible ordering can always be found, e.g. by depth-first-search going from the leafs to the root and back.

However, the commutative property that makes it possible to move the maximum operation exploited in for the reordering in equation (2.21) holds only for clique trees with junction tree property, i.e. if for any clique C_1 and C_2 all cliques on the unique path of the clique tree joining them contain their separator set S_{12} [77], also called junction tree. In clique graphs with at least one circle, the interdependence of the variables contravenes a reordering. In this case, either a clique tree needs to be constructed by enlarging the cliques such that the new cliques have tree shape [77], or approximate methods [100], called loopy MP, have to be applied for this task. The results and methods to obtain a solution in the first approach are equal to the approaches for GMs that initially have a junction tree if the enlarged clique tree is a junction tree.

For loopy MP all messages m_{ij} are initialized with a value, e.g. 1, and are iteratively updated by variations of equation (2.22). Thus, loopy MP messages m_{ij} are initially wrong and include the same potential functions multiple times, both is opposed to the optimality of the result. Still, loopy MP performs incredibly well in praxis, e.g. for error decoding [48], and for some GM topologies and potential functions the theoretic correctness can even be derived. For Gaussian GM, i.e. GM with Gaussian potential functions, Weiss and Freeman [134] showed that if MP converges, it will result in the correct posteriors; qualitatively similar results have been shown in [131, 114]. The $-\log$ relates the maximization problem (2.23) for Gaussians to unconstrained Quadratic Programming (QP) minimization, for which [97] proved that the equivalent MP algorithm solution is optimal, given the existence of an appropriate quadratic approximating sequence. For discrete variables, quotient metric arguments can derive sufficient conditions for convergence of loopy MP [99, 67]. An intuitive way to understand these results is that the initial bias in the messages is marginalized by multiple propagation and normalization $1/Z$.

This implied error correction cannot be directly transferred to the application in Chapter 3, since the functions C_v do not need to sum/integrate to an a-priory known value, particularly not unity. Nevertheless, in radial grids' GM representations that are described in Section 3.2, cliques directly correspond to nodes in the electrical graph. Thus, the representations' clique graph is of tree shape and none such iterative procedure is required. But for multiple time-steps, each step and node would add a mesh with each neighboring node. This is in contradiction to the clique tree property, which is why the work presented here focuses on a single time-step dispatch.

For continuous random variables computing the message functions is in many cases intractable. Therefor, approximate inference algorithms for GMs then try to approximate those messages with standard function classes, e.g. [130] for an introduction. Projecting message updates onto a given class of functions in each iteration has been done, e.g. in [102].

Using probabilistic GM methods in power grids has been proposed before for state estimation applications [63, 87]. In [139], a one-to-one matching of customers and generators is computed using GM algorithms. Nevertheless, the MP techniques, proposed in Section 3.2, for solving the ED problem is novel, foremost to the specially adapted approximated inference application in Section 3.5.

2.5 Dynamic Programming

Dynamic programming is an optimization approach that splits a complex large problem up into a sequence of smaller subproblems interconnected by common variables. On these common variables defined linking terms, i.e. the cost-to-go functions J and reduced decision spaces Y, are used to coordinate the subproblems. The cost-to-go functions J aggregate the optimal objective value of the subsequent subproblems given the common variables. The reduced decision spaces Y capture the interdependence of the subproblem conditions. The key property for dynamic programming is that the subproblems' optimal solutions need to assemble the overall solution [58]. Thus, the optimization technique required to solve these subproblems is not a determining factor for dynamic programming, but the subproblems trace back towards optimality.

The above concept is the common ground for a number of different views on dynamic programming taken by computer scientists, economists, and engineers. In computer science the crucial point is the efficient use of storage and non-iterative propagation [36], based on the function representation J. However, the cost-to-go functions J of continuous variables take in generally infinitely many values that can neither be fully computed nor be stored, except if J coincides with an element of a finite-dimensional function class or can be well approximated by such[8]. Approximation approaches for the cost-to-go function J are qualitatively similar to the approximated messages in GM, discussed in Section 2.4. Economists and engineers often reduce dynamic programming to a tool to handle multiple time-period decision processes, which can either be in a deterministic or uncertain environment [17]. The common approach for deterministic processes is forward processing [109] whereas uncertain processes can better be approached in a backwards manner, e.g. [30]. The focus on multiple time period decision processes arises from a classic temporal interpretation of Bellman's principle of optimality, which states that each single decision has to be optimal with regard to the resulting state from the initial decisions [14][9]. However, Bellman's optimality principle allows for a broader concept of dynamic programming, e.g. [123, Fundamentals], and is detailed in the following.

The dynamic programming concept [123, Fundamentals] works as follows: (1) restructure the problem into subproblems, (2) obtain the link between those subproblems,

[8] In the machine learning community obtaining such function class combinations is known as deep learning.

[9] "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decisions." [14]

(3) solve the subproblems and (4) recover the solution of the initial problem from the subproblems. In this more general view the optimality principle of Bellman reads as:

Principle 2.5.1. *The optimal solution of the subproblem conditioned on its linking terms must contribute to the optimal overall solution given whatever arises from their common variables.*

The general formulation of dynamic programming for a minimization problem at one subproblem $l \in \mathcal{L}$ reads as

$$J_l(x) = \min_{\mathbf{y}} \quad F(x, \mathbf{y}) + \sum_{k \in \mathcal{K}(l)} J_k(y_k), \tag{2.24}$$

$$s.t. \quad \mathbf{y} \in Y(x), \tag{2.25}$$

where $Y(x)$ is the decision space and $\mathcal{K}(l)$ are subproblems with link to l and their optimal solutions are included in the cost-to-go function J_k, $k \in \mathcal{K}(l)$. In the temporal view l is the time t and $\mathcal{K}(l)$ reduces to $\{t+1\}$.

Restructuring a problem into smaller subproblems is not always advantageous. Often there is a trade-off between simplicity of extremely small problems and the mere number of such problems to solve. The maps J and Y hint at the resulting complexity reduction of this reformulation [123, Fundamentals]. An additional point is the complexity of the subproblems in relation to the initial problem, e.g. an optimization problem with multiple discrete decisions scales exponentially, in contrast to one with a single discrete variable.

An intuitive way to find subproblems satisfying the optimality Principle 2.5.1 is to exploit the commutative property of conditionally independent subproblems, similar to the example in Section 2.4. From a dynamic programming perspective, GMs are a special version of dynamic programming with extremely small subproblems and the message m_{vw} equals the cost-to-go J.

Using dynamic programming to solve electrical engineering problems has been standard for a long time, e.g. UC [62, 104, 106] and for VVC [91, 41]. For the UC problem, the on-off decision is the relationship between the per time point subproblems. The tap position evolution over time of an OLTC takes this role in [91]. The broader view discussed above was taken in [41], where the location, number and size of capacitors in radial grids is determined by subproblems in line with the grid implied independence. However, the author is not aware of literature using dynamic programming to enable an optimal VVC solution for a multitude of interfering discrete components, e.g. OLTC and capacitor banks, as proposed in Chapter 4.

Chapter 3

Economic Dispatch as Graph Problem

In this Chapter the dispatch problems, introduced in Section 2.1 and Section 2.2, are viewed from the Graphical Model (GM) angle and are solved utilizing GM methods. The focus is on the Economic Dispatch (ED) problem that is to find the active power consumption and production such that the sum of production cost and negative utility is minimized in the grid $\mathcal{G}(\mathcal{V}, \mathcal{E})$. This should be done subject to line capacities and feed-in restrictions.

In the past, the ED problem was a problem for few utilities. They had to operate their portfolio of large centrally placed power plants optimally to supply demand. Yet, the widespread introduction of new technologies has been changing this. Renewable energies, e.g. Photovoltaic (PV), wind and biomass, are typically small in size and spread throughout the distribution grids. Moreover, heater, Electric Vehicle (EV) or some industrial processes are flexible with regard to their consumption. Thus, a large number of Distributed Energy Resources (DER) with a diverse owner structure need to be dispatched in distribution grids. These DER are changing the requirements on ED methods to be scalable, quickly adaptable, incorporate grid limits, preserve privacy and have minimal communication effort.

The vast number of to control units is one of the challenges for ED in this environment. Another challenge is the complex interdependence structure of the DER due to the grid topology as well as the DER implied limits. GM methods are known to be able to handle a large number of interrelated variables efficiently, e.g. the number of pixels in a picture for pattern recognition. However, the interdependences for this application are local but the energy conservation (2.2) is global, in the sense that all variables are coupled. Thus, the ED is equivalently reformulated to a local problem by the nodal balances. The resulting local problems are then solved independently and coordinated by messages, in analogy to Message Passing (MP) in Section 2.4.

For this the problem formulation (2.1) subject to (2.2) is transferred to a graph formulation, Section 3.1. Then MP method, introduced in Section 2.4, is reformulated to solve this problem, Section 3.2. Section 3.3 discusses the properties this method inherits, before the discrete approach and continuous approach are detailed, Section 3.4 and Section 3.5.

3.1 Graph Problem Formulation

The underlying idea for the graph formulation is the naturally implied decomposition for local problems. To obtain a local problem, the global problem (2.1) subject to (2.2) is equivalently reformulated in terms of local active power flows p_{vw}. This section is based on [73].

In the problem (2.1) subject to (2.2), the power balance (2.2) is the global constraint, it couples all p_v decision variables and hinders a simple decomposition of the problem. Equivalent to (2.2), however, is a set of power flow equations, e.g.,

$$\sum_{w \in \mathcal{N}(v)} p_{vw} = p_v, \tag{3.1}$$

where p_{vw} is the power transport from v to w and $p_{vw} = -p_{vw}$. If those equations are satisfied at each node, no energy is lost in the network and the overall power balance constraint (2.2) is satisfied.

Then integrating (3.1) in the problem (2.1) subject to (2.2), the problem reads as

$$\min_{p_{vw} \in \mathcal{P}_{vw}} \quad \sum_{v \in \mathcal{V}} C_v\left(\sum_{w \in \mathcal{N}(v)} p_{vw}\right), \tag{3.2}$$

$$s.t. \quad \sum_{w \in \mathcal{N}(v)} p_{vw} \in \mathcal{P}_v, \tag{3.3}$$

$$p_{vw} = -p_{wv}. \tag{3.4}$$

Like in Section 2.1, $C_v(p_v)$ is the cost or negative utility function associated with the active power p_v and $\mathcal{P}_v$, $\mathcal{P}_{vw}$ the feasibility sets for the active power p_v and the active power flow p_{vw}, respectively. This formulation corresponds to a simple lossless power flow model and is equivalent to the common DC-approximation of the Alternating Current (AC) power flow equations for radial grids, see Section 2.1.

Note that using formulation (3.1), the optimization problem (3.2), subject to (3.3) and (3.4) is local in the sense that each term of the objective and the constraints only contains variables from a neighborhood of v in the graph. Such a graph problem formulation of the ED problem can be tackled by local operations on the graph, trying to find optimal local assignments while guaranteeing local consistency between neighboring sites that share an interrelating active power flow p_{vw} variables.

In some sense, that is exactly what GM Message Passing (MP) method presented in Section 3.2 and detailed for discrete variables in Section 3.4 and for continuous variables in Section 3.5 does. The rewriting of (2.2) into (3.1) is, while almost trivial, important for the proposed solution strategy.

3.2 Message Passing Methodology

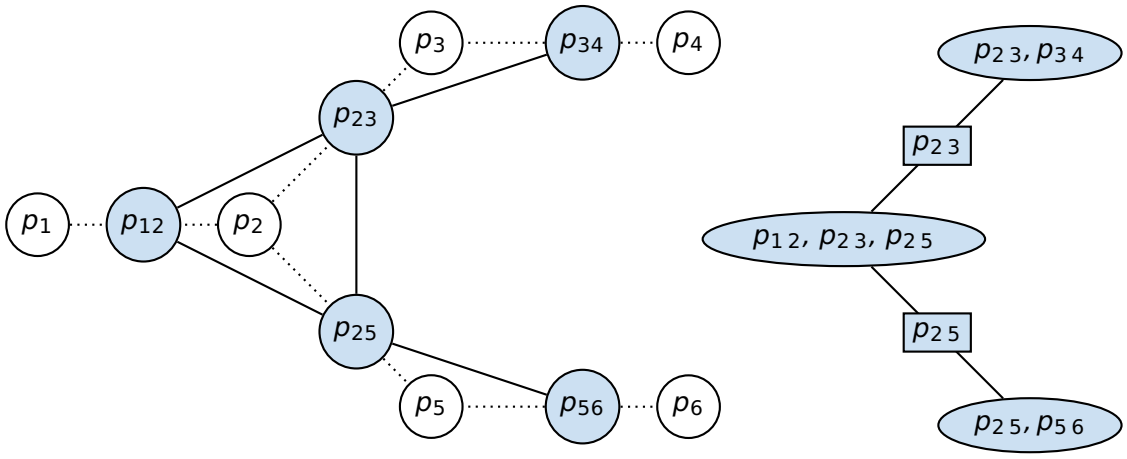
In this section is shown how the GM framework can be applied for ED problems. Matching the GM notation introduced in Section 2.4 and using the graph formulation of the ED problem (3.2), subject to (3.3) and (3.4), Section 3.1, a GM inspired solution methodology is presented. This idea was firstly proposed by the author in [73] and this section is based on that.

At first note that due to monotonicity $\operatorname{argmax} \prod_c \phi_c = \operatorname{argmin} \sum_c -\log \phi_c$ and the graph ED problem (3.2) subject to constraint (3.3) can be re-written as

$$\min_{p_{vw} \in \mathcal{P}_{vw}} \sum_{v \in \mathcal{V}} C_v \left(\sum_{w \in \mathcal{N}(v)} p_{vw} \right), \tag{3.5}$$

where the nodal active power domain constraints $\sum_{w \in \mathcal{N}(v)} p_{vw} = p_v \in \mathcal{P}_v$ have been included in the cost function $C_v(p_v)$, by setting it to infinity if $p_v \notin \mathcal{P}_v$.

After this is set an undirected GM is constructed from the electric network as demonstrated in Figure 3.1. A node is created for each transport variable p_{vw} and edges are added between these nodes, if and only if the corresponding electric lines connect to the same electric bus. Buses in the electric network then matching the cliques in a GM and the clique potentials ϕ_C are defined by identifying the cost function of the electrical node, $C_v\left(\sum_{w \neq v} p_{vw}\right)$, with the negative logarithm of the clique's potential function in the GM $-\log \phi_C$. Since the electrical nodes directly translate to cliques, the electrical network topology is equivalent to the clique graph. Thus, radial electrical networks lead to clique trees, Figure 3.1.



(a) Electrical Grid (b) Clique Graph

Figure 3.1: An electrical grid with six power nodes $p_1, \ldots, p_6$ and five power lines (dotted, solid) with power transport variables $p_{12}, \ldots, p_{56}$ (solid lines, filled circles) that corresponds to the five nodes GM shown in Figure 2.6. The clique graph implied by the cost functions, where the cliques are depict oval and separator set, i.e. intersection of two adjacent cliques, as rectangle.

With these findings and connections the following statement holds:

Theorem 3.2.1 ([73])**.** *The maximum probability assignment on the constructed undirected probabilistic GM is equivalent to the economic dispatch problem in the original electric network.*

This equivalence implies that all methods for the decoding of the maximum probability in probabilistic GMs can also be applied to ED. For the introduced MP the message computation (2.22) becomes

$$m_{vw}(p_{vw}) = \min_{p_{kv} \in \mathcal{P}_{kv}\, k \in \mathcal{N}(v)\setminus\{w\}} C_v\left(\sum_{l \in \mathcal{N}(v)} p_{kv}\right) + \sum_{k \in \mathcal{N}(v)\setminus\{w\}} m_{kv}(p_{kv}), \tag{3.6}$$

and (2.23) then reads

$$p_{vw}{}^* = \operatorname*{argmin}_{p_{vw}} m_{vw}(p_{vw}) + m_{wv}(-p_{vw}). \tag{3.7}$$

Thus, using the local power flow equations (3.1) in replacement of the global constraint (2.2) naturally impose the line limits on power transport, e.g., $-\overline{p}_{vw} \leq p_{vw} \leq \overline{p}_{vw}$ where $\overline{p}_{vw}$ is the line capacity.

The notation above hints at directly using the active power transport p_{vw} as optimization variables. However, the full AC power flow equations (2.13) can also be employed. Then the voltage magnitudes $\hat{u}_v$ and phase angle differences θ_{vw} are the optimization variables. Here, phase angle differences $\theta_{vw} = -\theta_{wv}$ take the active power transport p_{vw} role in equation (3.1). Thus, such a formulation preserves the locality property as defined above and would be amenable to the solution technique shown in this section. Only the number of real variables in each neighborhood would increase when using the three variables θ_{vw}, $\hat{u}_v$ and $\hat{u}_w$ defining the complex power flow s_{vw} instead of one active power transport variable p_{vw} per line. Using the full AC power flow equations (2.13) then leads to the following messages

$$m_{vw}(\hat{u}_v, \hat{u}_w, \theta_{vw}) = \tilde{m}_{vw}(s_{vw}(\hat{u}_v, \hat{u}_w, \theta_{vw})). \tag{3.8}$$

Following this message formulation, the here proposed divide and conquer trick (2.21) could still be used. However, these multi-dimensional messages are more complex to obtain and to store so an especially suited approach is needed.

The implementation of this methodology in two fashions is discussed in Section 3.4 and Section 3.5. In Section 3.4 messages are represented via discretization of the active power flow optimization variables p_{vw}, denoted as Discrete Message Passing (DMP). This allows for an exact inference on tree-structured distribution grids and for arbitrary cost functions C_v. However, for variables of significantly different scales a common fine discretization becomes computationally infeasible. Thus, a continuous approach might be more suitable. An such approach is proposed in Section 3.5, called Continuous Message Passing (CMP). The approach utilizes the optimality principles to efficiently obtain the exact messages and a functional approximation to represent them.

3.3 Properties

The equivalence of the ED problem to the graph problem (3.2) subject to (3.3) and (3.4), Section 3.1, that is solvable by the MP messages (3.6), Section 3.2, has several important consequences for ED in smart distribution networks. They concern the in linear run-time obtained optimal ED result, Section 3.3.1, the inherited information privacy principles, Section 3.3.2, and the additional local cost knowledge, Section 3.3.3.

3.3.1 Decomposition

The decomposition in line with the commutative property of conditional independent subproblems and their coordination via messages, described in Section 3.2, leads to optimal results. Moreover, the MP method obtains the result in linear time, independent of the ED cost function shapes but dependent on the grid topology.

MP is known to find the maximum probability assignments on trees. For discrete variables this holds also for the computational calculation of the maximum probability assignments. An implementation for continuous variables require the message to be of a finite-dimensional function class or well approximated by a such, in order to make that hold. Similar results hold for the broader class of clique trees with junction tree property, i.e. for each two cliques with common variables these variables need to be part of each separator set on the path from one to the other. By the generation of the grid respective GM, the cliques directly relate to nodes in the grid. Applying this algorithm to ED will result in the globally optimal assignment, if the electrical network is radial since the respective clique graph is a tree and satisfies by its generation, detailed in Section 3.2, the junction tree property, i.e. junction tree. Unlike most optimization algorithms, the cost functions and domain constraints may take any shape, at least for the discrete variable implementation. But the grid's topology is the determining factor, since a junction tree is only directly implied for radial grids.

Low and medium voltage networks are in practice often operated in a tree-shaped configuration [110], i.e. radial. Low Voltage (LV) feeders in rural regions tend to be composed of several strings connected to one transformer. Here meshed LV feeders rarely exist because of the operation complexity and higher required investments. Unlike, densely populated suburban and urban regions low voltage feeders are often built as rings and feed by two transformers, similar holds for medium voltage. So customers can be supplied from both sides in the case of line faults. But during normal operation, which is of major interest to ED, an opened breaker typically splits the loops.

The run-time of this methodology is basically dependent on the number of messages and their computation. Note that the number of messages is exactly twice the number of separator sets between cliques, which are in this case equal to buses. Assuming for now discrete active power flow variables p_{vw} and let the maximum number of such discrete values that p_{vw} can take be k. Moreover, let n be the number of electrical buses in the system and d denote the maximal clique size of the GM, which is equal to the degree of the electrical network. Then, for each message at most k^d values have to

be calculated/compared for the minimum operation (3.6), and two such messages have to be calculated for each of the n buses. Thus, the total run-time of the MP method is $O(2nk^d)$ by using full search for the optimizations in (3.6). Similar, however more complex and approach dependent, relations hold for continuous flow variables p_{vw}. Thus, the total run-time scales linearly in the size of the network, which renders the proposed methodology suitable for very large networks. Besides, for discrete variables the run-time is also precisely predictable and independent of the shape of cost functions and constraints. Similiar, holds just approximately for continuous variables.

The argument above showed that the degree of the electrical buses d has a significant influence on the computational complexity. Since for high-degree nodes in the electrical network the minimum calculation in the message optimization problem (3.6) will then extend over many dimensions. This exponential scaling factor k^d is often referred to as curse of dimensionality. However, a reduction of this factor k^d for high-degree nodes in the electrical network to at most k^3 can be achieved by re-writing the electrical network without changing its physical meaning. Therefore, each node with a degree larger than three is iteratively split into two new nodes that are connected with an appropriate capacity and lossless line. The neighborhood connections are distributed between the two new nodes and thus the original degree of the node is reduced to roughly one half. This procedure is repeated until all nodes in the graph have a maximum degree of three, i.e. turning the message calculation (3.6) into a two-dimensional problem. The price of this reformulation is a few additional edges and hence message computations. However, the computational effort is constant per edge, whereas the reduction of the neighborhood dimension leads to an exponential computation cost decrease.

Note that the local approach Section 3.2 is just consistent if the minimizer in each assignment calculation (3.7) is unique. The reason is that the individual optimal active power flows $p_{vw}*$ are computed locally and independently from each other so a non-consistent selection of one of the solution sets may result for non-unique minimizers, which lead to non-optimal or even infeasible node injections p_v. An example for this is a chain of three nodes where each of the end nodes could supply a load in the middle at identical cost. Independent decisions on each of the two edges could then lead to the situation where the load in the middle is covered twice or not at all. It is recommended to break such potential symmetries to ensure solution uniqueness, e.g. via adding small random terms onto otherwise identical cost functions.

3.3.2 Information Hiding

To preserve the component models local is the key for the information hiding principle of the proposed MP method. This privacy preserving property is detailed based on the two systematically different cases, the self-containded message from a leaf node and the aggregated message from a non-leaf node.

Solving ED via the MP method means that the component models do not have to be exchanged if the algorithm is implemented in a distributed fashion. The only exchanged information is the messages between neighboring nodes. The individual cost functions

typically cannot be reconstructed from these messages, since the messages contain highly aggregated information from several users. This built-in information hiding principle is helpful for competitive, multi-owner environments, where the true cost structure of a participant may comprise a privacy issue.

To what degree it is possible to reconstruct the individual cost functions from the transmitted messages strongly depends on two attributes: a) the prior knowledge of the shape of the true cost functions and b) the number and position of messages that are used for reconstruction. A complete characterization under which conditions which knowledge can be obtained goes beyond the scope of this thesis. Here, only two systematically different situations are discussed – leaf node messages and non-leaf node messages.

First, the message from a leaf node is the cost function of the leaf node itself

$$m_{vw}(p_{vw}) = C_v(p_{vw}), p_{vw} \in \mathcal{P}_{vw} \cap \mathcal{P}_v,$$

but C_v includes also the current demand δ_v. For assumed quadratic cost functions, the message

$$C_v(p_{vw}) = \alpha_v {p_{vw}}^2 + \beta_v p_{vw} + \gamma_v,$$

for no local demand, i.e. $\delta_v = 0$, results in exactly the same message m_{vw} as a cost function with coefficients

$$\hat{\alpha}_v = \alpha_v, \hat{\beta}_v = 1/(1 + 2\delta_v)\beta_v \text{ and } \hat{\gamma}_v = \gamma_v - {\delta_v}^2\alpha_v - \delta_v/(1 + 2\delta_v)\beta_v$$

for some demand δ_v. Thus, already in this extremely simple case the cost function can not be exactly reconstructed.

For message up the tree-hierarchy, i.e. a message from a non-leaf node, the message is already a mixture of at least two cost functions. The minimum operator in the message computation (3.6) is not one-to-one, meaning that only the lowest cost for each value p_{vw} is visible in the message and the information about the higher "bids" is lost. Knowing the potential shapes of cost functions up to a scaling factor, one can eventually estimate the individual, single cost functions from this partial information. But the assignment of these function to a particular node, however, would still not be given. In a realistic scenario with limited prior knowledge one can hence assume that from the messages of a larger subgraph no exact reconstruction of each node's contribution is possible.

3.3.3 Local Influence Knowledge

The messages to a node equip the node with a knowledge of their influence on the overall problem. This nodal insight is similar to the local knowledge in dynamic programming and is systematically different to marginal knowledge supplied by other approaches, e.g. linear programming.

The proposed MP framework applied to a string like graph is equivalent to a temporal dynamic programming approach, introduced in Section 2.5, over several time steps for one generator. Here the nodes on the string correspond to the time steps, and the messages to the cost-to-go function. Not time steps are links but lines between electrical nodes for

MP approach hence MP can also be applied in none string topologies. The proposed framework thus forms an extension of the temporal dynamic programming approach to ED for tree-like graphs. Moreover, the MP nodal decomposition is an extreme case for the general dynamic programming approach, as discussed in Section 2.5.

This observation has an important practical consequence, having received all messages from the neighborhood each node can locally deduce its influence on the global problem. For any finite step-size local change, the minimum of the incoming messages and local cost function gives the global cost when all other variable feed-in and demands stay constant. In contrast the optimal dual variables of the local power flow equation for linear programming and Mixed Integer Linear Programming (MILP) return only the distinctive marginal costs, i.e. the total cost changes for infinitesimal changes in local demand.

This result can be understood via the similarity between MP, Section 2.4, and dynamic programming, Section 2.5, discussed above. From a dynamic programming view, messages m_{vw} are the coordinating cost-to-go functions J_l. The values $m_{vw}(p_{vw})$ thus express the minimum cost-to-go that would be accumulated in the subtree behind edge vw conditioned on the active power flow p_{vw} on that line. This viewpoint also explains the formula (3.7), where the minimum cost of the left half-tree is added to the minimum cost of the right half-tree in order to obtain the globally optimal decision. This is systematically equal to the forward-backward pass in dynamic programming. Interpreting the messages as minimal cost of active power flow p_{vw}, the finite step-size marginal costs can be computed by deciding locally at each node for each additional demanded unit from which source it should be taken, either via transport from one of the neighbor sub-networks or from a local production. For this computation only the messages from the neighbors and the node's own cost function are needed but it results the true global finite step-size cost-changes of the overall solution.

3.4 Discrete Message Passing (DMP)

In this section the MP method of Section 3.2 is introduced for discrete power flow p_{vw} variables, $p_{vw} \in \mathcal{P}_{vw}$. Then, the messages m_{vw} take values for each entry in discrete set $\mathcal{P}_{vw}$ that are stored in a vector and straightforwardly computed by index operations. This approach leads to optimal results independent of the function shape C_v and the feasible set $\mathcal{P}_v$, only dependent on the discrete flow sets $\mathcal{P}_{vw}$. This flexibility is valuable for a multitude of diverse convex and non-convex component models. The method is denoted in this thesis as the Discrete Message Passing (DMP) algorithm.

In this discrete setting, Section 3.4.1 describes the Message Passing (MP) algorithm and the local message calculations. In Section 3.4.2 the DMP approach is compared to the central MILP with regards to the result and run-time. Finally, a conclusion for DMP is drawn in Section 3.4.3.

3.4.1 Implementation

Section 3.2 derived the theoretic MP formulation for ED not mentioning the computational implementation. Now, this is done under the assumption that the variables p_{vw} on lines $\{vw\} \in \mathcal{E}$ take values in the discrete set $\mathcal{P}_{vw}$. Then the messages m_{vw} are vectors of size $|\mathcal{P}_{vw}|$ resulting straightforwardly from index operations.

For a transport p_{vw} in the discrete set $\mathcal{P}_{vw}$ the messages (3.6) read as

$$m_{vw}(p_{vw}) = \min_{\mathbf{p} \in \bigtimes_{k \in \mathcal{N}(v)\setminus\{w\}} \mathcal{P}_{kv}} C_v\left(p_{vw} + \sum_{k \in \mathcal{N}(v)\setminus\{w\}} p_{kv}\right) + \sum_{k \in \mathcal{N}(v)\setminus\{w\}} m_{kv}(p_{kv}), \quad (3.9)$$

where $\mathbf{p}$ is a $|\mathcal{N}(v)|-1$ dimensional vector referencing by p_{kv} an entry of the discrete set $\mathcal{P}_{kv}$, $k \in \mathcal{N}(v)\setminus\{w\}$. For $\mathcal{N}(v)\setminus\{w\} = \{k_1, k_2\}$ this is the minimum of a $|\mathcal{P}_{k_1v}| x |\mathcal{P}_{k_2v}|$ matrix, given by the local costs C_v of all possible combinations of $p_{k_iv} \in \mathcal{P}_{k_iv}$, $i \in \{1, 2\}$, plus the respective message $m_{k_iv}(p_{k_iv})$, $i \in \{1, 2\}$, illustrated in Figure 3.2.

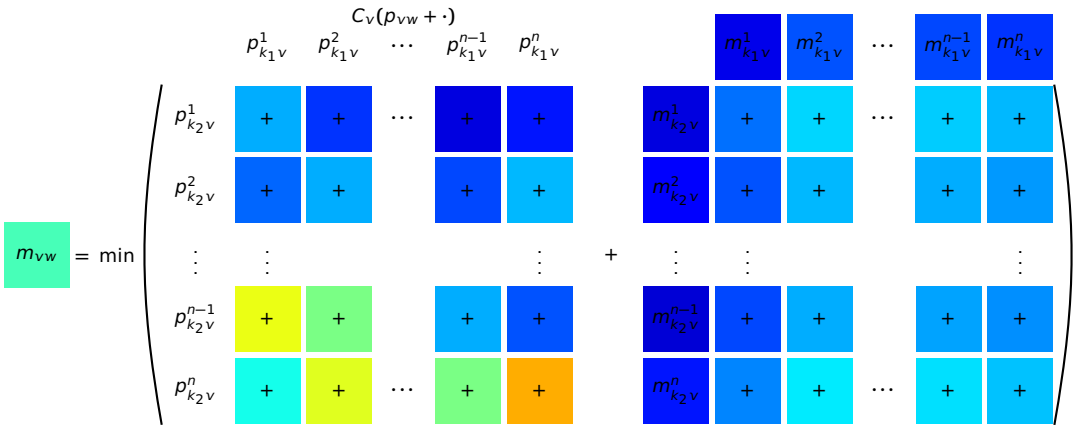


Figure 3.2: Illustration of the index operation to solve (3.9) when $\mathcal{N}(v) \setminus \{w\} = \{k_1, k_2\}$. The colors represent the function values in range from blue low to red high.

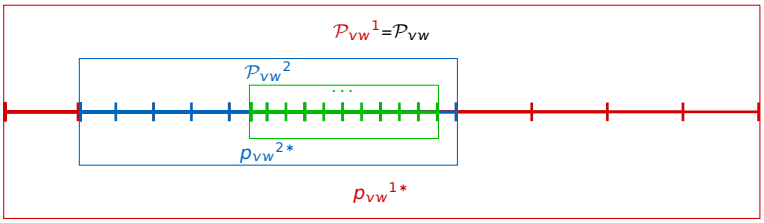


Figure 3.3: Propagation of the discrete sets $\mathcal{P}_{vw}{}^i$ for three iterations, $i \in 1, 2, 3$, of the proposed iterative procedure when the capacities are of strongly different sizes, first red, second blue and third green.

The UGM toolbox [117], which name was derived from the GMs it handles Undirected Graphical Models, implementation works exactly in this fashion. It allows for maximal two variables in one potential, called pairwise potentials. The degree of a node in the electrical grid corresponds to their potential/cost function scope, see Figure 3.1. Thus, the potential representation has be separated into auxiliary nodes aggregating two nodes until the initial potential spans only two nodes, where at least one is an auxiliary node. This strategy, explained in [133, Appendix], shows that every GM has a pairwise equivalent. Thus, the restriction to maximal pairwise potentials do not limit the solvable grid topologies.

However, working with non-pairwise potentials, i.e. higher dimensional, is in general possible and has been implemented. But the results were equal and the run-time not significant lower. Thus, for matters of comparability, a slightly modified standard algorithm of the UGM toolbox [117] was used in Section 3.4.2. Since the only interest is the costs, i.e. log probability functions, the max-product implementation (2.22) was replaced with the min-sum equations (3.9) for improved numerical stability.

In Section 3.4.2 first example, the line capacities are all equal and their size is on the same order of magnitude as the generators' and consumers' power intake. Thus an equal and equidistant discretization is taken on all lines. In the second example, the line capacities are very different from each other and relatively large in comparison to the loads and generators. In this case an iterative procedure is applied. A fixed number of discretization points is used for each line and the iteration is started with a large line-capacity dependent discretization step-size at first, red marks in Figure 3.3. Then a provisional solution, i.e. red $p_{vw}{}^{1*}$ in Figure 3.3, with one run of DMP is determined where the power balance at each node is assumed to hold up to discretization errors only. Afterwards, the algorithm is rerun with a finer grid, blue $\mathcal{P}_{vw}{}^2$ in Figure 3.3, around the initial solution $p_{vw}{}^{1*}$. This procedure is repeated until an equidistant minimal discretization is reached for all variables. Hence, the run-time of the algorithm increases by a small factor and is not guaranteed anymore to find the globally optimal solution, as claimed in Section 3.2. However, it works well in practice as shown in Section 3.4.2. It allows for tackling complex optimization problems with variables with largely different value ranges.

As a baseline in Section 3.4.2, the MILP formulation of [31] was implemented. The implementation is written using the GAMS modeling system calling either the CPLEX or the SCIP solver. CPLEX is a highly optimized commercial MILP solver whereas SCIP represents one of the best available open source implementations for MILP. The piece-wise linear cost functions for the MILP formulation use the same discretization values

as the DMP approach. While convex generator cost functions with a minimum load require only one binary variable per generator an additional binary variable is necessary for each discretization interval for non-convex functions. The additional binary variables indicate whether the previous interval is fully used. Moreover, since transmission constraints are of larger interest in distribution grids the formulation in [31] slightly extend by additional linear constraints on the transported power.

3.4.2 Experiments

The proposed DMP algorithmic framework is tested with two experiments in this section, exactly as it was done by the author in [73]. First the computational performance of the DMP framework is compared to MILP for sampled distribution networks that can be scaled to a very large size. Second, an exemplary smart distribution network of the future is examined in closer detail.

Computational Scaling Performance

For testing the computational performance of the DMP algorithm in comparison to the MILP approach, test distribution networks are sampled with different sizes n as illustrated in Figure 3.4. From a Medium Voltage (MV) ring, $n/100$ radial LV feeders dissect at busbar nodes S_i. This MV ring is fed by two High Voltage (HV)/MV transformers M_i. The number of households H_i in a line is chosen randomly in such a manner that the total number of households equals n and each feeder connects to at least one household.

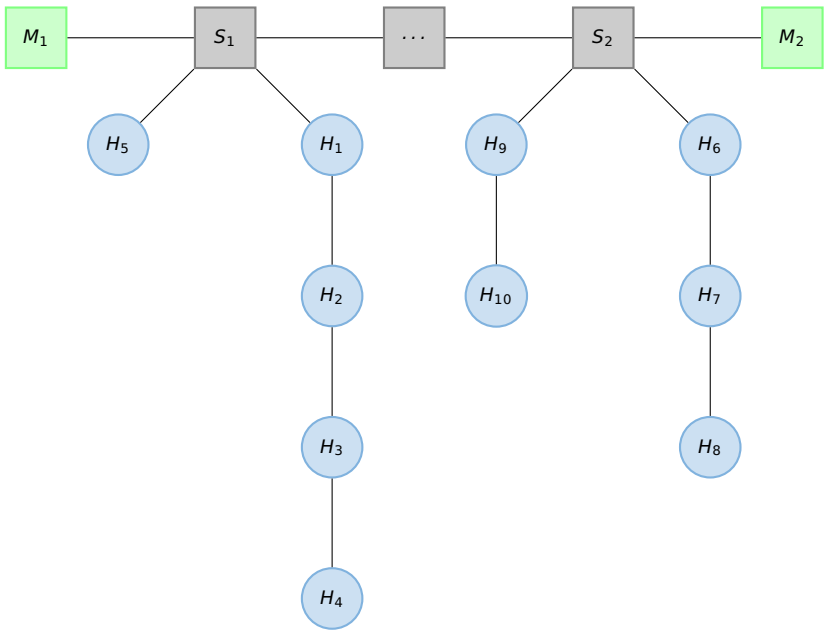


Figure 3.4: Test system for scaling experiments: Two HV/MV transformers M_1 and M_2 feed a MV ring of differing size. Radial LV lines with households H_i dissect at busbar nodes S_i.

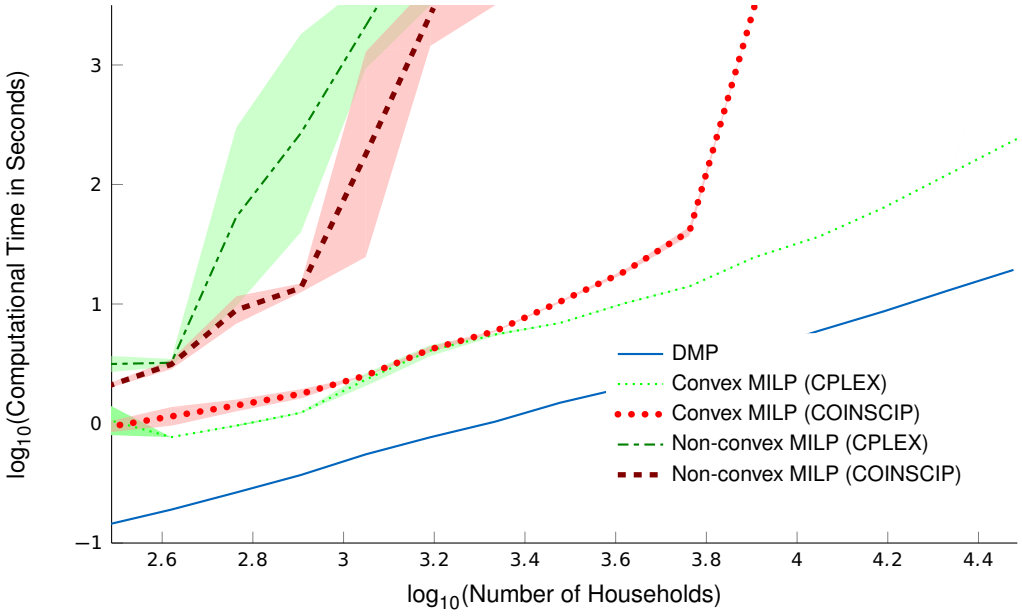


Figure 3.5: Computational time of ED for the test distribution systems shown in Figure 3.4, for convex (dotted) and non-convex (dashed) generator cost functions. Lines represent the mean computational time over 10 network random samples, shaded areas plus-minus one standard deviation. DMP algorithm (solid line) is compared to MILP using the CPLEX solver (fine lines) and the SCIP solver (bold lines).

Cost functions are designed as follows:

1. All households consume one power unit with infinite utility.
2. 70% of households additionally possess a decentralized power generation system able to produce between three and six units. The cost functions C_v are polynomials with maximal degree three and randomized parameters.
3. Busbar nodes do not produce or consume any power.
4. The HV/MV transformers act as slack nodes and supply any needed power at a constant-per-unit cost.

All LV branches possess a capacity of 3 power units, MV lines a capacity of 6. The discretization interval is uniformly 1 unit. Both optimization approaches are fed the same 10 independent network samples per network size $n \in [300, 30000]$. Timing tests are run on a virtual windows machine with 8 GB RAM and a 3.20 GHz CPU. To ensure a comparison on equal terms the solvers called by GAMS are restricted to run on one processor only and to solve the problems to optimality, i.e., with no remaining optimality gap. The run-time per problem was limited to 3000s.

When running these problems, all methods yielded the same optimal values up to machine precision. The different timings are shown in Figure 3.5. In the first set of experiments, shown as dotted lines, the generator cost functions are chosen to be convex within the operating range. This convexity is a common assumption in ED applications. It allows MILP to run with one binary variable per generator node only. Nevertheless, the SCIP branch-and-bound algorithm still scales exponentially in the network size. The CPLEX algorithm performs much better due to highly optimized pre-solve routines. The

DMP approach outperforms both methods in absolute numbers over the whole tested range of problem sizes. In the second set of experiments, shown as dashed lines, non-convex generator cost functions are examined. These are for example needed to describe the cost curves of real gas engines [61]. The MILP formulation then needs additional binary variables, one for each discretization interval per generator. In this case the run-time of the MILP solvers soars, while for the DMP approach no difference can be observed. Note also that the MILP approaches show a large variation in run-time, due to the fact that the pre-solving might, but is not guaranteed, to find a good starting point for the branch-and-bound algorithm that is applied afterward. In contrast, the DMP method has a deterministic run-time. In sum, the DMP method is the only one whose run-time scales linearly in the problem size. With the given time limit of 3000s, it is the only method to be able to compute optimal solutions for non-convex networks with more than 10k nodes. This may not be unrealistic in future smart distribution grids.

An Exemplary Smart Distribution Grid in Detail

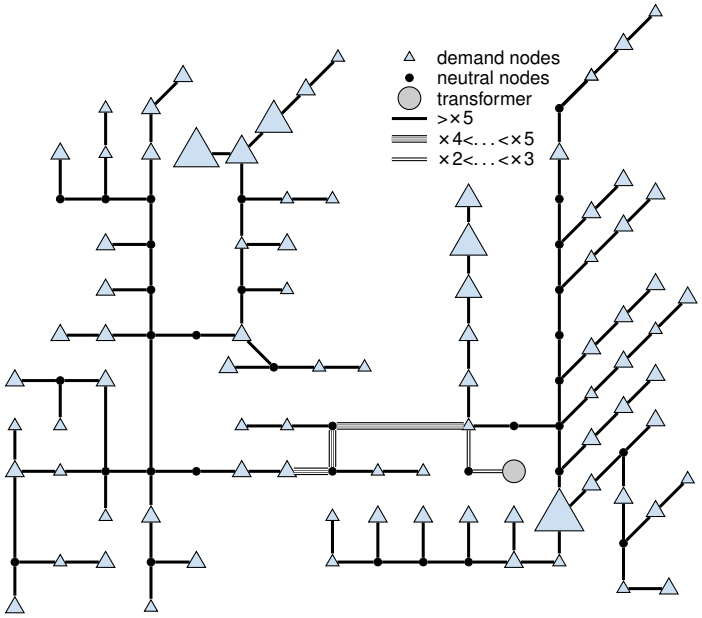
In this subsection, a smart distribution network of the future is examined in closer detail. Grid is constructed based on experience with real projects in Southern Germany. Due to confidentiality reasons, however, the original grid data could not be used.

The IEEE 123 node test distribution feeder [66] forms the basis for this example. Power line transmission limits are adjusted using Europe's 230V distribution grid voltage and the given consumer demands are scaled such that the smallest consumption of any node is 1.2 kW, a typical German household consumption. Note though grid has loops, the network switch states specified in [66] turn it into a tree-structure. The resulting grid model, see Figure 3.6(a), shows features of a typical distribution grid of today. The line capacities are dimensioned to cover all demands from the transformer node alone and are rather large in comparison with individual household consumptions. Line capacities are thus unlikely to pose an active limitation for dispatch problems. This allows applying common algorithms such as MILP formulation without line constraints. The situation, however, changes in the future when additional generation and consumption units are deployed decentralized throughout the grid, as is demonstrated with this example.

In the following are the prospects for technologies in these grids, Figure 3.6(b), detailed:

1. One EV per household chargeable with a maximal 20 kW and a utility function with marginal costs in the range 0.6-1 Euro per kW.
2. 60% of all households have an installed PV capacity of $14.9\,kW_{p_{el}}$ and offer their power output at prices between $0 - 0.2$ Euro per kW.
3. All PV owners additionally have a battery of equal size to the PV plant, generating a utility between $0.4 - 0.5$ Euro per kW for charging.
4. Ten Combined Heat and Power Plant (CHP) plants with a power rating of 140 kW_{el} are present in the network, preferably located at nodes with large demand. The generator model follows [61] where the cost functions are neither convex nor concave and a minimum load of 35 kW_{el} applies in running state.
5. Households without CHP have an electric heater consuming maximally $10\,kW_{P_{el}}$ while providing a utility of $0.2 - 0.3$ Euro per kW.

(a) Today's Distribution Grid



(b) Future Smart Grid

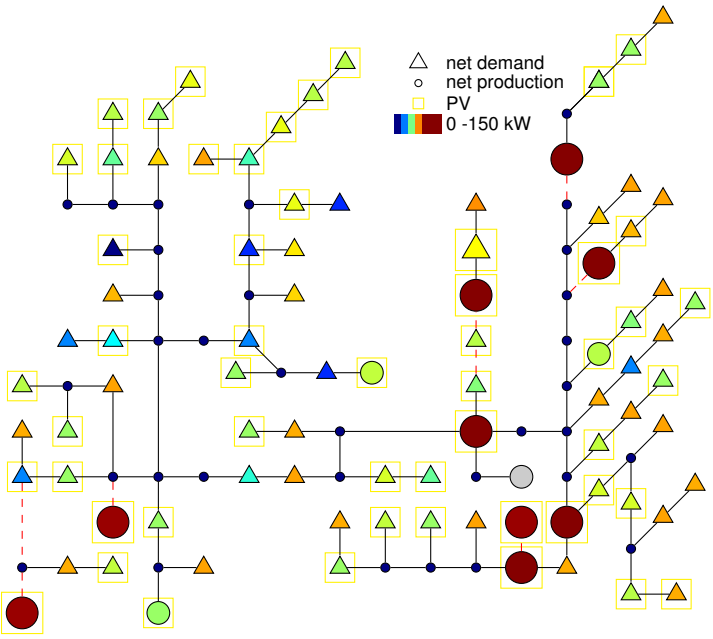


Figure 3.6: Exemplary Distribution Grid. (a) Structure of the passive distribution network supplied through the transformer. Triangle size encodes household demand and the grey round node denotes the transformer. Understanding the network as a tree with the transformer as the root node, lines are drawn according to the ratio of their capacity and the accumulated power demand of the subtree below the line. Two lines denote a capacity more than twice the accumulated demand of the subtree, three lines more than threefold, four lines more than fourfold and solid lines more than the fivefold accumulated demand. (b) Structure of a future active distribution grid. Circles denote net production nodes, where large circles are decentralized CHP plants of different sizes, yellow boxes PV installations and triangles net demand nodes. The color-coding denotes the result of the proposed DMP. Note that when using MILP scheduling without line constraints, six line capacity violations result (red dashed lines).

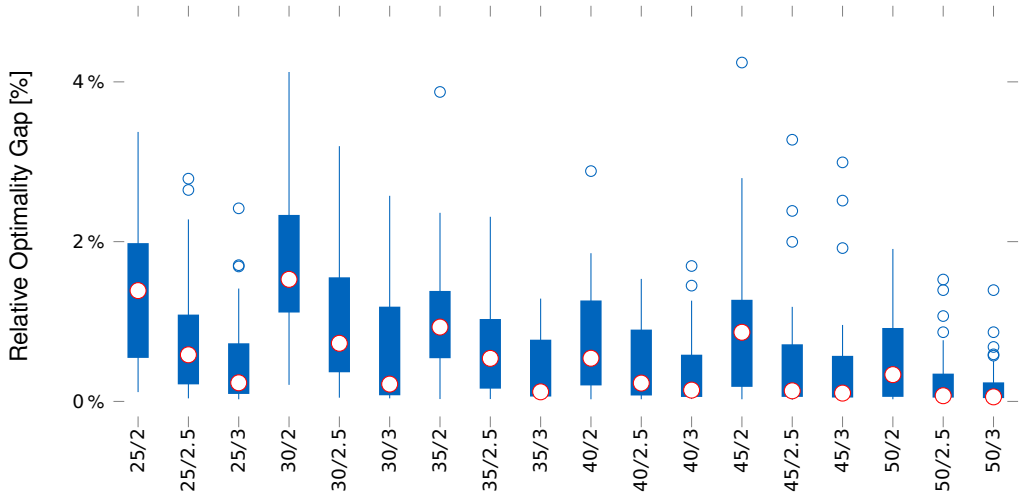


Figure 3.7: Relative optimality gap for the iterative implementation of the proposed DMP approach for 100 independent realizations of the exemplary smart distribution grid, see Figure 3.6. Each interval is discretized with 25-50 points and the re-discretization in each iteration takes into account a band of 2, 2.5 or 3 times the previous discretization interval. The blue boxes envelope the 98% quantiles, red circles the median and blue circles outliers.

The addition of these new prosumers to the distribution grid means that the network's line capacities can be violated for some dispatch assignments. For example in Figure 3.6(b) it is shown that MILP algorithm without additionally-introduced line constraints leads to 6 line violations for this network instance. In contrast, the DMP approach always yields feasible dispatch results.

Due to the large difference between line and generator capacities, the implementation had to resort in this example to an iterative re-discretization scheme as described in Section 3.4.1. Thus the procedure is not guaranteed to find the global optimum, but it mostly finds a solution very close to the optimum as is demonstrated by the following test, see Figure 3.7. For 100 independently sampled realizations of the exemplary smart distribution grid the optimal dispatch was computed with the MILP and with the proposed DMP approach. Since the MILP approach with line constraints is guaranteed to find the optimal solution, it was used as benchmark the quality of the DMP algorithm's results. With 50 discretization points for each interval and a re-discretization band of 2.5 times the discretization interval of the previous iteration, the suboptimality of the iterative approach is less than 1% in more than 75% of the cases, and less than 4.2% in the worst case.

The proposed DMP approach – despite hiding individual cost functions – delivers very broad information about the sensitivity of the solution. Remember that the optimal dual variables of a MILP approach denote the incremental total system cost for an infinitesimal demand change in one node. In contrast, the messages in the DMP approach allow one to deduce the exact total system cost increments for all possible demand changes. The difference is especially large in the case where a small step in the local demand leads the MILP problem to change its set of active constraints and the local linearization as expressed in the dual variables is thus no longer valid.

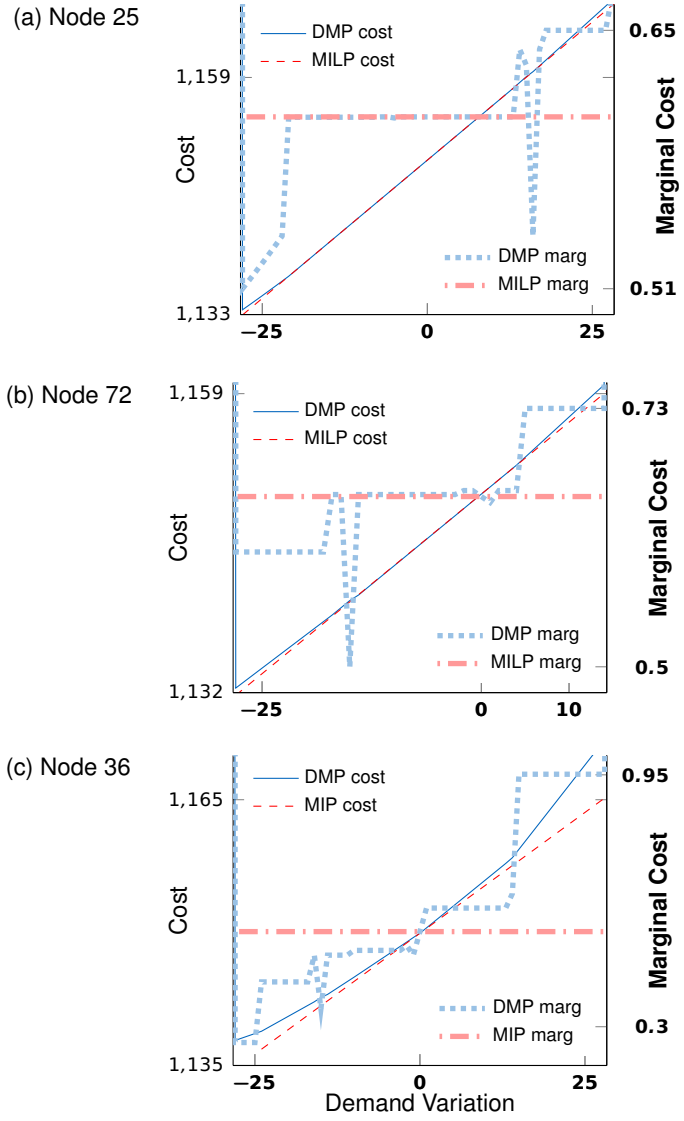


Figure 3.8: Total cost change for a local demand variation – assuming that the rest of the network stays the same. For three nodes from the network in Figure 3.6, the marginal cost approach derived from the MILP optimal dual variables is compared with the exact, locally computable result based on the discrete messages of the DMP approach.

In Figure 3.8, for three exemplary nodes the sum of the incoming messages plus the local node potential as a function of the local demand, minimized over all other local variables in the expression are plotted. This local operation derives the curve of total system cost changes as well as its derivatives. For comparison also the local linearization as derived from the optimal dual variables of the MILP approach are plotted. The difference between the local linearization of MILP and the holistic view derived from the local messages of the DMP approach is especially obvious for nodes close to highly loaded lines, see e.g. Figure 3.8(b). In this case even a small change in local demand may lead to infeasibility, and the linear continuation of the MILP costs is not valid any more.

3.4.3 Conclusion

The ED algorithm detailed in this section is a GM methods inspired DMP approach proposed in [73] and presented in [72]. The algorithm utilizes the grid topology and decomposes the ED problem into decentralized computable parts. Only a finite number of messages are passed between the nodes of the network, e.g. two on each edge. The algorithm finds the globally optimal solution for discretized variables and does so in linear time with respect to the grid size. The approach protects the privacy of local cost functions while giving a comprehensive picture as to how the system cost would change for local, finite demand variations. These features make this algorithm optimally suited for ED in future large, multi-owner smart distribution grids.

The classic ED problem is formulated with continuous dispatch variables in contrast to the discrete variable assumption in this DMP approach. The suitability of this assumption strongly depends on the structure of the individual problem. The first example in Section 3.4.2 shows that if the sizes of producers and consumers are relatively similar and the line capacities are not too large in comparison the DMP approach works well. However, the second example points out the arising problems with strongly differing line, generator as well as consumer sizes. For such problems a continuous approach utilizing the same divide and conquer method (2.21) could be used. Then, the messages are functions of continuous parameters and will typically grow to arbitrarily complex functions for large grid sizes. As discussed in Section 2.4 for approximated inferences, neither their exact representation nor their exact computation can be obtained in general. An exception are quadratic functions, since a one-dimensional slice of a multidimensional quadratic function is a quadratic function so exact message updates can be computed, see Gaussian belief propagation [77]. This findings support the basic idea for the Continuous Message Passing (CMP) approach proposed in the following Section 3.5.

3.5 Continuous Message Passing (CMP)

In this section the discrete approach of Section 3.4 is extended to continuous p_v, p_{vw} variables, denoting the result as the Continuous Message Passing (CMP) algorithm. The costs $C_v(p_v)$ and messages $m_{vw}(p_{vw})$ are assumed to be convex quadratic functions on compact intervals. This function class choice for the cost functions is motivated by generator cost curves that, in running mode, can often be approximated well using quadratic functions in the range between allowed minimum and maximum power production. Using the same function class for the messages follows naturally, where the interval restrictions enable additionally the representation of transport capacity limitations. The requirement of convexity of the cost functions is based on the idea that the chosen function class for the messages can only represent uni-modal functions well.

This specially adapted approximated inference scheme is built on an efficient message update mechanism, Section 3.5.1. Based on the approximation results different ways to derive globally feasible solutions are proposed, Section 3.5.2. Section 3.5.3 demonstrates CMP algorithm's features with two simulation experiments and Section 3.5.4 concludes.

3.5.1 Message Update

To derive (approximate) message updates according to equation (3.6) first the exact message m_{vw} is efficiently computed (Phase 1) and then the exact message is re-projected onto a chosen function class (Phase 2). The resulting procedure is detailed in Algorithm 1 and graphically explained in Figure 3.9.

In this section, the ideas behind the CMP algorithm are described and its correctness is shown. Starting point is the characterization of the exact message m_{vw} on the attainable subset $\hat{\mathcal{P}}_{vw}$ as being a convex, piece-wise quadratic function in Proposition 3.5.1. This holds under the assumption that the incoming (approximate) messages $\hat{m}_{kv}$ are convex quadratic functions on intervals and, as stated above, the local cost function C_v is convex and quadratic. In Proposition 3.5.2, the CMP is validated for computing the parameters of the exact message m_{vw} from Algorithm 1 Phase 1 and the best L^2 approximation given $\hat{m}_{vw}$ is derived in Proposition 3.5.3. Finally, in Proposition 3.5.4, it is shown that the best quadratic approximation of m_{vw}, the new approximate message $\hat{m}_{vw}$, will again be convex. This result closes the circle to the convexity assumption on the incoming (approximate) messages $\hat{m}_{vw}$ for the next message computation step.

Proposition 3.5.1. *Given that all incoming messages $\hat{m}_{kv}$ are convex quadratic functions on intervals $\hat{\mathcal{P}}_{kv}$ and that the local cost function C_v is convex on interval $\mathcal{P}_v$, the exact message update $m_{vw}(x)$ in* (3.6) *is a convex, piece-wise quadratic function.*

Proof. Denote the nodal cost function by $C_v(p_v) = \alpha_v {p_v}^2 + \beta_v p_v + \gamma_v$ on $\mathcal{P}_v = [\underline{p}_v, \overline{p}_v]$ and incoming messages by $\hat{m}_{kv}(p_{kv}) = \hat{\alpha}_{kv} p_{kv}^2 + \hat{\beta}_{kv} p_{kv} + \hat{\gamma}_{kv}$ on $\hat{\mathcal{P}}_{kv} = [\underline{p}_{kv}, \overline{p}_{kv}]$. Equation (3.6) can then be written as

$$m_{vw}(x) = \min_{\mathbf{p}} f(\mathbf{p}, x) \tag{3.10}$$

$$\text{s.t. } \mathbf{A}_{vw}\mathbf{p} + \boldsymbol{\alpha}_{vw}^{x} x + \boldsymbol{\alpha}_{vw} \leq 0, \tag{3.11}$$

Algorithm 1 Approximate Message Update

Input: Parameters of $\mathcal{P}_v$, $\hat{m}_{kv}$, $\hat{\mathcal{P}}_{kv}$ and $\mathcal{P}_{vw}$

Phase 1: *// Compute the exact piece-wise quadratic message parameters.*

Set $i \leftarrow 0$,

$\hat{\mathcal{P}}_{vw} \leftarrow [\underline{p}_v - \sum_{u \neq w} \underline{p}_{kv}, \overline{p}_v - \sum_{u \neq w} \overline{p}_{kv}] \cap \mathcal{P}_{vw}$

$x_0 \leftarrow \min(\hat{\mathcal{P}}_{vw})$

Solve (3.10),(3.11) for x_0 to obtain $\mathcal{A}_1 \leftarrow \mathcal{A}(x_0)$.

while $x_i < \overline{p}_{vw}$, $i \leftarrow i+1$ **do**

Determine the coefficients of $\mathbf{p}_i(x) = \mathbf{p}_i + \mathbf{p}_i^x x$, $\lambda_i(x) = \lambda_i^x x + \lambda_i$ via (3.12).

Determine and store $(\alpha_{vw}^i, \beta_{vw}^i, \gamma_{vw}^i)$ from $f(\mathbf{p}_i(x), x)$.

// Find the maximal x where $\mathbf{p}_i(x)$, $\lambda_i(x)$ is still primal/dual feasible for current active index set $\mathcal{A}_i$.

$x_i^{\mathcal{A}} \leftarrow \min\left[(-\boldsymbol{\alpha}_{vw}(k) - \mathbf{A}_{vw}(k,:)\mathbf{p}_i)/(\mathbf{A}_{vw}(k,:)\mathbf{p}_i^x + \boldsymbol{\alpha}_{vw}^x(k))\right]$
where $k \notin \mathcal{A}_i$, $\mathbf{A}_{vw}(k,:)\mathbf{p}_i^x + \boldsymbol{\alpha}_{vw}^x(k) > 0$

$x_i^{\lambda} \leftarrow \min(\lambda_i(k)/\lambda_i^x(k)$ where $k \in \mathcal{A}_i, \lambda_i^x(k) > 0)$

$x_i \leftarrow \min(x_i^{\mathcal{A}}, x_i^{\lambda})$

// Define the candidate sets for entering, leaving $\mathcal{A}_i$.

$\mathcal{A}_i^+ \leftarrow \{j \in \mathcal{A}_i^c : x_i = (-\boldsymbol{\alpha}_{vw}(j) - \mathbf{A}_{vw}(j,:)\mathbf{p}_i)/(\mathbf{A}_{vw}(j,:)\mathbf{p}_i^x + \boldsymbol{\alpha}_{vw}^x(j))\}$

$\mathcal{A}_i^- \leftarrow \{j \in \mathcal{A}_i : x_i = \lambda_i(j)/\lambda_i^x(j)\}$

if $\mathcal{A}_i^+ \neq \emptyset$ **then**

// Test full row-rank subsets of $\mathcal{A}_i \cup A_i^+$ to find $\mathcal{A}_{i+1}$ with respective $x_{i+1} > x_i$.

for all subsets $\hat{\mathcal{A}}_{i+1} \subset \mathcal{A}_i \cup A_i^+$ with $rank(\mathbf{A}_{vw}(\hat{\mathcal{A}}_{i+1}, :)) = |\hat{\mathcal{A}}_{i+1}|$ **do**

Solve $\left[\mathbf{A}_{vw}{}^T(:, \hat{\mathcal{A}}_{i+1})\right]\hat{\lambda} = -(\mathbf{c}_{vw} + \mathbf{c}_{vw}^x x_i + 2\mathbf{H}_{vw}\mathbf{p}_i(x_i))$.

if solution exist and all $\hat{\lambda}(j) > 0$ **then**

Find the maximal primal/dual feasible $\hat{x}_{i+1}$ for the possible active index set $\hat{\mathcal{A}}_{i+1}$ as above.

Break if $\hat{x}_{i+1} > x_i$.

end if

end for

$\mathcal{A}_{i+1} \leftarrow \hat{\mathcal{A}}_{i+1}$

else

$\mathcal{A}_{i+1} \leftarrow \mathcal{A}_i \setminus \mathcal{A}_i^-$

end if

end while

Phase 2: *// Determine approximation $\hat{m}_{vw}(x) = \hat{\alpha}_{vw}x^2 + \hat{\beta}_{vw}x + \hat{\gamma}_{vw}$ in L^2-sense.*

$$\begin{pmatrix}\hat{\alpha}_{vw}\\ \hat{\beta}_{vw}\\ \hat{\gamma}_{vw}\end{pmatrix} \leftarrow (\mathbf{X}_n - \mathbf{X}_0)^{-1}\left[\mathbf{X}_n\begin{pmatrix}\alpha_{vw}^n\\ \beta_{vw}^n\\ \gamma_{vw}^n\end{pmatrix} - \mathbf{X}_0\begin{pmatrix}\alpha_{vw}^1\\ \beta_{vw}^1\\ \gamma_{vw}^1\end{pmatrix} + \sum_{i=1}^{n-1}\mathbf{X}_i\left(\begin{pmatrix}\alpha_{vw}^i\\ \beta_{vw}^i\\ \gamma_{vw}^i\end{pmatrix} - \begin{pmatrix}\alpha_{vw}^{i+1}\\ \beta_{vw}^{i+1}\\ \gamma_{vw}^{i+1}\end{pmatrix}\mathbf{H}_{vw}\right)\right]$$

where $\mathbf{X}_i \leftarrow \begin{pmatrix}1/5\,x_i^5 & 1/4\,x_i^4 & 1/3\,x_i^3\\ 1/4\,x_i^4 & 1/3\,x_i^3 & 1/2\,x_i^2\\ 1/3\,x_i^3 & 1/2\,x_i^2 & x_i\end{pmatrix}$

Output: $(\hat{\alpha}_{vw}, \hat{\beta}_{vw}, \hat{\gamma}_{vw})$

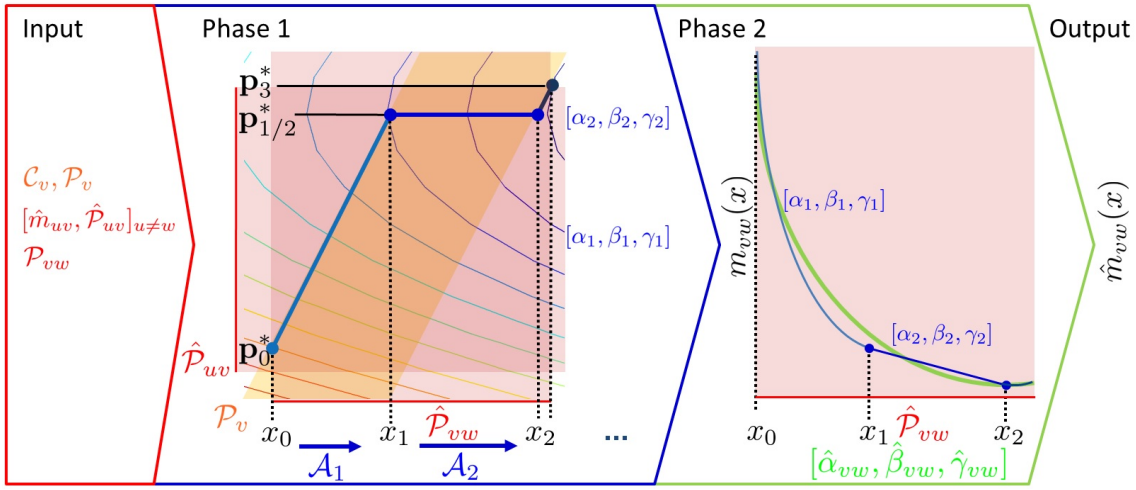


Figure 3.9: Schematic illustration of the message update, see Algorithm 1 Phase 1 and Phase 2 (Propositions 3.5.1/3.5.4).

where variable $x = p_{vw}$, the vector $\mathbf{p} = (p_{kv})_{k\in\mathcal{N}(v)\setminus\{w\}}$ and the objective function $f(\mathbf{p}, x) = \mathbf{p}^T\mathbf{H}_{vw}\mathbf{p} + (\mathbf{c}_{vw} + \mathbf{c}^x_{vw}x)^T\mathbf{p} + \text{const}$, for $x \in \hat{\mathcal{P}}_{vw} = [\underline{p}_{vw}, \overline{p}_{vw}]$. Here, $\underline{p}_{vw}$ is the maximum of the lower boundary of $\mathcal{P}_{vw}$ and the minimal x for which there exists $\mathbf{p}$ satisfying constraint (3.11). $\overline{p}_{vw}$ is defined respectively. The parameters of the objective function are $\mathbf{H}_{vw} = \mathbf{e}\alpha_v\mathbf{e}^T + diag(\hat{\alpha}_{kv})_{k\in\mathcal{N}(v)\setminus\{w\}}$, $\mathbf{c}_{vw} = (\hat{\beta}_{kv})_{k\in\mathcal{N}(v)\setminus\{w\}} + \beta_v$ and $\mathbf{c}_{vw}{}^x = 2\,\alpha_v\mathbf{e}$, where $\mathbf{e}$ is a vector of ones in dimension $\mathbf{p}$. Moreover, the parameters of constraint (3.11) are given by $\mathbf{A}_{vw} = \begin{pmatrix}\mathbb{1} & -\mathbb{1} & \mathbf{e} & -\mathbf{e}\end{pmatrix}^T$, $\boldsymbol{\alpha}_{vw} = \begin{pmatrix}-\overline{\mathbf{p}} & \underline{\mathbf{p}} & -\overline{p}_v & \underline{p}_v\end{pmatrix}^T$ and $\boldsymbol{\alpha}^x_{vw} = \begin{pmatrix}\mathbf{0} & \mathbf{0} & 1 & -1\end{pmatrix}^T$, where $\mathbb{1}$ is the identity matrix in the dimension of $\mathbf{p}$. Since both the incoming messages and the local cost function are by assumption convex, the above problem is convex, i.e. $\mathbf{H}_{vw} \succeq 0$.

For $x \in \hat{\mathcal{P}}_{vw}$ and the corresponding optimal (feasible) $\mathbf{p}^*(x)$ the KKT conditions imply that there exists at least one minimal active index set $\mathcal{A}(x)$ (i.e. $\mathbf{A}_{vw}(\mathcal{A}(x), :)$ has full row-rank, the constraints indexed with $\mathcal{A}(x)$ are satisfied with equality and

$$[\mathbf{A}_{vw}{}^T(\mathcal{A}(x), :)]\lambda = -(\mathbf{c}_{vw} + \mathbf{c}^x_{vw}x + 2\mathbf{H}_{vw}\mathbf{p}(x))$$

with $\lambda \geq 0$) and optimal, active dual variables $\lambda^*(x) \geq 0$ that fulfill

$$\begin{pmatrix} 2\mathbf{H}_{vw} & \mathbf{A}_{vw}{}^T(:, \mathcal{A}(x)) \\ \mathbf{A}_{vw}(\mathcal{A}(x), :) & 0 \end{pmatrix}\begin{pmatrix}\mathbf{p}^*(x) \\ \lambda^*(x)\end{pmatrix} = \begin{pmatrix} -\mathbf{c}_{vw} - \mathbf{c}^x_{vw}x \\ -\boldsymbol{\alpha}_{vw}(\mathcal{A}(x)) - \boldsymbol{\alpha}^x_{vw}(\mathcal{A}(x))x \end{pmatrix}. \quad (3.12)$$

This only holds point-wise for each x. However, by matrix inversion one can still derive an expression for $\begin{pmatrix}\mathbf{p}(x) & \lambda(x)\end{pmatrix}^T$ that is formally linear in x.

Now, it is argued that the thus computed values $\begin{pmatrix}\mathbf{p}(x) & \lambda(x)\end{pmatrix}^T$ are in fact optimal on intervals $[x_1, x_2]$ with equal $\mathcal{A}(x)$. To see this, assume that x_1, x_2 have $\mathcal{A}(x_1) = \mathcal{A}(x_2)$. Then for each convex combination $x = \mu x_1 + (1-\mu)x_2$ with $\mu \in [0, 1]$ the linearity of (3.12) implies that

$$\begin{pmatrix}\mathbf{p}(x) \\ \lambda(x)\end{pmatrix} = \mu\begin{pmatrix}\mathbf{p}^*(x_1) \\ \lambda^*(x_1)\end{pmatrix} + (1-\mu)\begin{pmatrix}\mathbf{p}^*(x_2) \\ \lambda^*(x_2)\end{pmatrix}.$$

Such $\begin{pmatrix}\mathbf{p}(x) & \lambda(x)\end{pmatrix}^T$ would be primal feasible, by the linearity of (3.11), and dual feasible, i.e. $\lambda(x) \geq 0$, and thus be optimal.

On intervals with equal $\mathcal{A}(x)$, $m_{vw}(x) = f(\mathbf{p}^*(x), x)$ thus is a quadratic, convex function due to the linearity of $\mathbf{p}^*(x)$ in x. Moreover, $m_{vw}(x)$ is also convex on all of $\hat{\mathcal{P}}_{vw}$, i.e. including points where $\mathcal{A}(x)$ changes. To see this, choose $x_1, x_2 \in \hat{\mathcal{P}}_{vw}$. The linear combination $\mu\mathbf{p}^*(x_1) + (1-\mu)\mathbf{p}^*(x_2)$, $\mu \in [0, 1]$, is again feasible due to the linearity of (3.11) and the convexity of the objective $f(p, x)$ (3.10) implies that

$$\begin{aligned} f(\mu\mathbf{p}^*(x_1) + (1-\mu)\mathbf{p}^*(x_2), \mu x_1 + (1-\mu)x_2) \leq \\ \mu f(\mathbf{p}^*(x_1), x_1) + (1-\mu) f(\mathbf{p}^*(x_2), x_2) = \\ \mu m_{vw}(x_1) + (1-\mu) m_{vw}(x_2). \end{aligned}$$

Since $m_{vw}(\mu x_1 + (1-\mu)x_2) \leq f(\mathbf{p}, \mu x_1 + (1-\mu)x_2)$ for all feasible $\mathbf{p}$, it follows that $m_{vw}(x)$ is convex. □

Proposition 3.5.2. *The full parameter set* $\{[\alpha^i_{vw}, \beta^i_{vw}, \gamma^i_{vw}, x_{i-1}, x_i]\}_{i\in[1,n]}$ *of the piecewise quadratic message function*

$$m_{vw}(x) = \sum_{i=1}^{n} \mathbf{1}_{[x_{i-1}, x_i)}(x)(\alpha^i_{vw} x^2 + \beta^i_{vw} x + \gamma^i_{vw}), \tag{3.13}$$

where $\mathbf{1}_{[x_{i-1}, x_i)}(x)$ *denotes the indicator function of the interval* $[x_{i-1}, x_i)$*, can be obtained by solving one quadratic optimization problem and a sequence of matrix vector operations, see Phase 1 of Algorithm 1. The number of pieces is bounded by* 2^{2d}*, where* d *is the degree of node* v*.*

Proof. The idea behind Phase 1 of Algorithm 1 is as follows: first an initial active index set $\mathcal{A}_1$ is determined by solving the quadratic problem (3.10) subject to (3.11) for the left boundary x_0 of $\mathcal{P}_{vw}$. Then it is searched for the largest possible x_1 whose KKT conditions still yield the same active index set $\mathcal{A}_1$. At this point, the KKT conditions start voting for another minimal active index set $\mathcal{A}_2$. This set is computed and iterate until the right boundary of $\mathcal{P}_{vw}$ is reached.

To prove the proposition's claims it is needed to argue that if it has been followed the rules of Phase 1 of Algorithm 1 the minimal active index set $\mathcal{A}_i$ is valid until x_i and that the computed next $\mathcal{A}_{i+1}$ is a valid minimal active index set for the KKT conditions from there on. Moreover, the computational effort for each such iteration as well as the number of iterations needs to be bounded as claimed.

Determining the next x_i is done by rewriting equation (3.12) as a linear vector function of the one-dimensional parameter x and checking how long the thus computed $\mathbf{p}_i(x)$ and $\lambda_i(x)$ are still primal and dual feasible. The computation can be performed by solving one small linear system and a series of minimum operations for determining the smallest x for which at least one primal/dual feasibility constraint is fulfilled with equality. Since $\mathbf{p}_i(x)$ and $\lambda_i(x)$ are primal and dual feasible for all x up to x_i and condition (3.12) is fulfilled by construction, the KKT conditions imply that $\mathbf{p}_i(x)$ is optimal and $\mathcal{A}_i$ valid throughout.

The number of iterations is limited to 2^{2d} since each problem has $2d$ constraints (i.e. index set elements) where d is node's v degree and no index set can occur more than

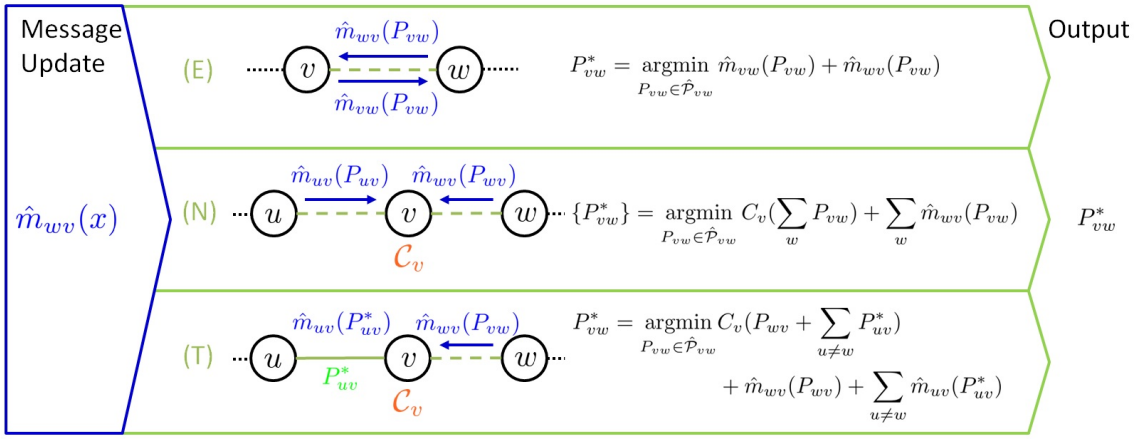


Figure 3.10: Methods for determining dispatch decisions from the approximated messages $\hat{m}_{vw}$, see Figure 3.9.

once. This is because a minimal active index set that is valid for two values of x is also valid for all values in-between, see the proof of Proposition 3.5.1.

Now it is argued that the algorithm's rules for computing $\mathcal{A}_{i+1}$ actually yield the minimal active index set implied by the KKT conditions for the next interval of non-zero length. Denote the true next index set by $\overline{\mathcal{A}}_{i+1}$. For all indices that are in $\overline{\mathcal{A}}_{i+1}$ the corresponding primal feasibility constraints are fulfilled with equality for all x in the interval, also at $x = x_i$. Indices in $\overline{\mathcal{A}}_{i+1}$ that are not part of $\mathcal{A}_i$ will then be in the candidate set $\mathcal{A}_i^+$ by definition and it has to hold $\overline{\mathcal{A}}_{i+1} \subseteq \mathcal{A}_i \cup \mathcal{A}_i^+$. If $\mathcal{A}_i^+$ is not empty, the algorithm checks all subsets $\hat{\mathcal{A}}_{i+1}$ of $\mathcal{A}_i \cup \mathcal{A}_i^+$ whether they allow to satisfy the KKT conditions at x_i and lead to an $x_{i+1} > x_i$. The KKT conditions for $\hat{\mathcal{A}}_{i+1}$ are met at x_i if $\mathbf{A}_{vw}(\hat{\mathcal{A}}_{i+1}, :)$ has full row rank and the $\hat{\lambda}$ satisfying $[\mathbf{A}_{vw}{}^T(:, \hat{\mathcal{A}}_{i+1})]\hat{\lambda} = -(\mathbf{c}_{vw} + \mathbf{c}^x_{vw} x_i + 2\mathbf{H}_{vw}\mathbf{p}_i(x_i))$ are all positive. The algorithm is thus guaranteed to find the next true active index set $\overline{\mathcal{A}}_{i+1}$. If $\mathcal{A}_i^+$ is empty, it follows that $\overline{\mathcal{A}}_{i+1} \subset \mathcal{A}_i$. In this case, it is set $\mathcal{A}_{i+1} = \mathcal{A}_i \setminus \mathcal{A}_i^-$, i.e. deleting all those indices from $\mathcal{A}_i$ whose dual feasibility constraints become tight at x_i. If this guess for $\mathcal{A}_{i+1}$ were to deleted too many indices from $\mathcal{A}_i$, the algorithm's next iteration would yield $x_{i+1} = x_i$ and would add the needed indices again. This step would then also guarantee that $x_{i+1} > x_i$, implying that the algorithm as a whole cannot cycle without making progress on x. □

Proposition 3.5.3. *The quadratic approximation $\hat{m}_{vw}(x) = \hat{\alpha}_{vw}x^2 + \hat{\beta}_{vw}x + \hat{\gamma}_{vw}$ of the piece-wise quadratic function $m_{vw}(x) = \sum_{i=1}^{n} \mathbf{1}_{[x_{i-1},x_i]}(x)(\alpha^i_{vw}x^2 + \beta^i_{vw}x + \gamma^i_{vw})$ with minimal L^2-norm distance is*

$$\begin{pmatrix} \hat{\alpha}_{vw} \\ \hat{\beta}_{vw} \\ \hat{\gamma}_{vw} \end{pmatrix} = (\mathbf{X}_n - \mathbf{X}_0)^{-1} \left[\mathbf{X}_n \begin{pmatrix} \alpha^n_{vw} \\ \beta^n_{vw} \\ \gamma^n_{vw} \end{pmatrix} - \mathbf{X}_0 \begin{pmatrix} \alpha^1_{vw} \\ \beta^1_{vw} \\ \gamma^1_{vw} \end{pmatrix} + \sum_{i=1}^{n-1} \mathbf{X}_i \left(\begin{pmatrix} \alpha^i_{vw} \\ \beta^i_{vw} \\ \gamma^i_{vw} \end{pmatrix} - \begin{pmatrix} \alpha^{i+1}_{vw} \\ \beta^{i+1}_{vw} \\ \gamma^{i+1}_{vw} \end{pmatrix} \right) \right]. \tag{3.14}$$

Moreover, $(\mathbf{X}_n - \mathbf{X}_0)$ has full rank as long $x_n \neq x_0$.

Proof. The L^2-norm distance is

$$L_{vw}(\hat{\alpha}_{vw}, \hat{\beta}_{vw}, \hat{\gamma}_{vw}) := \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \left[\begin{pmatrix} \hat{\alpha}_{vw} - \alpha_{vw}^i \\ \hat{\beta}_{vw} - \beta_{vw}^i \\ \hat{\gamma}_{vw} - \gamma_{vw}^i \end{pmatrix}^T \begin{pmatrix} x^2 \\ x \\ 1 \end{pmatrix} \right]^2 dx$$

$$= \sum_{i=1}^{n} \begin{pmatrix} \hat{\alpha}_{vw} - \alpha_{vw}^i \\ \hat{\beta}_{vw} - \beta_{vw}^i \\ \hat{\gamma}_{vw} - \gamma_{vw}^i \end{pmatrix}^T \int_{x_{i-1}}^{x_i} \begin{pmatrix} x^4 & x^3 & x^2 \\ x^3 & x^2 & x \\ x^2 & x & 1 \end{pmatrix} dx \begin{pmatrix} \hat{\alpha}_{vw} - \alpha_{vw}^i \\ \hat{\beta}_{vw} - \beta_{vw}^i \\ \hat{\gamma}_{vw} - \gamma_{vw}^i \end{pmatrix}.$$

For the minimum it holds that

$$\nabla L_{vw} = 2 \sum_{i=1}^{n} [\mathbf{X}_i - \mathbf{X}_{i-1}] \begin{pmatrix} \hat{\alpha}_{vw} - \alpha_{vw}^i \\ \hat{\beta}_{vw} - \beta_{vw}^i \\ \hat{\gamma}_{vw} - \gamma_{vw}^i \end{pmatrix} = 0.$$

Solving for $\begin{pmatrix} \hat{\alpha}_{vw} & \hat{\beta}_{vw} & \hat{\gamma}_{vw} \end{pmatrix}^T$ results in (3.14). Since the determinant $\frac{(x_n - x_0)^9}{2160}$ of $(\mathbf{X}_n - \mathbf{X}_0)$ has its only root at $x_0 = x_n$, $(\mathbf{X}_n - \mathbf{X}_0)$ has full rank for $x_n \neq x_0$. □

Proposition 3.5.4. *A quadratic function $\hat{m}_{vw}$ that minimizes the L^2-norm distance to a piece-wise quadratic, convex function m_{vw} is itself convex.*

Proof. Assume that $\hat{m}_{vw}$, described in Algorithm 1 Phase 2, is a non-convex, quadratic function and denote by $m_{vw}^l(x)$ the linear function either through two intersection points I, i.e. $\hat{m}_{vw}(x_i) = m_{vw}(x_i)$ for $i \in I$, or if it only one intersection point I exists the tangential line at $\hat{m}_{vw}(x_I)$. Then it holds that $|m_{vw}(x) - \hat{m}_{vw}(x)| > |m_{vw}(x) - m_{vw}^l(x)|$ for all $x \in \hat{\mathcal{P}}_{vw} \setminus \{x_I\}$, since $m_{vw}(x)$ is convex. Thus, ${m_{vw}}^l$ is a better approximation than $\hat{m}_{vw}$, which contradicts the assumption. □

As for classical belief propagation, the message update is computed for a node once all incoming messages are known. A feasible such ordering can always be found for trees, e.g. via depth-first search.

3.5.2 From Messages to Decisions

Having computed all messages one can derive the final dispatch decisions for the transport variables ${p_{vw}}^*$ via equation (3.7). The nodal injections ${p_v}^*$, i.e. the set points for generators and consumers, follow from ${p_v}^* = \sum_{w \in \mathcal{N}(v)} {p_{vw}}^*$. This approach is going to be called the edge based method (E).

However, it may then happen that the resulting nodal injections ${p_v}^*$ are not fully feasible since the each message update computation contained individually made approximations. This problem cannot easily be resolved completely. Thus, this section describes two further methods how the final dispatch decisions can be derived from the messages with different properties, depicted in Figure 3.10. The CMP methods are benchmarked with the examples in Section 3.5.3. Note that using exact message updates for discretized values as in Section 3.4 or [73] this problem does not occur.

In contrast to the edge based method, the node based method (N) asserts the nodal constraints, but allows for some and hopefully not to large violations of the edge constraints. It first determines the nodal injections $p_v{}^*$ by minimizing separately for each node the sum of the nodal costs and the incoming messages, a d-dimensional optimization problem where d is the node's degree. This may result in multiple decisions for the transport variables $p_{vw}{}^*$ that need not be consistent with each other. The true, optimal transport variables have to be determined by a power flow calculation afterwards, if needed.

The third approach tries to combine (E) and (N). The tree based method (T) calculates the optimal transports in a tree-aligned order from the leaves to the root. An one-dimensional optimization problem over p_{vw} is solved at each node if all $\{p^*_{kv}\}_{k \in \mathcal{N}(v) \setminus \{w\}}$ are already known. Method (T) will typically respect both node and edge constraints. However, both constraints can be violated as approximation implied non-optimal decisions in sub-trees are propagated upwards through the tree and lead to small infeasibilities further up. The transformer to the transmission grid is typically chosen as the root, since its large capacity and good connection allow to avoid many such problems.

3.5.3 Experiments

To evaluate the proposed CMP methods, first its advantages are highlighted in comparison to the decentralized methods DMP and Alternating Direction Method of Multipliers (ADMM), and a central Quadratic Programming (QP) solver for a stylized, easy-to-understand single distribution grid branch that includes generation nodes of largely different capacity. The second example then shows that the CMPs approach can solve a more complex, realistic future distribution grid scenario.

Single Branch Example

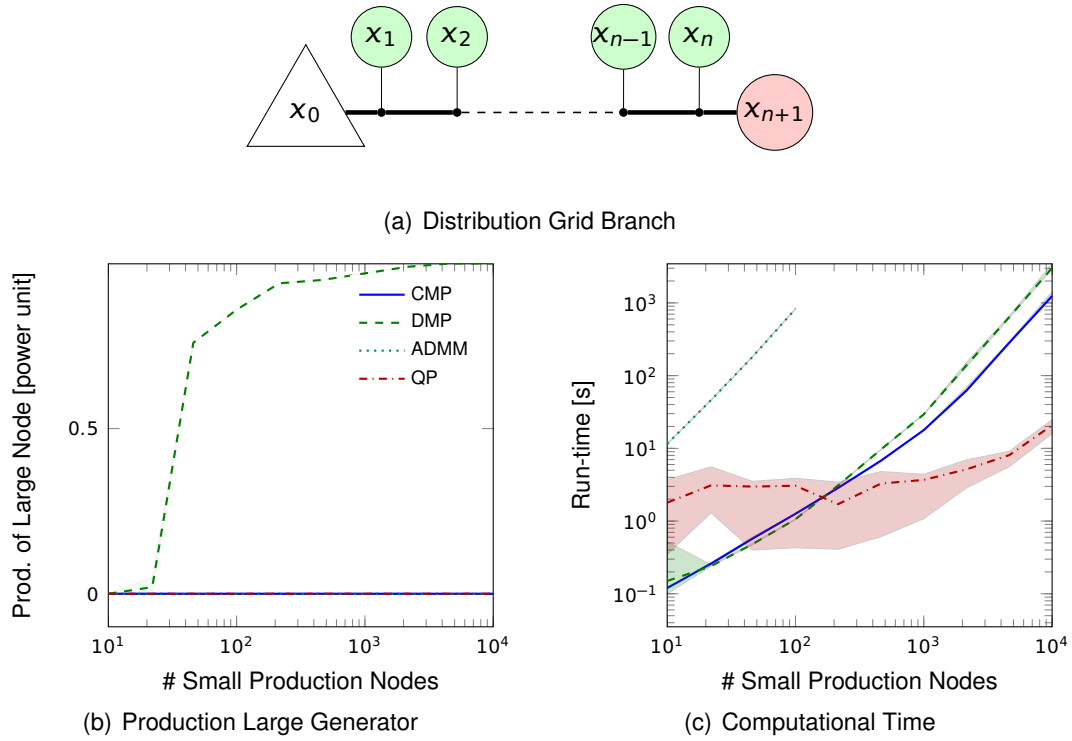
The structure of the single branch grid is shown in Figure 3.11 (a). The grid starts with a demand node △ with a power demand of one power unit. A sequence of n small production nodes ○ follows that are able to produce $1/n$ power unit, representing for example renewable production units at household connections. The chain is terminated by a large generation node ○ with capacity 1 power unit, e.g. a CHP unit. The small production node costs

$$C_i(p_i) = \alpha_i p_i^2 + \beta_i p_i, i \in 1, \ldots, n,$$

are sampled uniformly from $\alpha_i \in [0.1, 0.6]$ and $\beta \in [0.01, 0.1]$. While, the large production node cost

$$C_{n+1}(p_{n+1}) = \alpha_{n+1} p_{n+1}^2 + \beta_{n+1} p_{n+1},$$

is set ex-post such that it is costlier than any small node, i.e. $\alpha_{n+1} = \max_{i \in [1,n]} \alpha_i$ and $\beta_{n+1} = 2/n \max_{i \in [1,n]} \alpha_i + \max_{i \in [1,n]} \beta_i + 0.001$. Line capacities are chosen such that they do not pose any active restrictions. Given this setup, i.e. cheaper small production nodes, supplying all demand with the small production nodes is the optimal solution. All tests were repeated 100 times.



(a) Distribution Grid Branch

(b) Production Large Generator

(c) Computational Time

Figure 3.11: (a) Distribution grid branch where n small production nodes $x_1, \ldots, x_n$ with capacity $1/n$ power unit and one large generator x_{n+1} with capacity 1 power unit have to supply demand $x_0 = 1$ power unit Costs are set such that production at the small nodes is optimal. (b) Non-optimal production at the large generator for different dispatch methods. (c) Computational time for the different methods the shaded areas denote the minimum and maximum computation time. Method QP denotes solving the problem with CPLEX's quadratic programming solver.

DMP uses discretized messages and performs iterative rediscretization from coarse to fine, see Figure 3.3. It yields non-optimal results for the proposed problem for modest to large n, see Figure 3.11 (b). For growing n the capacity of the small generators will at some point be smaller than the discretization interval of the first, coarse discretization and the algorithm will then yield strongly misleading initial results. Later rediscretization can only improve things in a neighborhood of the initial solution. In contrast, the fully continuous approach of this section always yields the optimal result for this example.

Concerning computation time, CMP scales linearly, see Figure 3.11 (c). Computation times are slightly lower than for DMP since rerunning the model with finer discretization is avoided. For the non-convex problem setting in Section 3.4.2, the central solver is strongly outperformed by DMP, shown in Figure 3.5. This does not hold for the continuous, convex quadratic setting in this section where the QP solver performs much better. However, note that the QP approach requires full model knowledge in one location – a requirement that is typically not satisfiable in competitive energy market environments. The computational time of ADMM is comparably high, where all methods were implemented on a single processor here.

n	ADMM		DMP		CMP	
	Iter	# Com	Iter	# Com	Iter	# Com
10	35	2940	2	840	**1**	**210**
22	69	12420	3	2700	**1**	**450**
46	136	50592	3	5580	**1**	**930**
100	287	230748	3	12060	**1**	**2010**

Table 3.1: Number of iterations and values communicated between electrical nodes for ADMM, DMP and CMP.

The computations of ADMM, DMP and CMP could be parallelized in practice. On real field controllers the total execution time would then mostly depend on the communication pattern between the nodes of each method, see Table 3.1. This is due to typically long latencies and tight bandwidth restrictions. CMP requires just five values to fully describe message-functions and exchanges these just once. The messages of DMP are larger and more iterations are needed. ADMM exchanges only 4 variables between electrical nodes in each iteration, but requires many iterations to converge. It should be noted that since CMP messages are computed in tree-order, a CMP iteration requires g communication cycles where g is the tree-depth of the grid. This reduces CMP's benefit in comparison to ADMM in this example where $g = n + 2$, but ADMM still needs about three times as many iterations and cycles. The small communication load of CMP thus seems to be one of its major advantages.

Realistic Future Distribution Grid

In this section, it is shown that the proposed CMP algorithm can handle realistic future scenarios, where distributed PV installations together with larger CHP and new controllable consumers like EV, electric heating or battery storages challenge the electrical distribution grid. Two or more LV-areas of IEEE 123 [66] topology were simulated that are connected via a joint MV line to the HV grid and a wind turbine, see Figure 3.12. The future distribution grids for this experiment were generated as follows. First, it is described how single LV-areas are constructed from the IEEE 123 distribution grid. Then, the LV-areas' connection are described. Finally, additional future generators and consumer units and their parameters are defined that are integrated into the LV-areas.

Special interest lies on the distribution grid integration of renewable energies in Europe and especially Germany where those issues are most severe today. However, public, non-confidential material about distribution grids is only available for North America. Thus the IEEE 123 example feeder was used for the grid topology and the cable parameters, but adjust the lines' power limits and the connected demands to a European distribution setup with only 230V lines instead of the 4kV lines in Northern America. Power limits follow from the IEEE's ampacity limits times the distribution voltage. A connection of the IEEE test feeder supplies on average three houses via a local step-down transformer. In contrast, each house is connected directly in the European setup. The non-deferrable demands give in the IEEE test system were scaled such that the average load represents a typical German household demand.

Table 3.2: Power limits, cost function and installation rate parameters for all unit types.

type	min	max	α	β	γ	$\overline{\alpha}$	$\overline{\beta}$	$\overline{\gamma}$	cD	in	env
CHP	35	140	0.0095	$1.\overline{3}$	$-58.\overline{3}$	0.0195	$1.6\overline{3}$	$-58.\overline{3}$	3	0.05	1
EV	-10	0	0.125	4.5	0	0.325	5.3	0	3		
Stor	-5	5	0.075	3.95	0	0.155	4	0	3		
Heat	-7	0	0.0714	4	0	0.0714	4.45	0	3		
PV	0	14	0.0036	0.5	0	0.0061	1	0	3	0.7	0.3
Wind	0	1200	$8.\overline{3}E^{-5}$	0.5	0	$8.\overline{3}E^{-5}$	1	0	3	1	0.6
Gen_{Plus}	0	366	0	5	0	0	5	0	3		
Gen_{Neg}	-366	0	0	3	0	0	3	0	3		

The LV-areas are connected by a line and a transformer to the MV connector. This line has a power capacity matching the largest line capacity of the LV-area. The MV line has a capacity equal to the number of modeled IEEE LV-areas times the LV/MV transformer's capacity. It spans from the transformer node to the wind turbine, as shown in Figure 3.12.

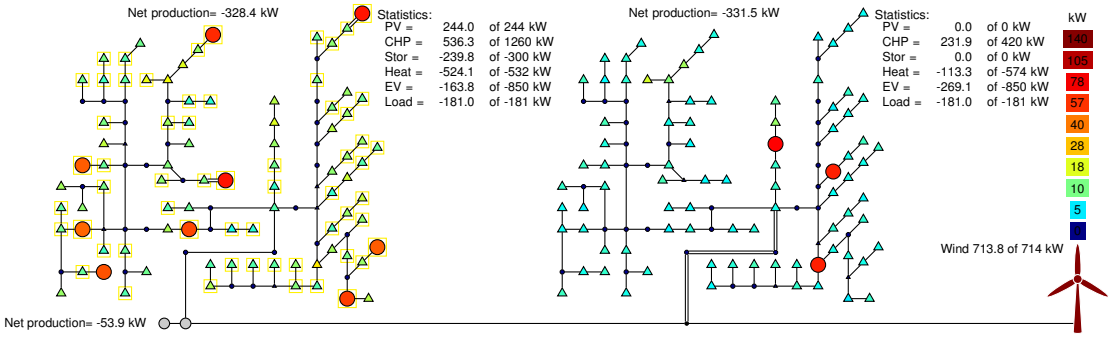
The single unit type specifications are detailed in Table 3.2, which row and column names are introduced in the following by the reference in brackets. Into this classical demand oriented setup decentralized renewable production, i.e. Photovoltaic (PV) and Wind (Wind), Combined Heat and Power Plants (CHP), as well as deferrable loads such as Electric Vehicles (EV), electric Heaters (Heat), and stationary Storages (Stor), were integrated. The transformer to the transmission grid can act as a Generator, i.e. importing energy into the modeled system (Gen_{Plus}), or as a consumer, i.e. exporting energy (Gen_{Neg}).

The parameters of each unit type are specified in Table 3.2. The minimum (min) and maximum (max) production capacity for environment-dependent generation types is scaled by the environmental factor (env). The cost functions of degree $cD - 1$ are given as

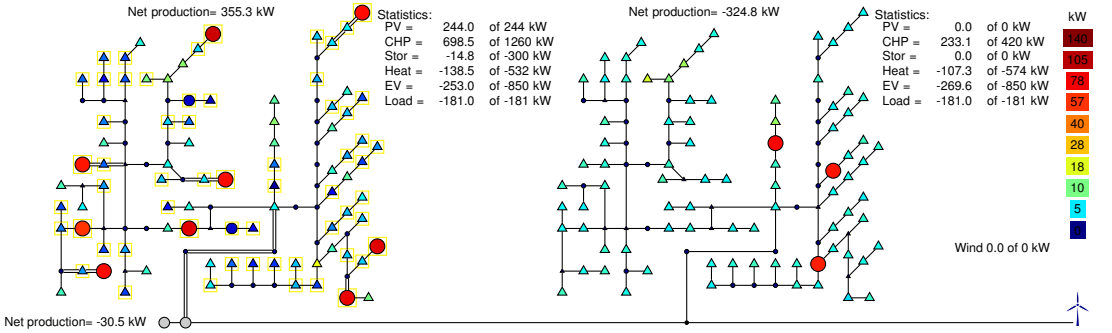
$$\begin{aligned} C_{type}(x) = &[\alpha_{type} + (\overline{\alpha_{type}} - \alpha_{type})\rho_1]x^2 + \\ &[\beta_{type} + (\overline{\beta_{type}} - \beta_{type})\rho_2]x + \\ &\gamma_{type} + (\overline{\gamma_{type}} - \gamma_{type})\rho_3, \end{aligned} \tag{3.15}$$

where ρ_i, $i \in 1, \ldots, 3$ is sampled uniformly from $[0, 1]$ independently for each unit. The column (in) specifies the fraction of demand nodes at which units of this type are going to be installed. Exceptions are the storage installed at each house with PV, the electric heater at all demand nodes without a CHP and EV at all nodes. Moreover, the PV-installations in the IEEE LV-areas are clustered based on individual qualities. This leads to to PV-dominated areas, solar areas, on the one hand, whereas in others PV installations are rare.

In the two LV-areas example, Figure 3.12, the left region is a solar area. This example is used to illustrate two situations, windy and no-wind. In the windy situation the environmental factor for wind is 0.6, like in Table 3.2. No wind is blowing at all in the no-wind situation, i.e. the environmental factor for wind is zero. The CMP method applied with variant (T), is able to produce a valid dispatch for both situations shown in Figure 3.12.



(a) Sunny + Windy

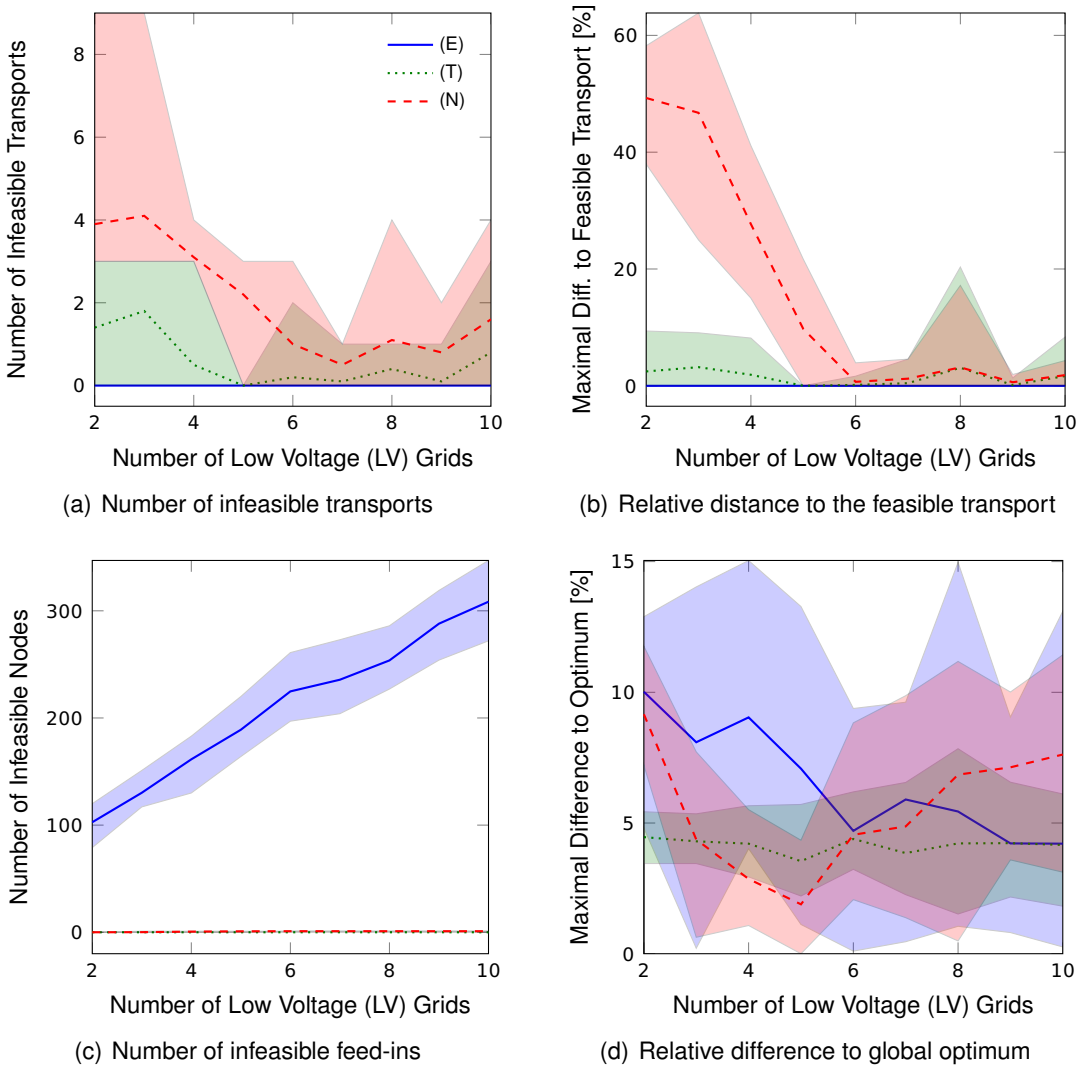


(b) Sunny + No Wind

Figure 3.12: A realistic future distribution grid with two residential, LV-areas connected via a MV line to the HV grid (transformer at the gray node). A wind turbine outside the residential areas is connected to the MV line. Many small PV installations (yellow boxes) are distributed throughout the left LV-area together with several decentralized CHPs (big circle nodes). The right area represents a classical consumption-dominated grid, again with a few CHPs. Dispatch results with CMP algorithm, variant (T), are coded as follows: node color encodes the absolute magnitude of the power injection (loads and generation connected to the same grid location is summed), triangles denote net consumption, circles net generation. Double lines denote a transport within 10% of the maximum line capacity.

In the windy situation, the wind turbine's power is almost fully consumed within the two LV-areas. In the no wind case, the left solar LV-area supplies the right one with energy and demand is generally reduced.

This example with more LV-areas is used to compare the different variants of CMP for deriving dispatch decisions from the approximate messages (all variants yield optimal results in the first example). Grids with different numbers of LV-areas were dispatched for the windy setup, ten random samples for each size. The resulting number and quality of infeasibilities as well as the relative sub-optimality compared to the globally optimal QP solution is shown in Figure 3.13. The number of infeasible edge and node constraints increases with the number of LV-areas, since the approximation errors accumulate with the growing number of nodes and edges. The tree-based variant (T) outperforms the other two with respect to both the number and severity of the resulting edge or node infeasibilities. The maximum edge violations for variant (T) is below 20% of the edge



(a) Number of infeasible transports

(b) Relative distance to the feasible transport

(c) Number of infeasible feed-ins

(d) Relative difference to global optimum

Figure 3.13: Comparison of quality of the dispatch decisions derived by CMPs algorithm's variants (E), (N), and (T) for the grid topology of Figure 3.12 and different numbers of connected, LV-areas. Lines depict the mean values and shaded areas the band between minimal and maximal value over 10 samples.

capacity for all computed grids; the maximal violation for an average network is only 3.2%. Such violations may be acceptable in practice. Moreover, the mean sub-optimality of the achieved true systems costs for variant (T) is below 4.5% for all network sizes; the worst-case sample is 8% sub-optimal. Thus, method (T) can be seen as the best option for future use.

3.5.4 Conclusion

This section proposes an ED algorithm utilizing methods from non-linear optimization and approximate inference. Approximated inference is a vital field of studies in the GM community. However, approximations using the proposed function class, i.e. quadratic plus interval constraints, for this purpose have apparently not been done before. The message based dispatch approach detailed in Section 3.1, proposed and evaluated for discrete variables in [73] or Section 3.4, has been extended to continuous variables. The DER representation was restricted to convex quadratic functions on compact intervals and each message projected on the same. Thereby, a similar nodal problem per message computation was ensured. It was shown based on the KKT-conditions that these nodal problems could be reduced to one quadratic problem and several linear problems, limiting the required distributed computational capability.

This CMP approach features a strongly increased practicality in comparison to DMP and ADMM. First of all, it can solve ED problems with largely different capacities since it requires no common fine discretization in contrast to DMP, as discussed in Section 3.4. The messages are given by five values, which significantly reduces communication load. This was shown in the first example, in Section 3.5.3, where the CMP outperformed ADMM and DMP with regard to the number of values needed to be communicated, and DMP also with regard to the dispatch results. Furthermore, the partial information security of the MP approach was further intensified by the message mapping onto a quadratic function on a compact interval. The CMP ability to evaluate finite step-size sensitivities and linear scaling in the network size remain the same as in DMP. The solar village example, the second example in Section 3.5.3, showed that the approach can balance demand and supply in large grids of different scales.

Chapter 4

Dynamic Programming Volt-VAr Control for Active Grids

In this chapter, a dynamic programming method for optimal Volt-VAr Control (VVC) is proposed. The VVC problem is to find the complex voltages $\mathbf{u}$ and a number of discrete transformer tap positions and/or capacitor bank positions τ such that the costs of power production minus the utility of power consumption in the grid $\mathcal{G}(\mathcal{V}, \mathcal{E})$ are minimized. This has to be done subject to the compulsory voltage limits and feasible sets for the nodal injections. The minimization directly results in loss minimal solution since higher losses are directly reflected in a higher objective function value.

The VVC has been the object of renewed interest due to the rise in intermittent renewable energy sources installed in distribution grids. The rapid fluctuations in contributions from the renewables that lead to a voltage rises and voltage drops challenges passive distribution grids. In order to satisfy the voltage limits, these rises and drops need to be actively counterbalanced. Thus, active controllable grid components, e.g. On-Load Tap Changer (OLTC) and capacitor banks, are required. But the number of active components and their electrical interaction increases the complexity of controlling them optimally and renders simple heuristics inadequate.

The major source of complexity in the VVC problem is the number of interacting discrete variables τ. As a rough rule, the runtime of an optimization problem rises exponentially with the number of discrete variables. To address this issue, the dynamic programming approach divides the VVC problem into subproblems of one discrete variable and one coupling variables. The subproblems are then successively solved and coordinated by functions of the common variables, once coupling variables and once slack variables. Accordingly, solving the small subproblems lead to an overall solution in line with the dynamic programming idea, introduced in Section 2.5. Moreover, the structure of the subproblems is utilized to further reduce the computational effort.

In Section 4.1 the theoretical foundation for the dynamic programming approach is derived. Based on this foundation, the numeric implementation is detailed in Section 4.2 and experiments in Section 4.3 compare the proposed approach with other commonly used approaches.

4.1 Volt-VAr Control as Dynamic Programming Problem

This section describes the theoretical foundation for the dynamic programming subproblems. Based on Alternating Current (AC) power flow equations, a dynamic programming approach for two subproblems is explicitly derived and formulated. The findings are then utilized to produce subproblems of at most one discrete variable that enable an efficient solution despite the discrete variable, e.g. by enumeration. Finally, the influence of the optimal discrete variable on the coordinating functions of the relating variable is discussed, which is the reason for using only one coupling variable per subproblem.

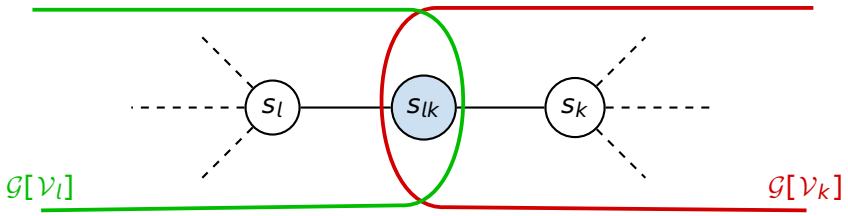
The AC power flow equations (2.17) are composed of node injection and branch flow equations. For each node $v \in \mathcal{V}$ the voltage magnitude $\hat{u}_v$ and the power flows s_{vw} to the neighboring nodes $w \in \mathcal{N}(v)$ determine the node injection as

$$s_v = \sum_{w \in \mathcal{N}(v)} s_{vw} + y_{v0}^H \hat{u}_v{}^2. \tag{4.1}$$

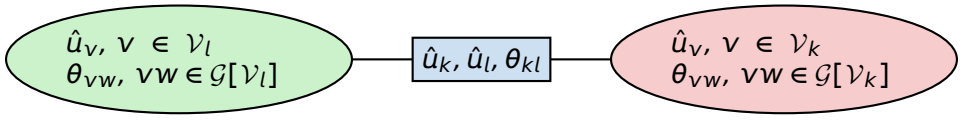
Power flows s_{vw} for their part are given by the angle difference θ_{vw} and voltage magnitudes $\hat{u}_v$,$\hat{u}_w$, i.e.

$$s_{vw} = \hat{u}_v y_{vw}{}^H (\hat{u}_v - \hat{u}_w e^{-j\theta_{vw}}). \tag{4.2}$$

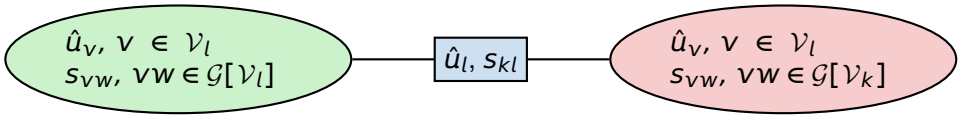
Since the variables voltage magnitude $\hat{u}_v$, active power p_{vw}, and reactive power flow q_{vw} uniquely determine the voltage magnitude $\hat{u}_w$ and angle difference θ_{vw} by the equation system (4.2), these could alternatively be used as separating, common variables. This set of variables will allow for a more intuitive reduction of the number of coupling subgrid variables later on.



(a) Edge Scope



(b) Separators ($\hat{u}_k$, $\hat{u}_l$, θ_{kl})



(c) Separators ($\hat{u}_l$, s_{kl})

Figure 4.1: Illustration of the subgrids $\mathcal{G}[\mathcal{V}_l]$ and $\mathcal{G}[\mathcal{V}_k]$ with the disconnecting edge $\{kl\} \in \mathcal{E}$ (a) and the enlarged clique graph based on AC power flow equations with the two equivalent separators (b) and (c).

Let $\{lk\} \in \mathcal{E}$ be a line that separates the grid $\mathcal{G}(\mathcal{V}, \mathcal{E})$ into two disconnected grids $\mathcal{G}[\mathcal{V}_l]$ and $\mathcal{G}[\mathcal{V}_k]$, where $\mathcal{V}_{l/k}$ are all nodes attainable from l/k without passing k/l and $\mathcal{G}[\mathcal{V}_{l/k}]$ are the subgrids of nodes $\mathcal{V}_{l/k}$ and all edges connecting these nodes. Then, for a fixed power flow on line lk, given by $\{\hat{u}_k, \hat{u}_l, \theta_{kl}\}$ or $\{\hat{u}_l, s_{kl}\}$, the node injection equation (4.1) for $v \in \mathcal{V}_{l/k}$ as well as the flow equations (4.2) for $\{vw\} \in \mathcal{E}[\mathcal{V}_{l/k}]$ depend only on variables in the respective subgrid $\mathcal{G}[\mathcal{V}_{l/k}]$, depicted in Figure 4.1. Thus, the solution of the VVC problem restricted to $\mathcal{G}[\mathcal{V}_l]$ and $\mathcal{G}[\mathcal{V}_k]$ are conditioned on the separator variables independent of each other. This proves the following theorem that is without loss of generality formulated for $\mathcal{G}[\mathcal{V}_k]$:

Theorem 4.1.1. *The optimal solution of the subproblem of $\mathcal{G}[\mathcal{V}_k]$ conditioned on the set $\{\hat{u}_l, s_{kl}\}$ or the set $\{\hat{u}_k, \hat{u}_l, \theta_{kl}\}$ contributes to an optimal overall solution given whatever values $\{\hat{u}_l, s_{kl}\}/\{\hat{u}_k, \hat{u}_l, \theta_{kl}\}$ arise from the optimal solution of $\mathcal{G}[\mathcal{V}_l]$.*

The three common variables, $\{\hat{u}_k, \hat{u}_l, \theta_{kl}\}/\{\hat{u}_l, s_{kl}\}$, between $\mathcal{G}[\mathcal{V}_l]$ and $\mathcal{G}[\mathcal{V}_k]$ requires a cost-to-go function J_k of three variables to incorporate the optimal $\mathcal{G}[\mathcal{V}_k]$ objective into $\mathcal{G}[\mathcal{V}_l]$. Three-dimensional functions are hard to computationally evaluate and represent. However, for distribution grids without actively controllable Distributed Energy Resources (DER), the active p_{lk} and reactive q_{lk} flow resulting from the deterministic feed-ins p_v and q_v, $v \in \mathcal{V}_k$, is only modified by the voltages $\mathbf{u}$ dependent losses. For $u_l = \hat{u}_l$ and known p_v and q_v, $v \in \mathcal{V}_k$, the voltages $\mathbf{u}$ are uniquely determined by the AC power flow equations (2.17) .Thus, for $\hat{u}_l$ reactive $Q_k(\hat{u}_l)$ and real power $P_k(\hat{u}_l)$ flow are determined. These functions set the power flow s_{lk} and reduce the coupling variables to $\hat{u}_l$, see separator in Figure 4.1 (c). Moreover, the optimal solution of $\mathcal{G}[\mathcal{V}_k]$ can be incorporated in the subproblem of $\mathcal{G}[\mathcal{V}_l]$ by a pseudo generator at node l with costs function $J_k(\hat{u}_l)$ for generation $Q_k(\hat{u}_l)$ and $P_k(\hat{u}_l)$.

Hence, the dynamic programming strategy to solve the VVC problem, (2.16) subject to (2.17) and (2.18), is to solve first the right and then the left of the following problems

$$\begin{aligned} \min_{\substack{\mathbf{u}\in\mathbb{C}^{|\mathcal{V}_l|}\\ \tau\in\mathcal{T}_l}} \quad & \sum_{v\in\mathcal{V}_l} c_v(s_v, u_v) + J_k(\hat{u}_l), \\ s.t. \quad & \mathbf{s} = \mathbf{u}\odot(\mathbf{Y}_l(\tau)\,\mathbf{u})^H, \\ & s_{lk} = P_k(\hat{u}_l) + jQ_k(\hat{u}_l), \end{aligned} \qquad \text{and} \qquad \begin{aligned} J_k(\hat{u}_l) = \min_{\substack{\mathbf{u}\in\mathbb{C}^{|\mathcal{V}_k|}\\ \tau\in\mathcal{T}_k}} \quad & \sum_{v\in\mathcal{V}_k\setminus\{l\}} c_v(s_v, u_v) \\ s.t. \quad & \mathbf{s} = \mathbf{u}\odot(\mathbf{Y}_k(\tau)\,\mathbf{u})^H, \\ & \hat{u}_l = |u_l|, \end{aligned}$$

here $\odot$ is the component-wise multiplication and $\cdot^H$ the complex conjugate transpose. The node l gives the coupling between the subproblems in this formulation. It serves as a slack node with fixed voltage magnitude in the right problem and as a pseudo generator that mirrors the subgrid $\mathcal{G}[\mathcal{V}_k]$ in the left problem. In general, for tree-structured subgrids' clique graphs, a graph where the subgrids $\mathcal{G}[\mathcal{V}_l]$ are nodes and edges $\{lk\}$ are given by $\mathcal{V}_l \cap \mathcal{V}_k \neq \emptyset$ for this example depicted in Figure 4.1 (b)/(c), an order exists in which the subproblems can be solved such that the pseudo generator functions J_k,P_k and Q_k are known when solving VVC subproblem on $\mathcal{G}[\mathcal{V}_l]$.

Thus far, the basis for a dynamic programming formulation has been derived without considering its motivation, i.e. the exponential rise in complexity for multiple discrete variables. Nevertheless, the optimality property in Theorem 4.1.1 based on the grid separating line $\{lk\}$ holds for any line that separates a grid or subgrid and, thus, can be applied multiple times to generate subgrids with at most one discrete variable. The subproblems of one discrete variable are solvable in linear time with regard to the discrete variable's feasible set, e.g. by full enumeration.

In such a subgrid $\mathcal{G}[\mathcal{V}_k]$ of one discrete decision variable $\tau_k \in \mathcal{T}_k$, the optimal discrete decision τ_k^* is solely dependent on $\hat{u}_k$ since it implies $\mathbf{u}$, i.e. there exists a optimal discrete decision map $T_k(\hat{u}_l)$. For a transformer, the direct influence on the voltage regime sets the tap τ_k^* that steps the voltage as high as possible without violating the voltage band. The capacitor banks influence is twofold: either by the reduction of the reactive power flows it reduces the voltage magnitude differences and enlarges the voltage angle differences or it increases the voltage magnitude differences drop and reduces the voltage angle differences. Regardless of the discrete controllable equipment type, the discrete variable only needs to be switched if the voltage band limit is approached in order to stay within the voltage band or a higher voltage regime becomes feasible within the voltage band. Moreover, the voltage band hinders this discrete value from recurring. Thus, the optimal discrete decision map $T_k(\hat{u}_l)$ is a monotone step function.

The step nature of the optimal discrete decisions $T_k(\hat{u}_l)$ is due to its influence on losses directly reflected in the pseudo generator functions $J_k(\hat{u}_l)$, $Q_k(\hat{u}_l)$ and $P_k(\hat{u}_l)$. This causes these functions to be defined only piece-wise, e.g. in Figure 4.7. The distortions that arise from various nodes with these piece-wise functions may conflict with an optimal subproblem solution. Thus, each subgrid should contain at most one coupling node l, i.e. a node with assigned pseudo generator functions. But with the nodal equation (4.1), a coupling node l can additively aggregate more than one set of pseudo generator functions. Due to the trade-off between the number of subproblems and their size, the subgrids should be as large as possible while spanning just over one discrete decision variable and aggregating node l.

4.2 Implementation

In line with the theoretical derivation in Section 4.1, a dynamic programming algorithm is proposed in this section. The outline of this algorithm, Algorithm 2, is as follows: (1) divide the overall problem into subproblems, Section 4.2.1, (2) sequentially solve all subproblems, Section 4.2.2, and (3) reassemble the overall solution from the subproblems, Section 4.2.3.

First, a method is deduced that decomposes the electric graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ into subgrids whose corresponding continuous VVC problem contains at most one discrete decision variable and at most one coupling node to other subgrids. This is done with regard to a root node, i.e. the highest voltage leaf node here[1]. Then, based on the thus derived subgrid structure, the subproblem for one subgrid is formulated. Second, an efficient solution procedure for such subproblems is derived. Based on the monotonicity of the optimal

[1] This is the convention here but it could be in general any node.

Algorithm 2 Overall Algorithm

Input: $\mathcal{G}(\mathcal{V}, \mathcal{E})$ Grid

```
// Split the overall problem into subproblems
{G[V_l]}_{l∈L} ←Grid2Subgrids(G(V, E), D) (Algorithm 3)
// Solve l ∈ L if all {T_k, J_k, P_k, Q_k, U_k}_{k∈K(l)} are known
Cons(L) ← 0
while Not all subgrids l ∈ L calculated do
  for {l ∈ L : (V_l ∩ (∪_{k∈K(l)∧Cons(k)=0} V_k)) = ∅} do
    // Determining {τ_l, J_l, P_l, Q_l, U_l} supporting points by reduced enumeration
    (T̂, Ĵ, P̂, Q̂, Û) ← BI-TAP({G[V_k]}_{k∈K(l)∪l}, n) (Algorithm 4)
    // Find approximation {T_l, J_l, P_l, Q_l, U_l} piece-wise functions
    (T_l, J_l, P_l, Q_l, U_l) ← AP-PIECE(T̂, Ĵ, P̂, Q̂, Û) (Quadratic Approximation)
    // Add supporting points by halving the largest n − N intervals
    Cons(l) ← 1
  end for
end while
// Reassemble solution for G(V, E) from the subproblem solutions
(τ, u) ← RE-TAP({T_l, J_l, P_l, Q_l, U_l}_{l∈L})(Figure 4.4)
```

Output: τ, $\mathbf{u}$

discrete variable in a subproblem family parameterized by the voltage magnitude $\hat{u}$, the to consider discrete variables for the solution of one problem in this family can be significantly reduced. Solutions of this subproblem family deliver the supporting points that serve as basis for the piece-wise approximation, i.e. pieces with equal optimal discrete variable, of the coupling pseudo generator functions. Finally, the overall solution is re-obtained from the successively solved subproblems in reverse order, i.e. from the root to the leaf nodes.

4.2.1 From the Overall Problem to the Subproblems

The subproblems, in this section, are given by subgrids $\mathcal{G}[\mathcal{V}_l]$ induced by a subset of nodes $\mathcal{V}_l \subset \mathcal{V}$, $l \in \mathcal{L}$. For the subset $\mathcal{V}_l$ the fundamental requirements are derived that imply a valid set of subgrids $\mathcal{G}[\mathcal{V}_l]$, $l \in \mathcal{L}$. Then, an algorithm is proposed to obtain such subgrids with at most one discrete variable and at most one coupling node as argued in Section 4.1. Finally, the subproblem of one subgrid is exemplarily formulated.

All subgrids $\{\mathcal{G}[\mathcal{V}_l]\}_{l\in\mathcal{L}}$ need to be connected. Hence each subgrid $\mathcal{G}[\mathcal{V}_l]$ needs to have at least one common node with another subgrid, i.e. there must exist $k \in \mathcal{L} \setminus \{l\}$ such that $|\mathcal{V}_l \cap \mathcal{V}_k| \geq 1$. The inducing sets $\mathcal{V}_l$, $l \in \mathcal{L}$, need to be a cover of $\mathcal{V}$, i.e.

$$\cap_{l\in\mathcal{L}} \mathcal{V}_l = \mathcal{V},$$

and pairwise overlapping in at most one node, $|\mathcal{V}_l \cap \mathcal{V}_k| \leq 1$ for all $l \neq k$, $l, k \in \mathcal{L}$, to minimize the number of variables covered several times. These fundamental properties ensure that grid $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is covered by connected subgrids $\mathcal{G}[\mathcal{V}_l]$, $l \in \mathcal{L}$, with minimal overlap for radial grids. In general, such a restriction to radial grids, is not necessary but eases the formulation. For loopy grids, in addition, each loop needs to be contained in

only one subset $\mathcal{V}_l$, since then also all edges are covered by the $\mathcal{G}[\mathcal{V}_l]$ edges and the subgrids clique graph[2] is of tree shape. This requirement can contradict the at most one discrete variable and at most one coupling node argument. Since most distribution grids are operated as radial grids, the loopy case is not considered in the following.

As requirement of Section 4.1, each subgrid $\mathcal{G}[\mathcal{V}_l]$ should contain at most one discrete variable and at most one coupling node, i.e. a node contained in at least one subgrid $\mathcal{G}[\mathcal{V}_k]$ further from the root node. The coupling nodes serve as slack nodes in the primary subgrid, namely the subgrid further from the root node, and as coordinating pseudo generators aggregating the optimal cost-to-go $J_k(\hat{u}_l)$, reactive $Q_k(\hat{u}_l)$, and real power $P_k(\hat{u}_l)$ of all downstream subgrids $\mathcal{K}(l)$ in the secondary subgrid.

Subgrids $\mathcal{G}[\mathcal{V}_l]$, $l \in \mathcal{L}$, satisfying the above mentioned requirements are obtained based on two observations – branching nodes and critical number of discrete decision entries in $\mathcal{G}(\mathcal{V}, \mathcal{E})$. Branching nodes, i.e. nodes with more than one child $\mathcal{C}(v)$ (neighbors further from the root node), are able to serve as primary node for more subgrids and bundle pseudo generator functions at a single node. Moreover, at least two discrete decision entries need to be downstream from a branching node in order to make it necessary to split the downstream subgrid into different subgrids of one discrete decision variable. Thus, the relevant set of branching nodes is

$$\mathcal{Y}^* = \{v \in \mathcal{V} : |\mathcal{C}(v)| \geq 2 \wedge |\{t \in \mathcal{D} : dt(t) \geq dt(v)\}| \geq 2\}, \tag{4.3}$$

where $dt(v)$ denotes the discovery time of node $v \in \mathcal{V}$ in the first depth search started at the root node and $dt(\{vw\})$, $\{vw\} \in \mathcal{E}$, denotes the lower discovery time of v and w. In case that on the path between two discrete variables no branching node $y \in \mathcal{Y}^*$ is situated then an additional node s is needed to separate the path into two of one discrete variable. Thus, the branching node set is augmented by the set $\mathcal{S}$ of these additional nodes to

$$\mathcal{Y} = \mathcal{Y}^* \cup \mathcal{S}. \tag{4.4}$$

These sets are illustrated by colored nodes in Figure 4.2.

Based on $\mathcal{Y}$ the subgrids $\mathcal{G}[\mathcal{V}_l]$, $l \in \mathcal{L}$, can be obtained as detailed in Algorithm 3 and depicted in Figure 4.2. Each subgrid $\mathcal{G}[\mathcal{V}_l]$ is determined by a primary node $y(l) \in \mathcal{Y}$ and a child node $c(l) \in \mathcal{C}(y(l))$ with at least one discrete variable or pseudo generator in the reachable downstream $c(l)$ subgraph, i.e. a graph implied by all nodes downstream from $c(l)$ and reachable without passing a node in $\mathcal{Y}$. So the subgrid $\mathcal{G}[\mathcal{V}_l]$ defining set of nodes $\mathcal{V}_l$ is given by all nodes reachable from $c(l)$ without passing a node in $\mathcal{Y}$ and if $c(l)$ has the lowest discovery time of these nodes additionally all nodes reachable in the same manner from the nodes in $\mathcal{C}(y)$ without reachable discrete variable and pseudo generator. Thus, the subgrid $\mathcal{G}[\mathcal{V}_l]$ spans between at most two nodes in $\mathcal{Y}$, namely its primary node $y(l)$ and possible a pseudo generator node $\hat{y}(l)$. Moreover, each subgrid $\mathcal{G}[\mathcal{V}_l]$ contains by definition of $\mathcal{Y}$ no more than one discrete variable.

[2]The clique graph for the two subgrids $\mathcal{G}[\mathcal{V}_l]$ and $\mathcal{G}[\mathcal{V}_k]$ is shown in Figure 4.1.

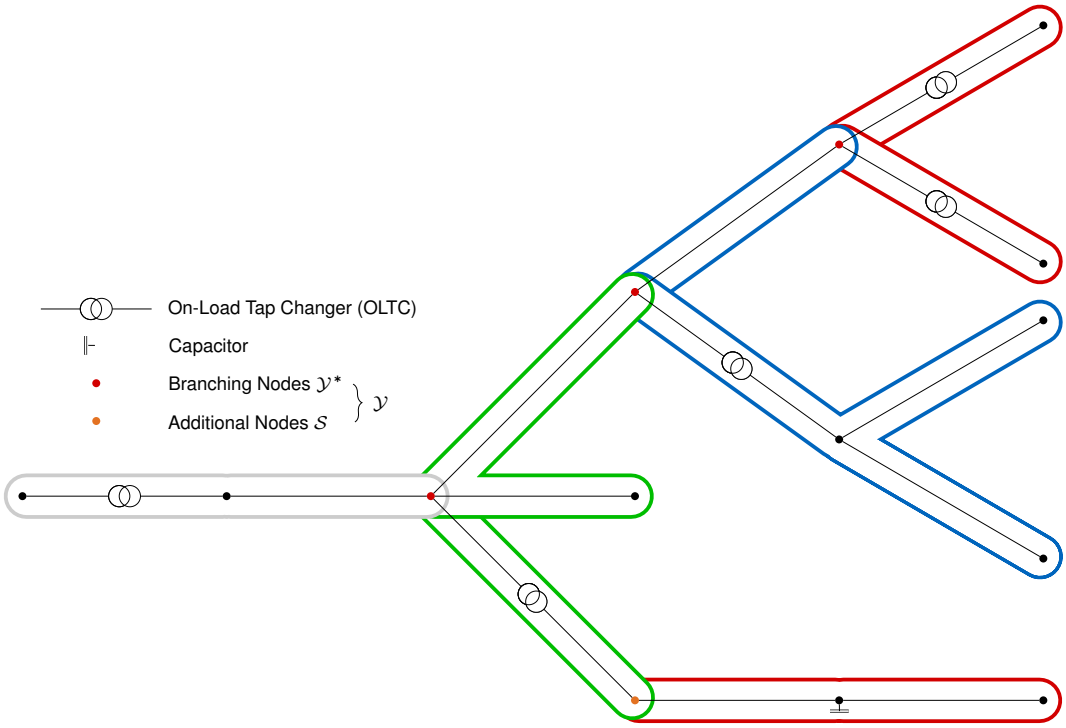


Figure 4.2: Resulting subgrids obtained by Algorithm 3. The colors indicate the iteration the subgrid is added, first red, second green, third blue and the root node grey.

Algorithm 3 Grid2subgrids

Input: $\mathcal{G}(\mathcal{V}, \mathcal{E})$, $\mathcal{Y}$ and $\mathcal{D}$

// Build a graph of branching nodes $\mathcal{Y}$
$G(\mathcal{V}_\mathcal{Y}, E_\mathcal{Y})$
// Set considered flag
$Con(\mathcal{Y}) \leftarrow 0$, $dt \leftarrow$ discovery time from first depth search
$ft \leftarrow$ finish time from first depth search
while $\exists\, y(l) \in \mathcal{Y} : Con(y(l)) = 0 \wedge \forall\, Con(\mathcal{C}_\mathcal{Y}(y)) = 1$ **do**
  // Generate subgrids for all children of $y(l)$
  $Con(y) \leftarrow 1$
  for $c(l) \in \mathcal{C}(y(l)) \wedge \exists d \in \mathcal{D} : dt(d) > dt(c(l))$ or $y \in \mathcal{Y} : dt(y) > dt(c(l))$ **do**
    // get nodes
    $\mathcal{V}_l \leftarrow \{v \in \mathcal{V} : dt(v) \geq dt(y) \wedge ft(v) \leq ft(y)\}$
    // updated discovery and finish time measures
    $dt(\mathcal{V}_l) \leftarrow -\inf, ft(\mathcal{V}_l) \leftarrow \inf$
  end for
end while
// Add root node subgrid
$\mathcal{V}_r \leftarrow \{v \in V : dt(v) \geq dt(r) \wedge ft(v) \leq ft(r)\}$

Since each $y(l) \in \mathcal{Y}$ is contained in at least two subgrids, it is employed to interrelate the respective subproblems. In $y(l)$'s primary subgrids it serves as slack node with fixed voltage magnitude $\hat{u}_{y(l)}$. In the other subgrid $\hat{y}(l)$ gives a pseudo generator with costs-to-go $J_k(\hat{u}_{\hat{y}(l)})$ for a throughput $Q_k(\hat{u}_{\hat{y}(l)})$ and $P_k(\hat{u}_{\hat{y}(l)})$ that results from its primary subgrid $\mathcal{G}[\mathcal{V}_k]$, $k \in \mathcal{K}(l)$, where

$$\mathcal{K}(l) = \{k \in \mathcal{L} \setminus \{l\} : (\mathcal{V}_k \cap \mathcal{V}_l) \setminus \{y(l)\} \neq \emptyset\}. \tag{4.5}$$

These functions are given by the optimal objective value, reactive and real power in $\hat{y}(l)$'s primary subgrid problem. The subproblem for the subgrid $\mathcal{G}[\mathcal{V}_l]$ with slack node $y(l) \in \mathcal{Y}$ and a discrete variable feasibility set $\mathcal{T}_l$ reads as

$$J_l(\hat{u}) = \min_{\substack{\mathbf{u} \in \mathbb{C}^{|\mathcal{V}_l|} \\ \tau_l \in \mathcal{T}_l}} \quad \sum_{v \in \mathcal{V}_l \setminus \hat{y}(l)} c_v(s_v, u_v) + \sum_{k \in \mathcal{K}(l)} J_k\left(\hat{u}_{\hat{y}(l)}\right), \tag{4.6}$$

$$s.t. \quad \mathbf{s} = \mathbf{u} \odot (\mathbf{Y}_l(\tau_l)\,\mathbf{u})^H, \tag{4.7}$$

$$s_{\hat{y}(l)} = \sum_{k \in \mathcal{K}(l)} P_k\left(\hat{u}_{\hat{y}(l)}\right) + j\,Q_k\left(\hat{u}_{\hat{y}(l)}\right), \tag{4.8}$$

$$\hat{u} = u_{y(l)}, \tag{4.9}$$

and defines by $s_{y(l)}$ the functions $Q_l(\hat{u}_{y(l)})$ and $P_l(\hat{u}_{y(l)})$. Moreover, this solution defines the discrete variable function $T_l(\hat{u}_{y(l)})$ by $\operatorname{argmin}_{\tau_l \in \mathcal{T}_l} J_l(\hat{u}_{y(l)})$ and voltage magnitude function $U_l(\hat{u}_{y(l)})$ by $\hat{u}_{\hat{y}(l)}$, which is necessary to re-obtain the overall solution in Section 4.2.3. Since physical flows depend solely on angle differences the voltage angles offset of the primary node, i.e. assigning any angle in constraint (4.9), leaves the optimal power flows unchanged.

For tree-structured subgrid clique graphs an order exists in which the subproblems l can be solved such that all $\mathcal{K}(l)$ pseudo generator functions, namely cost-to-go $J_k\left(\hat{u}_{\hat{y}(l)}\right)$, active power $P_k\left(\hat{u}_{\hat{y}(l)}\right)$, and reactive power $Q_k\left(\hat{u}_{\hat{y}(l)}\right)$, are known when the problem is solved. This order starts at leaf subproblems, i.e. subproblems $l \in \mathcal{L}$ with $\mathcal{K}(l) = \emptyset$, and then sequentially solves the subproblem l when all predecessor subproblems $\mathcal{K}(l)$ are solved. Moreover, radial grids $\mathcal{G}(\mathcal{V}, \mathcal{E})$ pass their tree-structure to the subgrids $\mathcal{G}[\mathcal{V}_l]$, $l \in \mathcal{L}$, clique graph.

Since for each subproblem l the pseudo generator functions integrate the optimal results for the voltage magnitude $\hat{u}_{\hat{y}(l)}$ of the subgrids $\mathcal{K}(l)$, the proposed approach results in an optimal overall solution.

4.2.2 Solve the Subproblems

In the following, the solution of the problem (4.6), subject to (4.7), (4.8) and (4.9), is derived. This problem is a family of problems parametrized by the voltage magnitude $\hat{u}_{y(l)} \in [\underline{\hat{u}}_{y(l)}, \overline{\hat{u}}_{y(l)}]$ of the primary node $y(l)$. Solving or representing the results of this family of problems for each continuous parametrization variable $\hat{u}_{y(l)}$ is in general computationally intractable. However, the monotonous behavior of the optimal discrete variable $T_l(\hat{u}_{y(l)})$ and its influence on the problem enables an efficient computational evaluation and representation of the functions J_l, P_l, Q_l and U_l, as detailed in the following.

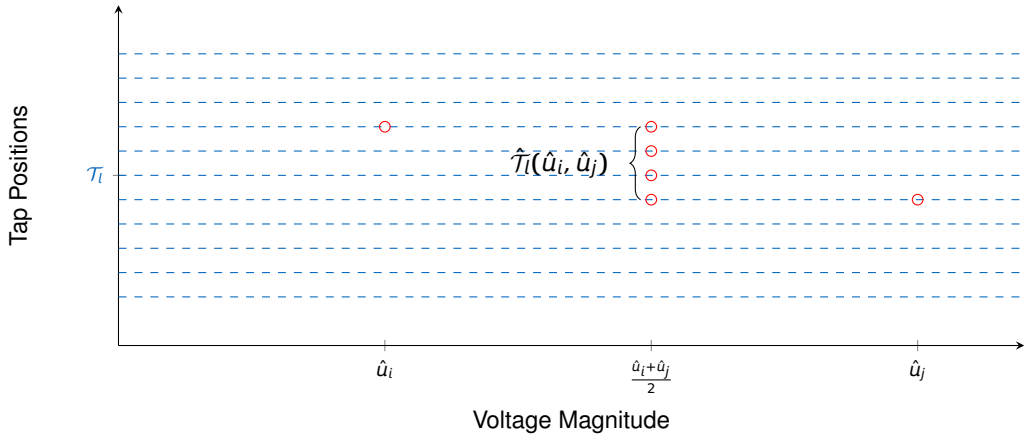


Figure 4.3: Restriction of the feasible tap positions based on $T_l(\hat{u}_i)$ and $T_l(\hat{u}_j)$.

For each subgrid l the optimal discrete variable $T_l(\hat{u})$ is a monotonous step function, as in Figure 4.7. The discrete variable τ_l needs to be switched only if the voltage band is approached or a higher voltage regime turns attainable within the voltage band for $\hat{u}_l$. The dependence on the voltage band also hinders the discrete variable from recurring, which leads to a monotonously changing optimal discrete variable. This monotonicity restricts the discrete variable candidates between two voltage magnitudes $\hat{u}_i$ and $\hat{u}_j$ to

$$\hat{\mathcal{T}}_l(\hat{u}_i, \hat{u}_j) = \{\tau_l \in \mathcal{T}_l : T_l(\hat{u}_i) \geq \tau_l \geq T_l(\hat{u}_j) \vee T_l(\hat{u}_i) \leq \tau_l \leq T_l(\hat{u}_j)\}, \tag{4.10}$$

which are the discrete variables in $\mathcal{T}_l$ between $T_l(\hat{u}_i)$ and $T_l(\hat{u}_j)$, illustrated in Figure 4.3.

On the one hand, the monotonous step nature decreases the number of discrete values τ_l to be considered for $\hat{u}_{y(l)}$. Starting with the voltage band $[\underline{\hat{u}}_{y(l)}, \overline{\hat{u}}_{y(l)}]$, the voltage magnitudes at which the tap position changes are obtained by bisection, i.e. by halving each interval $[\hat{u}_i, \hat{u}_j]$ for which $|\hat{\mathcal{T}}_l(\hat{u}_i, \hat{u}_j)| \neq 1$. Thus, each such evaluation of problem (4.6) subject to (4.7), (4.8) and (4.9) for $\hat{u} = (\hat{u}_j + \hat{u}_i)/2$ is first reduced to a full enumeration of $\hat{\mathcal{T}}_l(\hat{u}_i, \hat{u}_j)$ and finally further reduced to a power flow problem with known $T_l(\hat{u})$. This procedure is detailed in Algorithm 4.

Algorithm 4 BI-TAP

Input: $\{\mathcal{G}[\mathcal{V}_k]\}_{k\in K(l)\cup l}$, n

// Initialize discrete variables $T(u)$, $u \in [\underline{\hat{u}}_{y(l)}, \overline{\hat{u}}_{y(l)}]$

$(\hat{T}_{1/2}, \hat{J}_{1/2}, \hat{P}_{1/2}, \hat{Q}_{1/2}, \hat{U}_{1/2}) \leftarrow$ FI-TAP$(u, \mathcal{T}_l, \{\mathcal{G}[\mathcal{V}_k]\}_{k\in K(l)\cup l})$, $N \leftarrow 2$

// Get supporting points by bisection and reduced enumeration

while $\max(N) < n$ **do**

  for All neighboring $\hat{u}_i, \hat{u}_j$ with $\tau(\hat{u}_i) \neq \tau(\hat{u}_j)$ **do**

    $N \leftarrow N + 1$, $\hat{u}_N \leftarrow (\hat{u}_i + \hat{u}_j)/2$

    $(\hat{T}_N, \hat{J}_N, \hat{P}_N, \hat{Q}_N, \hat{U}_N) \leftarrow$ FI-TAP$(\hat{u}, \hat{\mathcal{T}}_l(\hat{u}_i, \hat{u}_j), \{\mathcal{G}[\mathcal{V}_k]\}_{k\in K(l)\cup l})$ (Algorithm 5)

  end for

end while

Output: $\hat{T}, \hat{J}, \hat{P}, \hat{Q}, \hat{U}$

Algorithm 5 FI-TAP

Input: $\hat{u}$, $\mathcal{T}_l$, $\{\mathcal{G}[\mathcal{V}_k]\}_{k\in K(l)\cup l}$

Solve (4.6), subject to (4.7), (4.8) and (4.9).

$J \leftarrow J_l(\hat{u})$, $\tau \leftarrow \operatorname{argmin}_{\tau_l\in\mathcal{T}_l} J_l(\hat{u})$, $P \leftarrow real(s_{v_l})$, $Q \leftarrow imag(s_{v_l})$, $U \leftarrow \hat{u}_{\hat{y}(l)}$

Output: τ, J, P, Q, U

On the other hand, the step-wise-nature of $T_l(\hat{u})$ is reflected in the cost-to-go $J_l(\hat{u})$, the optimal active $P_l(\hat{u})$, and reactive $Q_l(\hat{u})$ power, see Figure 4.7, since $\hat{u}$ and $T_l(\hat{u})$ influence the losses. The same holds true for the directly influenced $U_l(\hat{u})$. The pseudo generator functions are obtained piece-by-piece for the $T_l(\hat{u})$ steps. For each $T_l(\hat{u})$ step the best maximal quadratic approximation onto the supporting points which have been obtained by bisection represents the pseudo generator function piece, see dotted lines in Figure 4.7. The small physically attainable scope of $\hat{u}_{\hat{y}(l)}$ given $\hat{u}_{y(l)}$ resolves the contradiction of the piece-wise non-continuous functions and the optimality of the solution

4.2.3 Reassemble the overall Solution

This section details the reassembling of the overall solution given the subproblem solutions T_l, U_l, J_l, P_l and Q_l, $l \in \mathcal{L}$. It relies on a mapping between the secondary node and primary node that arises from the subgrid structure. This mapping starts at the root node subgrid $\mathcal{G}[\mathcal{V}_r]$ and progresses to the leaf subgrid nodes.

The voltage magnitude $\hat{u}_r$ at the root node is set either to $\hat{u}_r = 1$ or to the minimum argument of $J_r(\hat{u})$. Hence, the discrete decision $T_r(\hat{u}_r)$ and the voltage magnitude $U_r(\hat{u}_r)$ at the secondary node $\hat{y}(r)$ are fixed. This secondary node voltage magnitude $U_r(\hat{u}_r)$ sets consequently the voltage magnitude in its primary subgrid as well as the discrete decision variable and the secondary node voltage magnitude. Thus, the optimal value of $\mathcal{G}[\mathcal{V}_l]$ propagates the setting to $\mathcal{G}[\mathcal{V}_k]$, $k \in \mathcal{K}(l)$, see Figure 4.4, and the process stops when $\mathcal{K}(l) = \emptyset$. The angles could be included by consecutively adding the angle arising in U_l, since flows are solely dependent on angle differences.

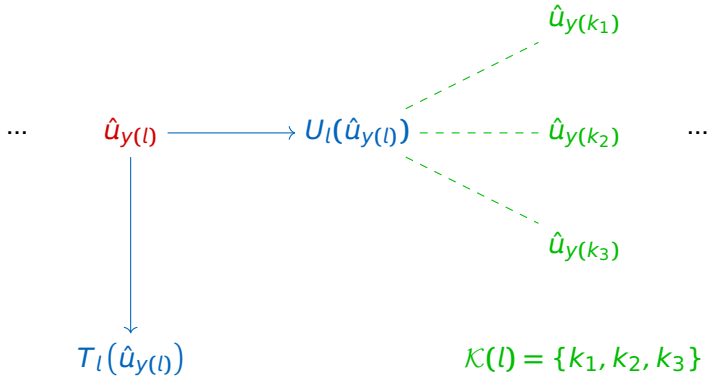


Figure 4.4: Propagation scheme to reassemble the overall solution based on $\mathcal{G}[\mathcal{V}_l]$ and the functions $T_l(\hat{u}_l)$ and $U_l(\hat{u}_l)$.

4.3 Experiments

The dynamic programming framework for the VVC problem is tested with two experiments in this section. A stylized distribution grid highlights the differences between the dynamic programming approach and other commonly used approaches, Section 4.3.1. Section 4.3.2 shows the ability of the approach to remedy certainly and optimally overvoltages in highly DER penetrated authentic future distribution grids.

4.3.1 Simplified Distribution Grid

The first example is a very concentrated exemplary distribution grid, shown in Figure 4.5. This distribution grid consists of a Medium Voltage (MV) branch and six diverging Low Voltage (LV) grids, which are connected to the MV branch by an OLTC. Each LV grid has an aggregated feed-in of 0.15 MW assigned to a single node and a line connecting this node to its MV/LV OLTC with equal impedance $z_{LV} = 3.125 - j0.5$. Also, the six line segments on the MV branch possess the same impedance $z_{MV} = 0.027 - j0.025$. The OLTC with tap positions $\{1.03, 1.02, \ldots, 0.97\}$ are of two types: the High Voltage (HV) to MV (110kV/30kV) transformer with 26 MVA capacity and MV to LV (20kV/0.4kV) transformer with 1 MVA capacity. The Line Drop Compensation (LDC) inspired method and hill climbing (HILL) method are considered as baselines for a comparison to proposed dynamic programming (DYN) approach.

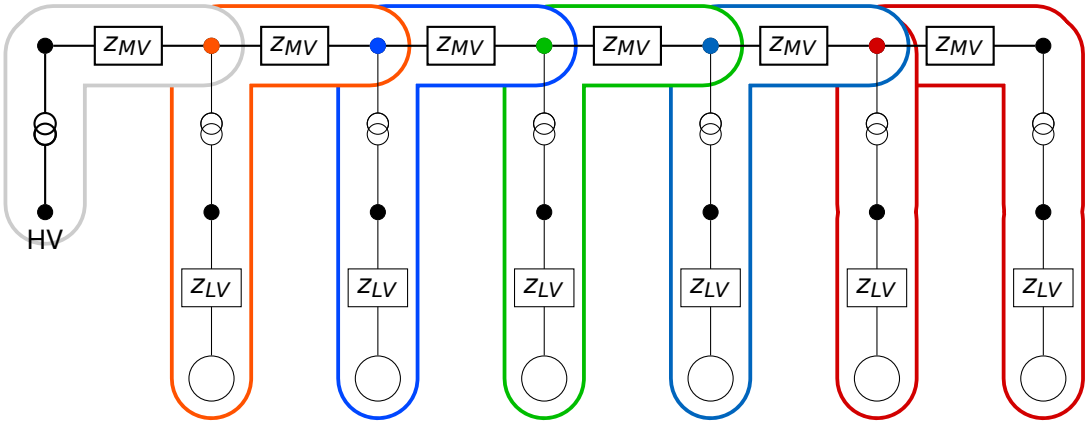


Figure 4.5: Simplified distribution grid of one MV branch with six LV branches and its seven subgrids implied subproblems. The subgrids are shown by the colored frames around the nodes and edges. The primary nodes $y(l)$ are colored in the respective subgrids color. These nodes serve as pseudo generator in the subgrid in the different colored frame around them.

The LDC approach measures the voltage u_l and current flow i_l at the lower voltage/ secondary side of the transformer [51, 47]. On basis of these measurements and an aggregated feeder impedance $\hat{z}_l = \hat{r}_l + j\hat{x}_l$ the voltage magnitude $\hat{u}_{end}$ at the feeder end is estimated as follows $\hat{u}_{end} \approx \hat{u}_l - (\hat{x}_l Im(i_l) + \hat{r}_l Re(i_l))$. In this example the feeder impedance $\hat{z}_l$ of the MV feeder is derived from the sum of the six line impedances times a factor that adjusts for the flow reduction caused by the in between dissecting LV branches [35], while the feeder estimate for the LV grids is the line impedance itself. According to [51] the transformers are set in the following order: first the HV/MV OLTC such that the voltage rise/drop on MV feeder stays within the voltage band in a voltage regime as high as possible and then the medium to MV/LV OLTCs are set such that the voltage rise/drop on their secondary side still satisfies the voltage band in a high voltage regime.

A method based on full knowledge instead of aggregation based estimates is the HILL approach [6, 91]. Here starting at the initial tap position the next higher and the next lower tap position are examined and the tap with the lowest objective value fixed. These comparisons are successively done for each OLTC in an order equal to the LDC approach, as proposed in [6] to reduce the number of switches, and is started over until the objective function does no longer improve. The quality of this approach's result strongly depends on the initial tap positions. When all initial taps are set to position 1, HILL1, the resulting OLTC tap positions are not optimal, while starting at 0.97, HILL0.97, the resulting taps are optimal, see Figure 4.6 and Table 4.1. In the HILL1 case the optimal higher tap at the HV/MV OLTC disagrees with the voltage band and thus the tap is not change, and the ex post set MV/LV OLTC have no incentive to lone lower their tap.

The proposed dynamic programming method splits the simplified distribution grid in seven subgrids, see Figure 4.5. These subgrids have been solved by reduced enumeration to obtain the discrete variable switches and the 40 supporting points for the pseudo generator functions. For the third subgrid the resulting supporting points and the there on approximated pseudo generator functions $T_l(\hat{u})$, $J_l(\hat{u})$, $P_l(\hat{u})$ and $Q_l(\hat{u})$ are depict in Figure 4.7. The marginal overall cumulated errors are illustrated by the difference between the recalculated, diamond, and reassembled, cross, power flow voltage, Figure 4.6.

HILL0.97 and the proposed dynamic programming approach found the optimal tap positions, while small inaccuracies in the estimator misled the LDC method step down the first MV/LV by one tap, shown in Figure 4.6, which resulted in slightly higher losses, Table 4.1. The HILL0.97, LDC and DYN method have utilized their additional knowledge to reduce the losses by $1.39 - 1.51\%$ in comparison to the reference feed-in tap setting, i.e. $\tau_{MV/LV} = 0.97$. The HILL1 method does not since the prevalent aim is to satisfy the voltage band starting from the initial position. So the HILL1 results in an overall lower voltage regime, see Figure 4.6, and thus higher losses than the reference feed-in tap setting, see Table 4.1. This shows the draw back of the HILL climbing approach: it might end up in local minima and the effort to obtain the additional information is lost.

Besides the voltage band, which is satisfied by all methods for this extreme feed-in scenario (see Figure 4.6), the voltage spread gives significant details of the grid state [80]. The voltage spread is the difference between the highest and lowest voltage in the

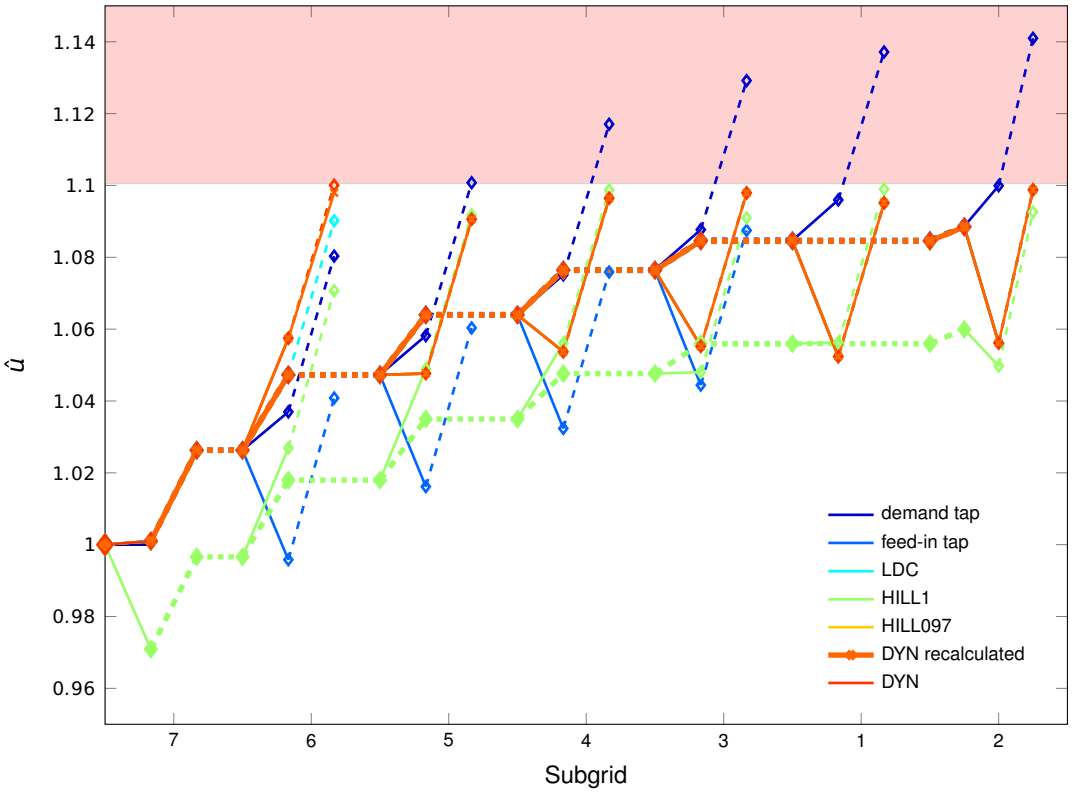


Figure 4.6: Simplified distribution grid node voltage magnitudes' evolution according to the containing subgrid and its topology for the tap positions set to demand tap ($\tau_{MV/LV} = 1.01$), dark blue, feed-in tap ($\tau_{MV/LV} = 0.97$), light blue, LDC, turquoise, HILL1, green, HILL0.97, yellow, the proposed DYN approach and respective recalculated power flow, orange. The line thickness represents the voltage level and dotted lines the link between the subgrids, i.e. coupling variables.

grid. Similar to the losses and the tap positions, the voltage spreads obtained by DYN, HILL0.97 and LPC are alike. This good performance of the LDC approach results from the homogenous feed-in and almost perfect estimator $\hat{z}_l$, which are the assumptions for this abstraction. However, feed-in and demand often fluctuate non homogenous. These fluctuations can be partially integrated in the approach by an in-feed reference cell, equal to the compounding in MV [42]. But that would require detailed measurements to evaluate the best reference cell and estimates.

Table 4.1: Five methods summary of the simplified grid example.

	$\tau_{MV/LV} = 0.97$	LDC	HILL1	HILL0.97	DYN
p_{loss} [MW]	0.0837	0.0826	0.0852	0.0825	0.0825
vs $\tau_{MV/LV} = 0.97$ [%]	0	-1.3947	1.7901	-1.5158	-1.5158
spread	0.1029	0.0988	0.1281	0.1001	0.1001

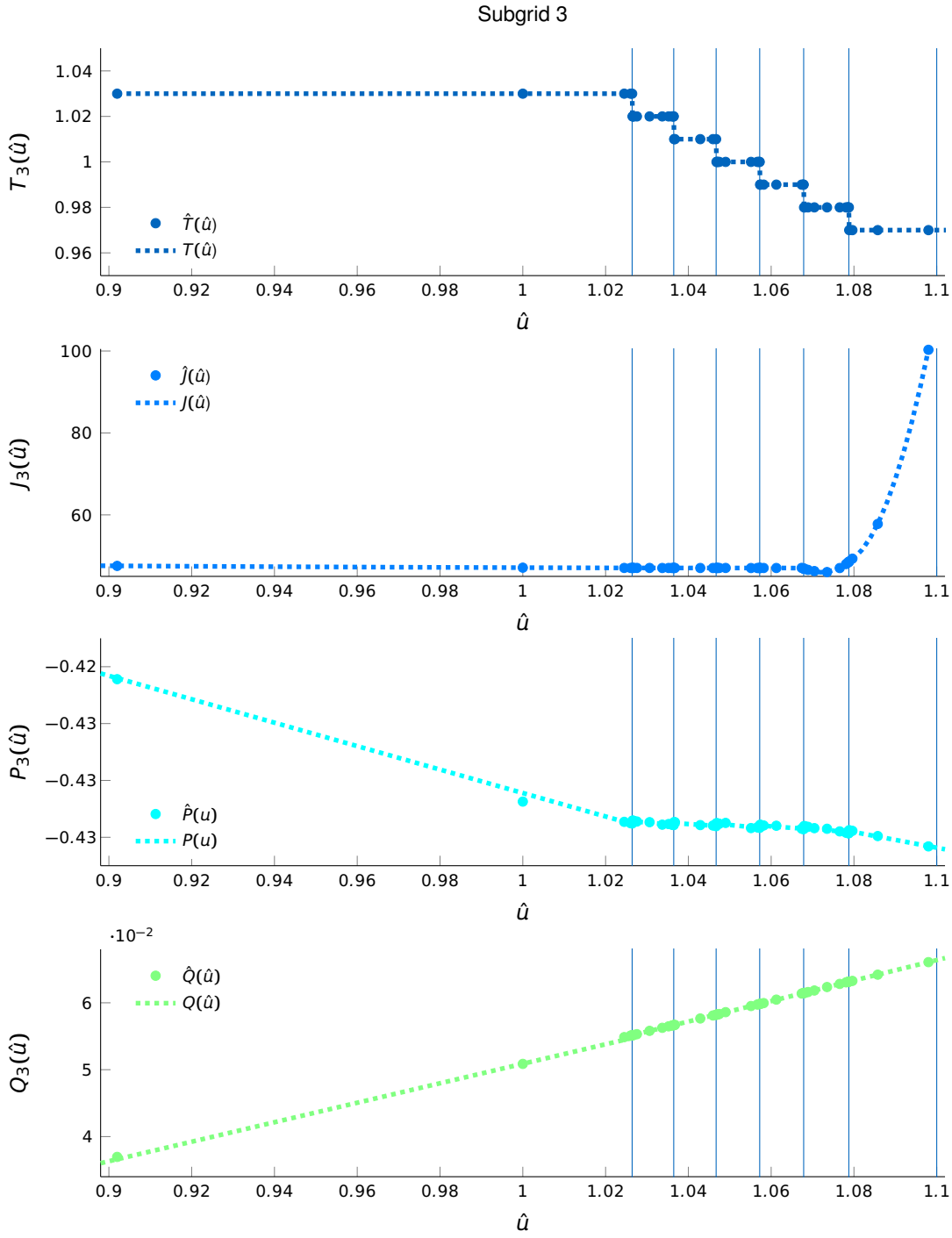


Figure 4.7: Illustration of the functions $\tau_3(\hat{u})$, $J_3(\hat{u})$, $P_3(\hat{u})$ and $Q_3(\hat{u})$ for the third subgrid in the concentrated distribution grid example depicted in Figure 4.5. The points are the supporting points and the dotted lines the piece-wise approximation, which are given by the optimal tap positions, vertical blue lines.

4.3.2 Realistic Medium Voltage-Low Voltage Distribution Grids

Real distribution grids are complex in comparison to the sketch in Section 4.3.1. The eight radial distribution grids evaluated in this section comprise around 8000-14000 nodes and 90-165 transformers and a part of a Siemens internal project. The number and size of the integrated DER are based on a 2020 scenario. Grid topology and DER reason intelligent infrastructure, like multiple OLTC, or grid extensions essential to guarantee the voltage band.

In the following a snap shot of the first time point in the year with maximal voltage magnitude of at least 1.12 is approached by the HILL and DYN. Each of the transformers was assumed a range of possible turn ratios $\mathcal{T} = \{0.9, 0.92, \ldots, 1.1\}$.

The resulting DYN tap positions solve the over voltage issue for the given generation and demand pattern in all eight grids, see Table 4.2. In contrast, the HILL fails in sixth distribution grids since it lifted the voltage level on the HV/MV too high and the last transformers do not have the flexibility to step further down. Moreover, the DYN approach frees more of the available voltage range, i.e. lower voltage spread see Table 4.2. At the same time, DYN reduces the losses in comparison to the HILL by up to 9.1% and in average by 3.6%, see Table 4.2. This is a result of the higher voltage regimes, e.g. Figure 4.8.

Table 4.2: HILL and DYN approach solutions summary of the eight distribution grids.

	Grid	1	2	3	4	5	6	7	8
BASE	# Nodes	8813	8382	13662	8142	11075	8487	8109	13085
	# OLTC	98	98	165	100	100	104	99	158
	$\max(\mathbf{u})$	1.1218	1.1465	1.1344	1.125	1.2161	1.1546	1.1683	1.1274
	$\min(\mathbf{u})$	0.9972	0.9966	0.976	0.9636	0.9553	0.9922	0.9942	0.9882
	spread	0.1246	0.1499	0.1585	0.1613	0.2608	0.1623	0.1741	0.1393
HILL	$\max(\mathbf{u})$	1.0964	1.0813	1.0948	1.0904	1.1163	1.1089	1.0954	1.0927
	$\min(\mathbf{u})$	0.9799	0.9237	0.938	0.9581	0.9279	0.9168	0.902	0.9557
	spread	0.1166	0.1575	0.1568	0.1323	0.1885	0.1921	0.1934	0.137
	p_{loss} [MW]	0.0509	0.0358	0.0489	0.0197	0.0451	0.4243	0.3995	0.0356
DYN	$\max(\mathbf{u})$	1.095	1.0945	1.0951	1.0948	1.0995	1.0909	1.0955	1.0944
	$\min(\mathbf{u})$	0.9799	0.9697	0.9424	0.96	0.9055	0.9522	0.9394	0.9631
	spread	0.1152	0.1249	0.1526	0.1348	0.1940	0.1387	0.1561	0.1313
	p_{loss} [MW]	0.0509	0.0331	0.0469	0.0194	0.0469	0.4068	0.3662	0.035

4.4 Conclusion

This chapter presents a novel dynamic programming VVC algorithm for active radial distribution grids. It is based on the computational complexity reduction arising from multiple optimizations over one discrete variable in contrast to an optimization over multiple discrete interdependent variables. The grid topology is used for a decomposition of the overall grid into subgrids of one discrete variable, which form together with the coordinating pseudo generators the subproblems. Pseudo generator functions of the voltage magnitude at the coupling nodes coordinate the subgrids subproblems such that each local solution

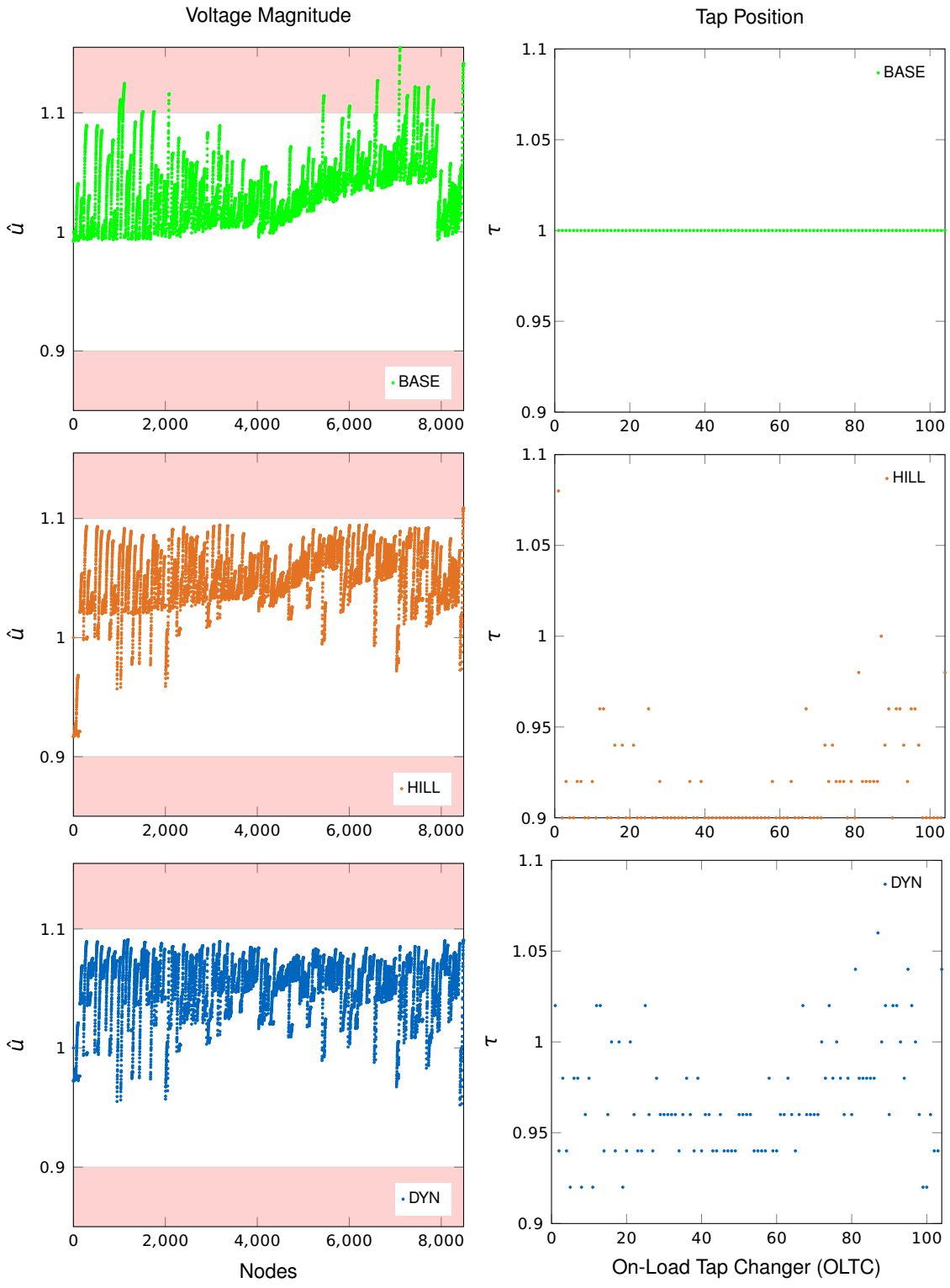


Figure 4.8: Illustration the voltage magnitudes and tap position of the sixth grid for the BASE, HILL and DYN approach. The BASE tap positions 1 lift the voltage over the upper voltage limit 1.1, first line of plots. The HILL approach consecutively set tap positions lead to only few voltage band violations, second line. The DYN approach tap positions solve the over-voltage issue completely, third line.

contributes its parts to the optimal overall solution. Thus, the proposed algorithm finds the optimal solution.

The simplified distribution grid example in Section 4.3.1 shows this optimality in a very simplified grid with a homogenous feed-in. In this case perfectly suited for the LDC method, the DYN approach showed how to find the optimal solution while the approximation within the LDC slightly weakens this method's results. Moreover, in this example the hill climbing heuristic shows its dependence on the initial discrete variables and the lack optimality for one initial setting even though it has full knowledge of the grid. These results stand out even more in the authentic distribution grid examples, in Section 4.3.2. Here the hill climbing approach results in slight over voltages for two out of the eight distribution grids considered, while the dynamic programming approach resulted in none at lower losses.

The pseudo generator functions give a comprehensive picture of the voltage magnitude dependent power losses and costs, while the discrete variable mapping function yields the optimal discrete variables. If the approach is not implemented online, an abstraction of these mapping functions can be the basis for local control tables given by a few characteristics, e.g. voltage magnitude, hour and weather. An appropriate set of such parameters should be the focus of future work.

Chapter 5

Synthesis and Outlook

The proposed theoretical methods require communication and measurements, which do not exist today in distribution grids. The Message Passing (MP) method for Economic Dispatch (ED), Chapter 3, is based on communication between the neighboring nodes in the grid topology. Dynamic programming for Volt-VAr Control (VVC), Chapter 4, demands measurements as well as communication between subgrids. Today distribution grids lack communication infrastructure [56] and measurement devices are rare [80]. However, the dispatch and control in distribution grids with a high number of Distributed Energy Resources (DER) will require these technical tools and they are gaining in research interest at the moment under the term Smart Grid. Thus, the requirements are just ahead of time and not standard yet.

Currently, a number of communication technologies are under discussion. The infrastructure for power line communication is in place and the connection to the actual device is in place. However, power lines are noisy channels that make it hard to modulate signals and the bandwidth is relatively low [56]. A high bandwidth alternative could be the digital subscriber lines (DSL) or optical fibre lines [132]. DSL has been used at Stadtwerke Emden, a municipal utility, as part of a pilot scheme in cooperation with the Deutsche Telekom [126]. The infrastructure for DSL is an integral component in inhabited areas but the connections to remote DER, e.g. wind turbines, are expensive. This is no issue for cellular network communication, for example as used in the IRENE project [70]. Nevertheless, cellular networks' reliability is limited in rural areas [56]. At the moment no communication technology exists that can be recommended for all regions and applications. Hence a specially adapted combination of technologies might be the method of choice for future applications [4]. An even more diverse candidate set exists for communication protocols; an extensive list of these as well as their qualities can be found in [32]. So far, none has prevailed. Europe made the first steps to define comprehensive standards by a guideline [55] issued in 2012. However, these standards are currently under review [32] and the standard committee expects this to be a long-term process [33].

Regardless of the yet to be assembled technological framework, communication is and will be an essential part of future smart distribution grids, not only with regard to dispatch but also with implications for default security [56, 132]. For the latter a detailed monitoring

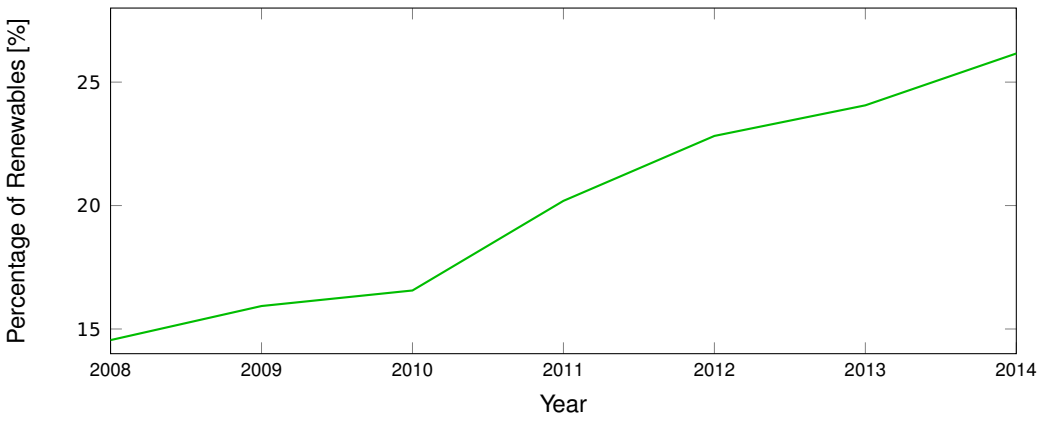
of the distribution grid status is vital [5] but not yet enabled.

At the moment distribution grids are sparsely monitored, in contrast to the transmission grids [80, 101]. The mass role out of smart meters can reverse this trend and turn the distribution grid into a comprehensively monitored part of the electrical grid. First regulatory steps have been taken towards a roll out of online metering in Germany [49]. Besides this the incentive structure still needs to be adjusted to foster this roll out [4]. Aside for measurements at connections, additional measurement units, e.g. phasor measurement units, at critical locations in the distribution grid can contribute important knowledge of the grid state [122, 101]. Aggregated and evaluated, e.g. by state estimation techniques [54, 63, 69, 87], a holistic picture of the distribution grid can be drawn, allowing for a specially suited control of devices and more efficient use of the existing infrastructure which leads to a reduction of required grid extensions. Moreover, new technologies might help to gather additional insights in the state of the grid in future. Autonomous sensors are currently tested for High Voltage (HV) overhead lines, to maximize the capacities, which are today operationally restricted by conservative worst case scenarios [81]. Similar approaches could temporarily enlarge the allowed flows of overhead lines in the distribution grid. Thus, the capacities may be further increased without grid extensions.

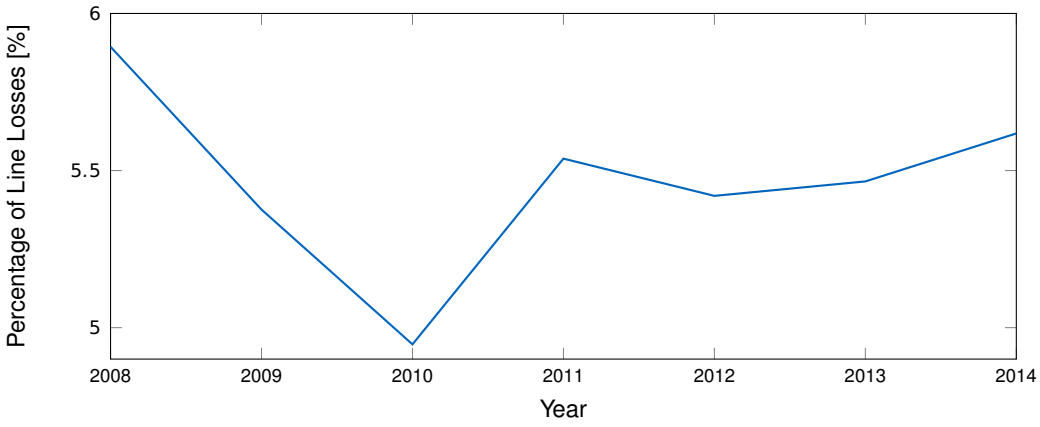
The potential to circumvent costly grid extensions has been the argument for the coordination of DER and active VVC, which necessitate the aforementioned changes. However, the required devices are expensive and operational costs occur [4, 3]. On-Load Tap Changer (OLTC) are approximately four-times as expensive as traditional transformers and might need to be more regularly maintained [3]. Moreover, meters and measurement equipment is costly and the data communication and evaluation add operational costs [4]. Additional values have the ability to countervail these operational costs but depend on additional gains of control beyond the voltage limits.

The reduction of line losses has the potential to be such an additional value. In centrally supplied grids the line losses were dependent on the distance the power was transported. The proximity of supply and demand in distribution grids with DER implies a loss reduction. However, losses due to non homogenous voltages that occur in a static infrastructure faced with intermittent sources counteract this reduction. Moreover, DER are also not aligned with load centers [115]. On the whole the line losses have not been declining in the last five years although the contributions from distributed renewables rose, see Figure 5.1. In Germany 85%[1] of the overall $5-6\%$ line losses occur in the distribution grid [71]. The accumulated line losses in distribution grids add up to around 30 TWh a year. This result in approximately 1.5 bn Euro a year in losses for distribution system operators based on a price of around 5 Cent/kWh. A small reduction can thus contribute to significant savings. The realistic distribution grids in Section 4.3.2 indicate that optimally controlled OLTCs can reduce the losses in comparison with heuristic approaches. Nevertheless, these results are based on theoretical models, and one should not model for numbers but for insights [65].

[1] This estimation is based on the average losses of the local utilities WW Netz, Netz Mittelrhein, Netz BW, Stadtwerke Rostock, München, Kaiserslautern und Duisburg. Moreover, these values are plausible with regard to the distribution grid line losses in Europe stated in [127].



(a) Percentage of electricity generation that comes from renewables in Germany [24]



(b) Percentage of the electricity generation lost due to line losses in Germany [71]

Figure 5.1: Comparison of the contribution of renewables and the line losses occurring in the electricity generation in Germany between 2008 and 2014.

Glossary

AC Alternating Current.

ADMM Alternating Direction Method of Multipliers. A decomposition method combining dual decomposition techniques with robust augmented Lagrangian approaches. An introduction can be found in [23].

CHP Combined Heat and Power Plant.

clique set of vertices where all vertices are pairwise adjacent.

clique graph a graph where cliques are nodes and edges are non empty separator sets between the variables of cliques.

clique tree singly-connected clique graph.

CMP Continuous Message Passing method detailed in Section 3.5.

DER Distributed Energy Resources refers to small-scale power generation, storage or deferrable load [88].

DMP Discrete Message Passing method detailed in Section 3.4.

ED Economic Dispatch. The economic dispatch problem is to schedule a number of electricity generation or consumption units such that the overall welfare, i.e. the summed utility of all consumption minus the summed cost of all production, is maximized. This is to be done subject to generation constraints and may include transportation constraints.

EV Electric Vehicle.

GM Graphical Model describes via a graph a family of multivariate probability distributions that share a common (conditional) independence structure.

HV High Voltage level, nominal system voltages greater than or equal 110 kV [28].

junction tree clique tree possessing the junction tree property.

junction tree property if for any clique C_1 and C_2 in a clique tree all cliques on the unique path joining them contain their separator set S_{12}.

LDC Line Drop Compensation.

LV Low Voltage level, nominal system voltages less than or equal 1 kV [28].

MILP Mixed Integer Linear Programming.

MP Message Passing.

MV Medium Voltage level, nominal system voltages greater than 1 kV and less than 110 kV [28].

OLTC On-Load Tap Changer.

OPF Optimal Power Flow.

PV Photovoltaic.

QP Quadratic Programming.

UC Unit Commitment. The unit commitment problem minimizes the cost of operation by optimal dispatch of available electricity generation units subject to load, generation and/or transmission constrains. Our formulation of ED includes the basic UC on/off constraint.

VVC Volt-VAr control.

Symbols

C_{v0} transport cost from v to auxiliary node0.

C_v cost or negative utility function.

P joint probability distribution of random variables x_V.

T turn ratio or capacitance function of the voltage magnitude $\hat{u}$.

U voltage magnitude function.

Z normalization factor.

$\mathcal{A}$ minimal active index set.

p_{vw} (active) power flow from node v to w.

p_v net (active) power injection of node $v \in \mathcal{V}$.

$\mathbf{A}_{vw}$ inequality constraint matrix of $m_{vw}(x)$.

$\mathcal{C}(v)$ set of all children of the vertex v in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$.

i_{vw} complex current flow from node v to w.

i_v complex current at node $v \in \mathcal{V}$.

P active power function.

J cost-to-go function.

Q reactive power function.

$\mathbf{H}_{vw}$ matrix for the quadratic part of $f(\mathbf{p}, x)$.

$\mathcal{K}(l)$ set of all sub-grid indices interconnected to $\mathcal{G}[\mathcal{V}_l]$.

$\mathbf{Y}$ admittance matrix given by (2.4).

$\hat{\mathcal{P}}_{vw}$ attainable subset of $\mathcal{P}_{vw}$.

s_n reference power at node $v \in \mathcal{V}$.

s_{vw} complex power flow from node v to w.

s_v complex power at node $v \in \mathcal{V}$.

q_{vw} reactive power flow from node v to w.

q_v reactive power at node $v \in \mathcal{V}$.

$\mathcal{L}$ index set of all sub-grids $\mathcal{G}[\mathcal{V}_l]$.

$\mathcal{T}_l$ set of feasible discrete turn ratios or capacitance τ_l.

$\mathcal{T}_{vw}$ set of feasible discrete turn ratios or capacitance τ_{vw}.

$\mathcal{T}$ set of feasible discrete turn ratios or capacitance τ.

$\mathcal{U}$ feasible set for $\mathbf{u}$.

u_v complex voltage at node $v \in \mathcal{V}$.

θ_{vw} voltage angle difference of the nodes v and w.

θ_v voltage angle at node $v \in \mathcal{V}$.

$[\underline{\hat{u}}_v, \overline{\hat{u}}_v]$ voltage band for $\hat{u}_v$.

$\hat{u}_{v,n}$ reference voltage at node $v \in \mathcal{V}$.

$\hat{u}_v$ voltage magnitude at node $v \in \mathcal{V}$.

$\hat{\alpha}_{vw}$ quadratic coefficient of $\hat{m}_{vw}$.

$\hat{\beta}_{vw}$ linear coefficient of $\hat{m}_{vw}$.

y_{vw} per unit admittance the line from v to w, $\{v, w\} \in \mathcal{E}$.

$\hat{\gamma}_{vw}$ constant coefficient of $\hat{m}_{vw}$.

$\boldsymbol{\alpha}_{vw}$ inequality constraint vector of $m_{vw}(x)$.

$\boldsymbol{\alpha}^x_{vw}$ inequality constraint vector dependent on x of $m_{vw}(x)$.

α_v quadratic coefficient of C_v.

α^i_{vw} quadratic coefficient of i-th piece of (3.13).

$\hat{m}_{vw}$ approximated message from node v to w.

β_v linear coefficient of C_v.

β^i_{vw} linear coefficient of i-th piece of (3.13).

$\cdot^H$ complex conjugate transposed.

$\odot$ component-wise multiplication.

$\mathbf{c}_{vw}$ vector for the linear part of $f(\mathbf{p}, x)$.

$\mathbf{c}_{vw}^{x}$ vector for the dependent on x linear part of $f(\mathbf{p}, x)$.

δ_v active power demand at node $v \in \mathcal{V}$.

γ_v constant coefficient of C_v.

γ_{vw}^{i} constant coefficient of ith piece of (3.13).

$x_{\mathbf{V}}$ set of all random variables x_v, $v \in V$.

x_v a random variables.

$\hat{\mathcal{T}}_l(\hat{u}_i, \hat{u}_j)$ set of restricted feasible discrete turn ratios or capacitance (4.10).

$\hat{z}_l$ impedance estimate of $\mathcal{G}[\mathcal{V}_l]$ in the Line Drop Compensation (LDC) approach.

$\hat{y}(l)$ pseudo generator node in $\mathcal{G}[\mathcal{V}_l]$.

$\hat{r}_l$ resistance of $\hat{z}_l$.

$\hat{x}_l$ reactance of $\hat{z}_l$.

$\mathbb{1}$ identity matrix.

$\mathbf{C}$ set (collection of maximal) cliques of $\mathcal{G}(\mathcal{V}, \mathcal{E})$, meaning maximal completely connected subsets of $\mathcal{V}$.

$\mathbf{W}$ variable for $\mathbf{u}\mathbf{u}^H$.

$\mathbf{e}$ vector of ones.

$\mathcal{E}^0$ set of transmission lines in the network and the shunt representing lines to 0.

$\mathcal{E}$ set of transmission lines in the network.

$\mathcal{G}(\mathcal{V}, \mathcal{E})$ graph (directed or undirected) with vertices $\mathcal{V}$ and edges $\mathcal{E}$.

$\mathcal{N}(v)$ set of all neighbors of the vertex v in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$.

$\mathcal{P}(v)$ set of all parents of the vertex v in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$.

$\mathcal{P}_{vw}$ feasible set for p_{vw}.

$\mathcal{P}_v$ feasible set for p_v.

$\mathcal{Q}_v$ feasible set for q_v.

$\mathcal{V}$ set of buses in the network.

$\mathcal{Y}^*$ index of the restricted branching nodes.

$\mathcal{Y}$ index of the restricted branching nodes and the chain separator nodes.

τ discrete turn ratio or capacitance variable vector in $\mathcal{T}$.

C a clique of $\mathcal{G}(\mathcal{V}, \mathcal{E})$, meaning maximal completely connected subsets of nodes $\mathcal{V}$.

u_v^n nominal voltage of node v.

$\overline{p}_{vw}$ maximal active power carrying capacity of line vw.

$\overline{p}_{vw}$ maximal active power carrying capacity of line vw.

$\overline{p}_v$ maximal (active) power output of node $v \in \mathcal{V}$.

$\mathbf{p}$ vector (active) power flows in m_{vw}, i.e. $(p_{kv})_{k \in \mathcal{N}(v) \setminus \{w\}}$.

ϕ_C potential function of the clique C.

$\mathcal{D}$ transformers and capacitor banks discrete variable subset $\mathcal{E}^0$.

$\mathcal{G}[\mathcal{V}_l]$ subgraph induced by $\mathcal{V}_l \subset \mathcal{V}$, meaning the graph consisting of vertices $\mathcal{V}_l$ and each edge in $\mathcal{E}$ connecting two vertices $\mathcal{V}$, in electrical grids each edge in $\mathcal{E}^0$ connecting vertices in $\mathcal{V} \cup 0$.

$\tilde{i}_{vw}$ variable for $|i_{vw}|^2$.

$\tilde{u}_v$ variable for $|u_v|^2$.

τ_l discrete turn ratio or capacitance variable in $\mathcal{T}_l$ for $\mathcal{G}[\mathcal{V}_l]$.

τ_{vw} discrete turn ratio or capacitance variable in $\mathcal{T}_{vw}$ for $\{vw\} \in \mathcal{D}$.

$\underline{p}_{vw}$ maximal active power carrying capacity of line vw.

$\underline{p}_v$ minimal (active) power output of node $v \in \mathcal{V}$.

0 reference/ground node.

$y(l)$ primary node in $\mathcal{G}[\mathcal{V}_l]$ and in $\mathcal{Y}$.

b_v binary variable associated with node $v \in \mathcal{V}$.

dt distance measure from the reference transformer.

$f(\mathbf{p}, x)$ objective function of $m_{vw}(x)$.

m_{vw} message from node v to w.

Bibliography

[1] Anwendungsregel:2011-08, Erzeugungsanlagen am Niederspannungsnetz, 8 2011. 27

[2] DIN EN 60038 VDE 0175-1:2012-04, 04 2012. 27

[3] A. Agricola, B. Höflich, P. Richard, J. Völker, C. Rehtanz, M. Greve, B. Gwisdorf, J. Kays, T. Noll, J. Schwippe, A. Seack, and J. Teuwsen. dena-Verteilnetzstudie. Ausbau- und Innovationsbedarf der Stromverteilnetze in Deutschland bis 2030. Deutsche Energie-Agentur GmbH (dena), 2012. URL http://www.dena.de/fileadmin/user_upload/Projekte/Energiesysteme/Dokumente/denaVNS_Abschlussbericht.pdf. 15, 82

[4] A. Agricola, P. Richard, H. Kobel, L. Einhellig, K. Behrens, L. Preysing, C. Rehtanz, B. Gwisdorf, A. El-Hadidy, V. Liebenau, and J. Teuwsen. Einführung von Smart Meter in Deutschland. Deutsche Energie-Agentur GmbH (dena), 2014. URL http://www.dena.de/fileadmin/user_upload/Publikationen/Energiesysteme/Dokumente/140709_dena-Smart-Meter-Studie_Endbericht_final.pdf. 81, 82

[5] E. Ahlers. Smart Grids und Smart Markets - Roadmap der Energiewirtschaft. In Christian Aichele and Oliver D. Doleski, editors, Smart Market, pages 97–123. Springer Fachmedien Wiesbaden, 2014. ISBN 978-3-658-02777-3. doi: 10.1007/978-3-658-02778-0_4. URL http://dx.doi.org/10.1007/978-3-658-02778-0_4. 81

[6] M. Ahmadipour. A new approach to control coordinated voltage in radial distribution feeders. International Research Journal of Applied and Basic Sciences, 4 (1): 243–251, 2013. 17, 29, 76

[7] R. Ahuja, T. Magnanti, and J. Orlin. Network Flows: Theory, Algorithms, and Applications. Prentice Hall, New Jersey, 1993. 23

[8] G. Andersson. Modelling and Analysis of Electric Power Systems. http://www.eeh.ee.ethz.ch/uploads/tx_ethstudies/modelling_hs08_script_02.pdf, Sept 2008. 27, 28

[9] X. Bai, H. Wei, K. Fujisawa, and Y. Wang. Semidefinite programming for optimal power flow problems. International Journal of Electrical Power & Energy Systems, 30:383 – 392, 2008. ISSN 0142-0615. 25

[10] R. Baldick and F.F. Wu. Efficient Integer Optimization Algorithms for Optimal Coordination of Capacitors and Regulators. Power Systems, IEEE Transactions on, 5(3): 805–812, Aug 1990. ISSN 0885-8950. doi: 10.1109/59.65909. 28

[11] R. Baldick and F.F. Wu. Approximation Formulas for the Distribution System: the Loss Function and Voltage Dependence. Power Delivery, IEEE Transactions on, 6 (1):252–259, Jan 1991. ISSN 0885-8977. doi: 10.1109/61.103745. 28

[12] M.E. Baran and F.F. Wu. Optimal capacitor placement on radial distribution systems. Power Delivery, IEEE Transactions on, 4(1):725–734, Jan 1989. ISSN 0885-8977. doi: 10.1109/61.19265. 23

[13] G. Barbose, S. Weaver, and Darghouth N. Tracking the Sun VII An Histrorical Summary of the Installed Price of Photovoltaics in the United States from 1998 to 2013. Berkeley Lab, 2014. 14

[14] R. Bellman. Dynamic Programming. Princeton University Press, Princeton, NJ, USA, 1 edition, 1957. 33

[15] A. Bergen and Y. Vittal. Power Systems Analysis. Prentice Hall, 2000. 20

[16] D. Bertsekas. Nonlinear Programming. Athena Scientific, 1999. 20

[17] D. Bertsekas. Dynamic programming and optimal control/1. Athena Scientific, 2005. ISBN 1-886529-26-4. 33

[18] A. Borghetti. A Mixed-Integer Linear Programming Approach for the Computation of the Minimum-Losses Radial Configuration of Electrical Distribution Networks. Power Systems, IEEE Transactions on, 27(3):1264–1273, Aug 2012. ISSN 0885-8950. doi: 10.1109/TPWRS.2012.2184306. 29

[19] A. Borghetti. Using mixed integer programming for the volt/var optimization in distribution feeders . Electric Power Systems Research, 98(0):39 – 50, 2013. ISSN 0378-7796. doi: http://dx.doi.org/10.1016/j.epsr.2013.01.003. URL `http://www.sciencedirect.com/science/article/pii/S0378779613000138`. 29

[20] S. Bose, S.H. Low, and K.M. Chandy. Equivalence of Branch Flow and Bus Injection Models. In Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on, pages 1893–1899, Oct 2012. doi: 10.1109/Allerton.2012.6483453. 26

[21] S. Bose, S.H. Low, T. Teeraratkul, and B. Hassibi. Equivalent Relaxations of Optimal Power Flow. Automatic Control, IEEE Transactions on, 60(3):729–742, March 2015. ISSN 0018-9286. doi: 10.1109/TAC.2014.2357112. 25, 26

[22] S. Boyd and L. Vandenberghe. Convex Optimization. Cambridge University Press, 2004. 20

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Foundation and Trends in Machine Learning, 3 (1):1–122, 2010. 85

[24] Bundesministerium für Wirtschaft und Energie. Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland. Technical report, Deutschland, 2015. 9, 13, 83

[25] Bundesministerium für Wirtschaft und Energie. Das Erneuerbare-Energien-Gesetz. BMWi, 2015. URL http://www.erneuerbare-energien.de/EE/Navigation/DE/Gesetze/Das_EEG/das_eeg.html. 13

[26] Bundesministerium für Wirtschaft und Technologie. Untersuchungen zur Notwendigkeit einer weitergehenden Systemsteuerung zur Einhaltung der Systembilanz. Technical report, Consentec GmbH and Ecofys Germany GmbH, 2013. URL http://www.dena.de/fileadmin/user_upload/Projekte/Energiesysteme/Dokumente/denaVNS_Abschlussbericht.pdf. 14

[27] Bundesministerium für Wirtschafts und Energy. The Energy Transition: Switch to the future. BMWi Newsletter, March, 2015. URL http://www.bmwi-energiewende.de/EWD/Redaktion/EN/Newsletter/2015/01/Meldung/topthema-the-energy-transition.html. 13

[28] Bundesnetzargentur. Definitionen und Erläuterungen zur Datenerhebung Vergleichsverfahren Strom / Anreizregulierung. Technical report, Bundesnetzargentur, 2005. URL http://www.bundesnetzagentur.de/SharedDocs/Downloads/EN/BNetzA/Areas/ElectricityGas/CollectionCompanySpecificData/Decisions/DefinitionOfDataAnnex1Id3319pdf.pdf?__blob=publicationFile. 15, 85, 86

[29] F. Carbone, G. Castellano, and G. Moreschini. Coordination and Control of Tap Changers Under Load at Different Voltage Level Transformers. Proceedings of 8th Mediterranean Electrotechnical Conference on Industrial Applications in Power Systems, Computer Science and Telecommunications (MELECON 96), 3: 1646–1648, 1996. doi: 10.1109/MELCON.1996.551269. 29

[30] R. Carmona and M. Ludkovski. Valuation of energy storage: an optimal switching approach. Quantitative Finance, 10(4):359–374, 2010. URL http://ideas.repec.org/a/taf/quantf/v10y2010i4p359-374.html. 33

[31] M. Carrion and J.M. Arroyo. A Computationally Efficient Mixed-Integer Linear Formulation for the Thermal Unit Commitment Problem. Power Systems, IEEE Transactions on, 21(3):1371 –1378, Aug 2006. ISSN 0885-8950. doi: 10.1109/TPWRS.2006.876672. 15, 22, 44, 45

[32] CCMC. SGCGM490G Smart Grid Set of Standards Version 3.1. Technical report, CEN-CENELEC-ETSI Smart Grid Coordination Group, 2014. URL https://www.dke.de/de/std/informationssicherheit/documents/sgcg_standards_report.pdf. 81

[33] CCMC. SG-CGM490I Smart Grid Interoperability . Technical report, CEN-CENELEC-ETSI Smart Grid Coordination Group, 2014. 81

[34] Y. Chistyakov, E. Kholodova, K. Netreba, A. Szabo, and M. Metzger. Combined central and local control of reactive power in electrical grids with distributed generation. In Energy Conference and Exhibition (ENERGYCON), 2012 IEEE International, pages 325–330, Sept 2012. doi: 10.1109/EnergyCon.2012.6347775. 14

[35] R. Cichowski. Systematische Netzplanung. VWEW-Verlag, Frankfurt am Main, 1994. 76

[36] T. Cormen, C. Leiserson, R. Rivest, and C. Stein. Introduction to Algorithms. The MIT Press, 2009. 21, 33

[37] K. Dallmer-Zerbe and B. Wille-Haussmann. Verteilnetzplanung mit dezentralen Blindleistungsreglern und rONT. In OTTI-Konferenz Zukünftige Stromnetze für Erneuerbare Energien 2014, pages 1–6, 2014. 29

[38] K. Dallmer-Zerbe, W. Biener, and B. Wille-Haussmann. Reactive Power Control in Low Voltage Distribution Grids : Comparison of Centralized and Decentralized Q (U)-controller Designs Based on Probabilistic Power Flow Analysis. Internationaler ETG-Kongress, 9:1–6, 2013. 15, 29

[39] V.N. Dieu and Weerakorn Ongsakul. Enhanced merit order and augmented Lagrange Hopfield network for hydrothermal scheduling . International Journal of Electrical Power & Energy Systems, 30(2):93 – 101, 2008. ISSN 0142-0615. doi: http://dx.doi.org/10.1016/j.ijepes.2007.06.022. URL http://www.sciencedirect.com/science/article/pii/S0142061507000828. 21

[40] T.S. Dillon, K.W. Edwin, H.-D. Kochs, and R.J. Taud. Integer Programming Approach to the Problem of Optimal Unit Commitment with Probabilistic Reserve Determination. Power Apparatus and Systems, IEEE Transactions on, PAS-97(6):2154–2166, Nov 1978. ISSN 0018-9510. doi: 10.1109/TPAS.1978.354719. 22

[41] H. Dura. Optimum Number, Location, and Size of Shunt Capacitors in Radial Distribution Feeders A Dynamic Programming Approach. Power Apparatus and Systems, IEEE Transactions on, PAS-87(9):1769–1774, Sept 1968. ISSN 0018-9510. doi: 10.1109/TPAS.1968.291982. 34

[42] A. Einfalt, F. Zeilinger, R. Schwalbe, B. Bletterie, and S. Kadam. Controlling active low voltage distribution grids with minimum efforts on costs and engineering. In Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE, pages 7456–7461, Nov 2013. doi: 10.1109/IECON.2013.6700374. 29, 77

[43] IAS/PSE Power Systems Engineering. 141-1993 - IEEE Recommended Practice for Electric Power Distribution for Intustrial Plants. Red Book - Red Book Working Group, 1999. URL `https://standards.ieee.org/findstds/141-1993.html`. 15

[44] P. Esslinger and R. Witzmann. Regulated Distribution Transformers in Low-Voltage Networks with a high Degree of Distributed Generation. In Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on, pages 1–7, Oct 2012. doi: 10.1109/ISGTEurope.2012.6465833. 15

[45] M. Farivar and S.H. Low. Branch Flow Model: Relaxations and Convexification;Part I. Power Systems, IEEE Transactions on, 28(3):2554–2564, Aug 2013. ISSN 0885-8950. doi: 10.1109/TPWRS.2013.2255317. 26

[46] M. Farivar and S.H. Low. Branch Flow Model: Relaxations and Convexification;Part II. Power Systems, IEEE Transactions on, 28(3):2565–2572, Aug 2013. ISSN 0885-8950. doi: 10.1109/TPWRS.2013.2255318. 26

[47] M. Fila, G.A. Taylor, M.R. Irving, J. Hiscock, P. Lang, and P. Aston. Systematic modelling and analysis of TAPP voltage control schemes. In Universities Power Engineering Conference, 2007. UPEC 2007. 42nd International, pages 349–356, Sept 2007. ISBN 1905593368. doi: 10.1109/UPEC.2007.4468973. 27, 29, 76

[48] B. Frey and D. MacKay. A Revolution: Belief Propagation in Graphs with Cycles. Advances in neural information processing systems, pages 479–485, 1998. 32

[49] Bundesamt für Sicherheit in der Informationstechnik, editor. Das Smart Meter Gateway Sicherheit für intelligente Netze. WM Druck + Verlag, 2014. 82

[50] D. Gamarnik, D. Shah, and Y. Wei. Belief Propagation for Min-Cost Network Flow: Convergence and Correctness. Operations Research, 60(2):410–428, 2012. 23

[51] C. Gao and M. Redfern. A Review of Voltage Control Techniques of Networks with Distributed Generations using On-Load Tap Changer Transformers. Universities Power Engineering Conference (UPEC), 2010 45th International, pages 3–8, 2010. 27, 29, 76

[52] J. Glover, M. Sarma, and T. Overbye. Power System Analysis and Design. CENGAGE Learning, 2008. 27

[53] A. Gomez Esposito and E.R. Ramos. Reliable load flow technique for radial distribution networks. Power Systems, IEEE Transactions on, 14(3):1063–1069, Aug 1999. ISSN 0885-8950. doi: 10.1109/59.780924. 26

[54] A. Gomez-Exposito, A. Abur, A. de la Villa Jaen, and C. Gomez-Quiles. A Multilevel State Estimation Paradigm for Smart Grids. Proceedings of the IEEE, 99(6): 952–976, June 2011. ISSN 0018-9219. doi: 10.1109/JPROC.2011.2107490. 82

[55] EN-CENELEC-ETSI Smart Grid Coordination Group. Framework for Smart Grid Standardization. Brussels, 397, 2012. 81

[56] V.C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G.P. Hancke. Smart Grid Technologies: Communication Technologies and Standards. Industrial Informatics, IEEE Transactions on, 7(4):529–539, Nov 2011. ISSN 1551-3203. doi: 10.1109/TII.2011.2166794. 81

[57] N. Gupta, A. Swarnkar, and K.R. Niazi. An Efficient Greedy Approach for Reactive Compensation on Large-Scale Distribution Systems. In Power and Energy Society General Meeting, 2011 IEEE, pages 1–7, July 2011. doi: 10.1109/PES.2011.6038931. 17, 29

[58] G. Hadley. Nonlinear and dynamic programming. Addison-Wesley Publishing Company, 1964. 33

[59] M. Haslbeck, M. Sojer, T. Smolka, and O. Brückl. Mehr Netzanschlusskapazität durch regelbare Ortsnetztransformatoren. Etz, 9:2–7, 2012. 27

[60] C. Heilek. Modellgestützte Optimierung des Neubaus und Einsatzes von Erzeugungsanlagen und Speichern für elektrische und thermische Energie im deutschen Energiesystem. Dissertation, Technische Universität München, München, 2015. 14

[61] C.A. Hernandez-Aramburo, T.C. Green, and N. Mugniot. Fuel consumption minimization of a microgrid. Industry Applications, IEEE Transactions on, 41(3):673 – 681, may-june 2005. ISSN 0093-9994. doi: 10.1109/TIA.2005.847277. 47

[62] W.J. Hobbs, G. Hermon, S. Warner, and G.B. Shelbe. An enhanced dynamic programming approach for unit commitment. Power Systems, IEEE Transactions on, 3(3):1201 –1205, aug 1988. ISSN 0885-8950. doi: 10.1109/59.14582. 34

[63] Y. Hu, A. Kuh, T. Yang, and A. Kavcic. A Belief Propagation Based Power Distribution System State Estimator. Computational Intelligence Magazine, IEEE, 6(3):36–46, Aug 2011. ISSN 1556-603X. doi: 10.1109/MCI.2011.941589. 32, 82

[64] Matthias Huber and Matthias Silbernagl. Modeling Start-Up Times in Unit Commitment by Limiting Temperature Increase and Heating. In 12th International Conference on the European Energy Market, Lisbon, May 2015. IEEE Power and Energy Society. 22

[65] H. Huntington, J. Weyant, and J. Sweeney. Modeling for insights, not numbers: the experiences of the energy modeling forum. Omega, 10(5):449–462, 1982. 82

[66] IEEE PES Subcommitee on Distribution System Analysis. Distribution Test Feeders, 2013. URL `http://ewh.ieee.org/soc/pes/dsacom/testfeeders/`. 47, 60

[67] A. Ihler, J. Fisher, Willsky A., and Chickering M. Loopy Belief Propagation: Convergence and Effects of Message Errors. Journal of Machine Learning Research, 6: 905–936, 2005. 32

[68] R.A. Jabr. Radial Distribution Load Flow Using Conic Programming. Power Systems, IEEE Transactions on, 21(3):1458–1459, Aug 2006. ISSN 0885-8950. doi: 10.1109/TPWRS.2006.879234. 26

[69] R.A. Jabr. Optimal Power Flow Using an Extended Conic Quadratic Formulation. Power Systems, IEEE Transactions on, 23(3):1000–1008, Aug 2008. ISSN 0885-8950. 26, 82

[70] K. Jopp. "Irene" tests the Smart Grid of the future: Wildpoldsried tests the future of energy ["Irene" testet das Smart Grid der Zukunft: Wildpoldsried erprobt die energiezukunft]. BWK- Energie-Fachmagazin, 64(10):41–44, 2012. URL http://www.scopus.com/inward/record.url?eid=2-s2.0-84891785237&partnerID=40&md5=f82d7d44b6dc0fb58a8658532060617b. cited By 0. 81

[71] Jörg Kaiser. Bilanz Elektrizitätsversorgung 2008-2014. Deutsches Statistisches Bundesamt, 2015. URL https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/Energie/Erzeugung/Tabellen/BilanzElektrizitaetsversorgung.html. 82, 83

[72] E. Kellerer and F. Steinke. Skalierbare Einsatzplanung für Smarte Verteilnetze. In VDE-Kongress 2014, pages 683–686. VDE, VDE Verlag GmbH, Oct 2014. ISBN ISBN 978-3-8007-3641-6. URL https://www.vde-verlag.de/proceedings-de/453641118.html. 51

[73] E. Kellerer and F. Steinke. Scalable Economic Dispatch for Smart Distribution Networks. Power Systems, IEEE Transactions on, 30(4):1739–1746, July 2015. ISSN 0885-8950. doi: 10.1109/TPWRS.2014.2358375. 16, 36, 37, 38, 45, 51, 57, 64

[74] B. Kim and R. Baldick. Coarse-Grained Distributed Optimal Power Flow. IEEE Transactions on Power Systems, 12(2):932–939, 1997. 27

[75] M. Klobasa, G. Angerer, A. Lüllmann, J. Schlich, T. Buber, A. Gruber, M. Hönecke, and S. von Roon. Lastmanagement als Beitrag zur Deckung des Spitzenlastbedarfs in Süddeutschland. Agora Energiewende, Berlin, Aug 2013. URL http://www.agora-energiewende.de/fileadmin/downloads/publikationen/Studien/Lastmanagementstudie/Agora_Studie_Lastmanagement_Sueddeutschland_Endbericht_web.pdf. 14

[76] K. Kok, B. Roossien, P. MacDougall, O. van Pruissen, G. Venekamp, R. Kamphuis, J. Laarakkers, and C. Warmer. Dynamic Pricing by Scalable Energy Management Systems - Field Experiences and Simulation Results using PowerMatcher. In IEEE PES General Meeting, 2012. 23

[77] D. Koller and N. Friedman. Probabilistic Graphical Models: Principles and Techniques. Adaptive Computation and Machine Learning. MIT Press, 2009. ISBN 9780262013192. 27, 29, 31, 32, 51

[78] M. Kopicka, J. Drapela, and D. Topolanek. Voltage Regulation Optimization in Low Voltage Network Based on Voltage Quality Index. In Electric Power Engineering (EPE), Proccedings of the 2014 15th International Scientific Conference on, pages 241–246, 2014. ISBN 9781479938070. doi: 10.1109/EPE.2014.6839541. 27

[79] M. Kraning, E. Chu, J. Lavaei, and S. Boyd. Dynamic Network Energy Management via Proximal Message Passing. Foundations and Trends in Optimization, 1(2): 70–122, 2013. 17, 27

[80] F. Kupzog, R. Schwalbe, W. Prüggler, B. Bletterie, S. Kadam, A. Abart, and M. Radauer. Maximising low voltage grid hosting capacity for PV and electric mobility by distributed voltage control. e & i Elektrotechnik und Informationstechnik, 131(6):188–192, 2014. ISSN 0932-383X. doi: 10.1007/s00502-014-0213-1. URL http://link.springer.com/10.1007/s00502-014-0213-1. 14, 15, 29, 76, 81, 82

[81] S. Kurth. ASTROSE – Autakes Sensorennetzwerk zur Überwachung von Hochspannungsleitungen. Fraunhofer ENAS, 2015. 82

[82] A. Lam, B. Zhang, and D. Tse. Distributed Algorithms for Optimal Power Flow Problem. IEEE Conference on Decision and Control (CDC), pages 430–437, December 2012. doi: 10.1109/CDC.2012.6427082. 25, 26, 27

[83] J. Lavaei and S.H. Low. Convexification of Optimal Power Flow Problem. In Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on, pages 223–232, Sept 2010. doi: 10.1109/ALLERTON.2010.5706911. 24, 25

[84] J. Lavaei and S.H. Low. Zero Duality Gap in Optimal Power Flow Problem. Power Systems, IEEE Transactions on, 27(1):92–107, Feb 2012. ISSN 0885-8950. 25

[85] F. Lee. Short-term thermal unit commitment-a new method. Power Systems, IEEE Transactions on, 3(2):421–428, May 1988. ISSN 0885-8950. doi: 10.1109/59.192892. 21

[86] B. Lesieutre, D. Molzahn, A. Borden, and C. DeMarco. Examining the Limits of the Application of Semidefinite Programming to Power Flow Problems. In Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on, pages 1492–1499, Sept 2011. doi: 10.1109/Allerton.2011.6120344. 26

[87] X. Li. Fully Distributed State Estimation of Smart Grids. In Communications (ICC), 2012 IEEE International Conference on, pages 6580–6585, June 2012. doi: 10.1109/ICC.2012.6364888. 32, 82

[88] J.A. Peças Lopes, N. Hatziargyriou, J. Mutale, P. Djapic, and N. Jenkins. Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities . Electric Power Systems Research, 77(9):1189 – 1203, 2007. ISSN 0378-7796. doi: http://dx.doi.org/10.1016/j.epsr.2006.08.016. URL http://www.sciencedirect.com/science/article/pii/S0378779606001908. Distributed Generation. 85

[89] S.H. Low. Convex Relaxation of Optimal Power Flow: A Tutorial. In Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium, pages 1–15, Aug 2013. doi: 10.1109/IREP.2013.6629391. 24, 25

[90] S.H. Low. Convex Relaxation of Optimal Power Flow Part II: Exactness. Control of Network Systems, IEEE Transactions on, 1(2):177–189, June 2014. ISSN 2325-5870. doi: 10.1109/TCNS.2014.2323634. 26

[91] F. Lu and Y. Hsu. Reactive power/voltage control in a distribution substation using dynamic programming. Generation, Transmission and Distribution, IEE Proceedings-, 142(6):639–645, Nov 1995. ISSN 1350-2360. doi: 10.1049/ip-gtd:19952210. 29, 34, 76

[92] T. Lühn, G. Schlömer, G. Schmidtmann, B. Lehde, J. Schmiesing, L. Hofmann, and J. Geldermann. Multi-Criteria Analysis of Grid Expansion Concepts on the Low Voltage Level. Zeitschrift für Energiewirtschaft, 38(3):183–200, 2014. ISSN 0343-5377. doi: 10.1007/s12398-014-0134-z. URL http://link.springer.com/10.1007/s12398-014-0134-z. 27

[93] C. Martin. China Installed a Record 12 Gigawatts of Solar in 2013. Bloomberg, January, 2014. 14

[94] M. Metzger, A. Szabo, and J. Bamberger. Control as a Key Technology for the Integration of Renewables. In IFAC World Congress, 2011. 23

[95] M. Minoux. A polynomial algorithm for minimum quadratic cost flow problems . European Journal of Operational Research, 18(3):377 – 387, 1984. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/0377-2217(84)90160-7. URL http://www.sciencedirect.com/science/article/pii/0377221784901607. 22

[96] M. Minoux. Solving integer minimum cost flows with separable convex cost objective polynomially. In Giorgio Gallo and Claudio Sandi, editors, Netflow at Pisa, volume 26 of Mathematical Programming Studies, pages 237–239. Springer Berlin Heidelberg, 1986. ISBN 978-3-642-00922-8. doi: 10.1007/BFb0121104. URL http://dx.doi.org/10.1007/BFb0121104. 22

[97] C.C. Moallemi and B. Van Roy. Convergence of Min-Sum Message Passing for Quadratic Optimization. Information Theory, IEEE Transactions on, 55(5):2413–2423, May 2009. ISSN 0018-9448. doi: 10.1109/TIT.2009.2016055. 32

[98] S. Mokhtari, J. Sing, and B. Wollenberg. A unit commitment expert system [power system control]. Power Systems, IEEE Transactions on, 3(1):272–277, Feb 1988. ISSN 0885-8950. doi: 10.1109/59.43211. 21

[99] J.M. Mooij and H.J. Kappen. Sufficient Conditions for Convergence of the Sum-Product Algorithm. Information Theory, IEEE Transactions on, 53(12):4422 –4437, dec. 2007. ISSN 0018-9448. doi: 10.1109/TIT.2007.909166. 32

[100] K. Murphy, Y. Weiss, and M. Jordan. Loopy Belief Propagation for Approximate Inference: An Empirical Study. In Proceedings of the Fifteenth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-99), pages 467–476, San Francisco, CA, 1999. Morgan Kaufmann. 32

[101] N. Neusel-Lange. Dezentrale Zustandsüberwachung für intelligente Niederspannungsnetze. PhD thesis, Bergischen Universitüt Wuppertal, 2013. 82

[102] N. Noorshams and M. Wainwright. Belief Propagation for Continuous State Spaces: Stochastic Message-Passing with Quantitative Guarantees. The Journal of Machine Learning Research (JMLR), 14(1):2799–2835, 2013. 32

[103] D. Oeding and B. Oswald. Elektrische Kraftwerke und Netze. Springer, 2011. 25, 27

[104] Z. Ouyang and S.M. Shahidehpour. An intelligent dynamic programming for unit commitment application. Power Systems, IEEE Transactions on, 6(3):1203 –1209, aug 1991. ISSN 0885-8950. doi: 10.1109/59.119267. 22, 34

[105] N.P. Padhy. Unit commitment-a bibliographical survey. Power Systems, IEEE Transactions on, 19(2):1196 – 1205, may 2004. ISSN 0885-8950. doi: 10.1109/TPWRS.2003.821611. 21

[106] J. Park, S. Kim, G Park, J. Yoon, and S. Lee. Modified Dynamic Programming Based Unit Commitment Technique. In Power and Energy Society General Meeting, 2010 IEEE, pages 1–7, July 2010. doi: 10.1109/PES.2010.5588184. 34

[107] J. Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann Publishers, 1988. ISBN 9781558604797. 32

[108] R. Pindyck. Microeconomics. Pearson/Prentice Hall, 2009. 20

[109] Y. Pochet. Mathematical Programming Models and Formulations for Deterministic Production Planning Problems. In Michael Jünger and Denis Naddef, editors, Computational Combinatorial Optimization, volume 2241 of Lecture Notes in Computer Science, pages 57–111. Springer Berlin Heidelberg, 2001. ISBN 978-3-540-42877-0. doi: 10.1007/3-540-45586-8_3. URL http://dx.doi.org/10.1007/3-540-45586-8_3. 33

[110] Witzmann R. and G. Kerber. Statistische Analyse von NS-Verteilungsnetzen und Modellierung von Referenznetzen. Energiewirtschaft, 107, Heft 6:22–26, 2008. 39

[111] S. Rahimi, S. Massucco, F. Silvestro, M.R. Hesamzadeh, and Y. Tohidi. Applying Full MILP Model to Volt-Var Optimization Problem for MV Distribution Networks. In Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES, pages 1–5, Oct 2014. doi: 10.1109/ISGTEurope.2014.7028854. 29

[112] Varun Rai, Ross Baldick, David Spence, Melinda Taylor, and Michael Webber. The Nexus of Markets and the Environment: Introduction. In Austin Electricity Conference, 2015. 13

[113] S. von Roon and M. Huck. Merit Order des Kraftwerksparks. https://www.ffe.de/download/wissen/20100607_Merit_Order.pdf, 2010. 20, 21

[114] P. Rusmevichientong and B. Van Roy. An Analysis of Belief Propagation on the Turbo Decoding Graph with Gaussian Densities. Information Theory, IEEE Transactions on, 47(2):745–765, Feb 2001. ISSN 0018-9448. doi: 10.1109/18.910586. 32

[115] K. Schaber. Integration of Variable Renewable Energies in the European power system: a model-based analysis of transmission grid extensions and energy sector coupling. PhD thesis, Technische Universität München, 2014. 82

[116] K. Schaber, F. Steinke, and T. Hamacher. Managing Temporary Oversupply from Renewables Efficiently: Electricity Storage Versus Energy Sector Coupling in Germany. Conference Paper at the International Energy Workshop, Paris, 2013. 14

[117] M. Schmidt. UGM: Matlab code for undirected graphical models. http://www.di.ens.fr/ mschmidt/Software/UGM.html, Aug 2012. URL http://www.di.ens.fr/~mschmidt/Software/UGM.html. 44

[118] A. Schrijver. Theory of Linear and Integer Programming. Wiley, 2000. 22

[119] A. Schwab. Elektroenergiesysteme. Springer Berlin Heidelberg, 2012. URL http://dx.doi.org/10.1007/978-3-642-21958-0_10. 15, 16, 28

[120] T. Senjyu, K. Shimabukuro, K. Uezato, and T. Funabashi. A Fast Technique for Unit Commitment Problem by Extended Priority List. Power Systems, IEEE Transactions on, 18(2):882 – 888, May 2003. ISSN 0885-8950. doi: 10.1109/TPWRS.2003.811000. 21

[121] F. Sensfuss, M. Ragwitz, and M. Genoese. The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. Energy Policy, 36(8):3086 – 3094, 2008. ISSN 0301-4215. doi: http://dx.doi.org/10.1016/j.enpol.2008.03.035. URL http://www.sciencedirect.com/science/article/pii/S0301421508001717. 21

[122] J. Sexauer, P. Javanbakht, and S. Mohagheghi. Phasor Measurement Units for the Distribution Grid: Necessity and Benefits. In Innovative Smart Grid Technologies

(ISGT), 2013 IEEE PES, pages 1–6, Feb 2013. doi: 10.1109/ISGT.2013.6497828. 82

[123] Moshe Sniedovich. Dynamic Programming, Foundations and Principles. Chapman & Hall/CRC Press, 2011. 33, 34

[124] S. Sojoudi and J. Lavaei. Physics of Power Networks Makes Hard Optimization Problems Easy to Solve. In Power and Energy Society General Meeting, 2012 IEEE, pages 1–8, July 2012. doi: 10.1109/PESGM.2012.6345272. 25, 26

[125] A. Subramanian, M.J. Garcia, D.S. Callaway, K. Poolla, and P. Varaiya. Real-Time Scheduling of Distributed Resources. Smart Grid, IEEE Transactions on, 4(4): 2122–2130, Dec 2013. ISSN 1949-3053. doi: 10.1109/TSG.2013.2262508. 23

[126] T-Systems. Smart Metering für Stadtwerke Emden. online, 2010. URL https://www.t-systems.de/umn/uti/754478_1/blobBinary/Stadtwerke-Emden-ps.pdf?ts_layoutId=754458. 81

[127] R. Turconi, C. Simonsen, I. Byriel, and T. Astrup. Life cycle assessment of the Danish electricity distribution network. The International Journal of Life Cycle Assessment, 19(1):100–108, 2014. ISSN 0948-3349. doi: 10.1007/s11367-013-0632-y. URL http://dx.doi.org/10.1007/s11367-013-0632-y. 82

[128] M. Ulbrich and S. Ulbrich. Nichtlineare Optimierung. Birkhäuser, 2012. 20

[129] L. Végh. Strongly Polynomial Algorithm for a Class of Minimum-cost Flow Problems with Separable Convex Objectives. In Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing, STOC '12, pages 27–40, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1245-5. doi: 10.1145/2213977.2213981. URL http://doi.acm.org/10.1145/2213977.2213981. 23

[130] M. Wainwright and M. Jordan. Graphical Models, Exponential Families, and Variational Inference. Foundations and Trends in Machine Learning, 1(1-2):1–305, 2008. 27, 32

[131] M.J. Wainwright, T.S. Jaakkola, and A.S. Willsky. Tree-Based Reparameterization Framework for Analysis of Sum-Product and Related Algorithms. Information Theory, IEEE Transactions on, 49(5):1120–1146, May 2003. ISSN 0018-9448. doi: 10.1109/TIT.2003.810642. 32

[132] W. Wang, Y. Xu, and M. Khanna. A survey on the communication architectures in smart grid . Computer Networks, 55(15):3604 – 3629, 2011. ISSN 1389-1286. doi: http://dx.doi.org/10.1016/j.comnet.2011.07.010. URL http://www.sciencedirect.com/science/article/pii/S138912861100260X. 81

[133] Y. Weiss. Correctness of Local Probability Propagation in Graphical Models with Loops. Neural Comput., 12(1):1–41, January 2000. ISSN 0899-7667.

doi: 10.1162/089976600300015880. URL http://dx.doi.org/10.1162/089976600300015880. 44

[134] Y. Weiss and W. Freeman. Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology. Neural Comput., 13(10):2173–2200, October 2001. ISSN 0899-7667. doi: 10.1162/089976601750541769. URL http://dx.doi.org/10.1162/089976601750541769. 32

[135] R. Wiser and M. Bolinger. 2013 WIND TECHNOLOGIES MARKET REPORT. Energy Efficent & Renewable Energy U.S. Department for Energy, 2014. 14

[136] P. Wolfrum, M. Kautz, and J. Schäfer. Smart Operation of CHP Units. In Power Plants and Power Systems Control, volume 8, pages 61–66, 2012. 23

[137] L. Wolsey. Integer Programming. Wiley Series in Discrete Mathematics and Optimization. Wiley, 1998. ISBN 9780471283669. URL https://books.google.de/books?id=x7RvQgAACAAJ. 22

[138] A. Wood and B. Wollenberg. Power Generation Operation and Control. John Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore, 1996. 20, 23

[139] L Zdeborová, A Decelle, and M Chertkov. Message Passing for Optimization and Control of a Power Grid: Model of Distribution System with Redundancy. Physical Review E, pages 1–10, 2009. 32