Technische Universität München
Zentrum Mathematik
Lehrstuhl für Mathematische Optimierung

Distributed and event-triggered optimization in multi-agent networks

Dipl.-Math. Univ. Martin Meinel

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Oliver Junge
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Michael Ulbrich
	2.	Univ.-Prof. Dr.-Ing. Sandra Hirche
	3.	Univ.-Prof. Dr. Sebastian Sager Otto-von-Guericke-Universität Magdeburg

Die Dissertation wurde am 09.07.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 08.12.2015 angenommen.

To Gabriella and my parents

Acknowledgments

I would like to thank Prof. Dr. Michael Ulbrich for supervising this thesis by giving valuable advice and active support whenever needed. His great expertise created a working climate that enabled a fruitful research.

Moreover, I want to express my gratitude to Prof. Dr.-Ing. Sandra Hirche, Prof. Dr. Sebastian Sager, and Prof. Dr. Michael Ulbrich for reviewing this thesis.

I am grateful for the financial support of the German Research Foundation (DFG) within the priority program SPP 1305 "Control Theory of Digitally Networked Dynamical Systems". In this context, I also would like to thank Prof. Dr.-Ing. Sandra Hirche and Frederik Deroo as well as Azwirman Gusrialdi for the successful and pleasant cooperation within this program. Further, I thank Prof. Javad Lavaei for the helpful personal communication.

Next, I would like to thank my colleagues for the good times we had in- and outside the university. Especially, I am grateful to had the pleasure of sharing an office with Alana, Christian, and Sebastian who never were reluctant to spread good mood and contagious enthusiasm. In particular, I thank Sebastian who not only always had a sympathetic ear but also a helpful answer.

Last and most importantly, I deeply thank my family for their belief and support, and above all I thank my love Gabriella for always being there.

Abstract

This thesis is concerned with the development of distributed optimization methods with adaptive step-size control and event-triggered communication, where the focus is on convex optimization problems with either nonseparable objective function but separable constraints or separable objective function but couplings in the constraints.
Regarding a practice related application of the developed algorithms, it is shown how the convex direct current optimal power flow (DC-OPF) problem can be solved distributedly with event-triggered and local communication in a multi-agent network. Moreover, the combined application with a decomposition technique for linear matrix inequalities is described which enables to distributedly solve a semidefinite dual of the nonconvex alternating current optimal power flow (AC-OPF) problem with (close to) local and event-triggered communication.
Numerical results for these applications confirm the good properties of the developed algorithms and show that event-triggered communication yields a considerable reduction of the information exchange in the optimization process.

Zusammenfassung

Diese Dissertation befasst sich mit der Entwicklung von verteilten Optimierungsverfahren mit adaptiver Schrittweitensteuerung und ereignisbasierter Kommunikation für konvexe Optimierungsprobleme, in denen entweder eine nicht separable Zielfunktion durch separable Nebenbedingungen beschränkt ist oder eine separable Zielfunktion mit gekoppelten Nebenbedingungen betrachtet wird.
Hinsichtlich einer praxisbezogenen Anwendung der entwickelten Algorithmen wird gezeigt, wie das konvexe DC-OPF Problem mit ereignisbasierter und lokaler Kommunikation verteilt in einem Multi-Agentensystem gelöst werden kann. Darüberhinaus wird die kombinierte Anwendung mit einer Dekompositionstechnik für lineare Matrixungleichen beschrieben, die es ermöglicht ein semidefinites duales Problem des nichtkonvexen AC-OPF Problems mit (fast) lokaler und ereignisbasierter Kommunikation zu lösen.
Die numerischen Ergebnisse zu diesen Anwendungen belegen die guten Eigenschaften der entwickelten Algorithmen und zeigen, dass ereignisbasierte Kommunikation im Optimierungsprozess zu einer deutlichen Reduzierung des Informationsaustauschs führt.

Contents

Abbreviations

AC-OPF	alternating current optimal power flow
ANA	adaptive Nesterov-Algorithm
CC	consensus communication
CV	constraint violation
DANA	distributed adaptive Nesterov-Algorithm
DANA-EC	distributed adaptive Nesterov-Algorithm with event-triggered communication
DAPCA-EC	distributed adaptive proximal center algorithm with event-triggered communication
DC	dual communication
DC-OPF	direct current optimal power flow
DNA	distributed Nesterov-Algorithm
DNA-EC	distributed Nesterov-Algorithm with event-triggered communication
DPCA-EC	distributed proximal center algorithm with event-triggered communication
IDG	information dependency graph
L_k-Up	L_k-updates
LMI	linear matrix inequality
MCTpA	maximal computation time per agent
NA	Nesterov-Algorithm
NoCI	number of consensus iterations
NoI	number of iterations
PCA	proximal center algorithm
PG	primal gap
SDP	semidefinite programming
SSU	step-size updates
TC	total communication

Notations

$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$\mathbb{R}_+$	set of nonnegative real numbers
$[x, y]$	closed subset of $\mathbb{R}$
(x, y)	open subset of $\mathbb{R}$
$\mathbb{R}^m$	real vector space of dimension m
$\mathbb{C}^m$	complex vector space of dimension m
$\mathbb{R}^m_+$	subset of the vectors in $\mathbb{R}^m$ with nonnegative components
$\mathbb{S}^n$	set of symmetric $n \times n$ matrices
$\mathbb{S}^n_+$	set positive semidefinite $n \times n$ matrices
$X \times Y$	Cartesian product of the sets X and Y
$\|X\|$	cardinality of the set X

X^T	transpose of $X \in \mathbb{C}^{m \times n}$
X^*	conjugate transpose of $X \in \mathbb{C}^{m \times n}$
$X \succeq Y$	$X - Y \in \mathbb{R}^{n \times n}$ is positive semidefinite
$\mathrm{Tr}(X)$	trace of $X \in \mathbb{C}^{n \times n}$
$X \bullet Y$	Frobenius product of $X, Y \in \mathbb{C}^{n \times n}$
$\\|X\\|_{\mathrm{F}}$	Frobenius norm of $X \in \mathbb{C}^{n \times n}$
$\mathrm{Re}\{X\}$	real part of $X \in \mathbb{C}^{m \times n}$
$\mathrm{Im}\{X\}$	imaginary part of $X \in \mathbb{C}^{m \times n}$

$x_l \in \mathbb{R}^{m_l}$	subblock l of $x \in \mathbb{R}^m$ with $m_l \leq m$
(x, y)	either concatenation of $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ to $(x^T, y^T)^T \in \mathbb{R}^{m+n}$ or row vector for $x, y \in \mathbb{R}$ (clear from the context)
$\|x\|_1$	1-norm of $x \in \mathbb{R}^m$
$\|x\|$	Euclidean norm (2-norm) of $x \in \mathbb{R}^m$
$\|x\|_{\max}$	maximum norm of $x \in \mathbb{R}^m$
$\langle x, y \rangle$	standard Euclidean scalar product of $x, y \in \mathbb{R}^m$
$[x]^+$	component-wise projection of $x \in \mathbb{R}^m$ onto $\mathbb{R}^m_+$
$\lvert x \rvert$	1-norm of $x \in \mathbb{R}$
$\lceil x \rceil$	smallest upper integer of $x \in \mathbb{R}$

$\cos(x)$	cosine at $x \in \mathbb{R}$
$\sin(x)$	sinus at $x \in \mathbb{R}$
$\exp(x)$, e^x	exponential function at $x \in \mathbb{C}$
$f(x, y)$	function value of f at $(x^T, y^T)^T) \in \mathbb{R}^{m+n}$
$\nabla f(x)$	gradient of f at $x \in \mathbb{R}^m$
$\frac{\partial f}{\partial x_i}$	partial derivative of f w.r.t $x_i \in \mathbb{R}$
$\nabla_{ij} f(X)$	partial derivative of f w.r.t $X_{ij} \in \mathbb{R}$
$\partial f(X)$	set of subgradients of f w.r.t $X \in \mathbb{R}^{m \times n}$
$f'(x)$	derivative of f w.r.t $x \in \mathbb{R}$
$\mathcal{O}(f(x))$	asymptotically bounded from above by f

1 Introduction

Distributed optimization is a relatively young topic in the optimization literature and receives growing attention as there is a need to find solutions for optimization problems related to large-scale networks which become increasingly important in an interconnected and globalized world.
A good overview of network optimization problems and relevant literature can be found in [Lem10, sec. 1.6] and to give a few examples, estimation problems are for instance network problems, where the nodes of the network are sensors that locally measure a disturbed signal which is then globally estimated by the collaboration of neighbored sensors (e.g. [SFJ06, SBG07]). An example for an application is target tracking, where the location of multiple targets is estimated by a comparatively small number of navigation satellites (for details see [SBG07]).
Another class of problems is referred to as network utility maximization (e.g. [LL99, PC06, NO09]), where in its basic form the nodes of the network represent sources that want to transmit data via a predetermined set of lines in the network at a rate that maximizes their utility. As the transmission capacities are limited and the lines are shared among the sources, the allocation of the permitted transmission rates is done in a way that the overall utility is maximized. A practical application is the TCP congestion control of the data transmission via the internet (for related literature see [PC06, Lem10]).
Finally, the network optimization problem, that is focused on in the numerical part of this thesis, is the problem of finding the minimal cost of real power generation in an electrical power system which is a network that connects power generating units and loads via branches (e.g. transmission lines). For the determination of the optimal real power generation, constraints such as a balanced power flow within the network, limits on the power generation, and limits on the power flow at the branches have to be considered and the resulting optimization problem is referred to as alternating current optimal power flow (AC-OPF) problem (e.g. [KB00, LL12, LZT12, DZG13]) or in its simplified and linearized version as direct current optimal power flow (DC-OPF) problem (e.g. [BB03, JDR08, WL10]).

The size and the spatial distribution of large-scale networks make the usage of centralized optimization methods obsolete, especially in the case of privacy concerns (e.g. [JDR08, DUAH12, DMUH14a, DMUH15]), where the subsystems in the network do not want to share sensitive data and information with a central processor. To give an example, in a power system network this sensitive data might be the cost of power production as nowadays the power generating units belong to competitive power suppliers (see section 4.1). Accordingly, in this respect the amount of produced power at a power generating unit might be a sensitive information. Another aspect is the availability of information in the optimization process that might not be given centrally as in wireless sensor networks according to [NO09, sec. 1.1]. Finally, regardless of whether the considered optimization problem is related to a large-scale network or not, a parallelization of the computations in an optimization process is obviously favorable with respect to the complexity (for a comprehensive work on parallel computation see [BT89].
For this reasons, the goal in distributed optimization is to design optimization methods that can be implemented in parallel by a number of agents (processors) placed at the nodes of the network for a distributed computation of an optimal solution to a network related problem, where each agent controls only a subblock of the optimization variable. Generally, the agents need to exchange the iterates of their subblocks in the optimization process in order to approach an optimal solution as they have no access to the optimization variable as a whole. Regarding this information exchange, it is desirable that the distributed algorithm is designed in a way that an agent does not need to communicate with every other agent in the optimization process, and in the favorable case that the communication topology of the multi-agent network coincides with the topology of the considered large-scale network, the information exchange is referred to as being local.
Moreover, a large communication traffic is undesirable especially for capacity limited wireless communication networks [WL10]. To this end, event-triggered communication finds application in distributed optimization (e.g. [WL09a, WL09b, ZC10, WL10]), where the agents use outdated information of other agents which is allowed to differ to a certain extent from the up to date information. This extent is determined by a given threshold that adjusts to the stage of the optimization process and thereby guarantees its convergence.
Regarding the design of a distributed algorithm, the structure of the considered network problem is crucial. Except for the AC-OPF problem which is nonconvex, the above mentioned problems share in their basic form a separable structure, where a convex (or concave) objective function is separable with respect to a partition of the optimization variable into disjoint subblocks and constrained by nonseparable linear (in)equality

constraints that, however, can be decoupled by forming the corresponding Lagrangian which is a classical approach in distributed optimization, called dual decomposition (e.g. [LL99, PC06, SBG07, NO09]). The Lagrangian is then separable with respect to the subblocks of the primal optimization variable in the same way as the objective function of the primal problem and this separability confers to the corresponding dual objective function which at worst is constrained by nonnegativity constraints. In other words, the dual function can be evaluated at given feasible dual multipliers by minimizing (or maximizing) the Lagrangian with respect to the primal subblocks in parallel, and moreover, the dual optimal function value coincides with the optimal primal function value under mild assumptions due the convexity (or concavity) of the primal problem and the linearity of the constraints (details are given in section 3.1).
Finally, the concave (or convex) dual function is continuously differentiable if the Lagrangian is minimized uniquely for any given feasible dual multipliers which is for instance the case if the primal objective function is strictly convex. Then, a first order algorithm (a (projected) gradient scheme or an accelerated first order method) can be applied to maximize (or minimize) the dual objective function in a distributed manner (e.g. [LL99, PC06, DMUH14a, DMUH15]). However, if the primal objective function is only convex, a subgradient scheme can be applied to the dual problem (e.g. [PC06, SBG07, NO09]) or the primal objective function can be regularized with strongly convex functions, yielding a smooth augmented dual objective function (e.g. [Nes05, NS08]).
In this context, the proximal center algorithm (PCA) by Necoara and Suykens [NS08] (see section 3.1) is an efficient dual decomposition method that is designed for the application to convex problems with a separable convex objective function and dually decomposable linear constraints as described above. The authors of [NS08] apply a smoothing technique from [Nes05], where the primal convex objective function is regularized with (strongly convex) prox-functions that maintain its separability, yielding a continuously differentiable dual augmented function that has the same separability features as the dual function. Moreover, the augmented dual function has a Lipschitz continuous gradient which allows the application of an optimal first order scheme by Nesterov [Nes05] (see section 2.1) and explains the efficiency of the PCA whose speed of convergence to the optimal primal objective function value is in the order of $\mathcal{O}(1/k)$, where k is the number of iterations, whereas the convergence speed of a subgradient scheme is in the order of $\mathcal{O}(1/\sqrt{k})$ according to [NS08].
Nesterov's optimal first order method from [Nes05] is designed for convex optimization problems, where a continuously differentiable convex objective function with Lipschitz continuous gradient is constrained by a closed and convex subset of a real vector space.

It is optimal in the way that its speed of convergence to the optimal objective function value is of the order $\mathcal{O}(L/k^2)$, where L is the Lipschitz constant, whereas the convergence speed of an applied gradient projection algorithm would be of the order $\mathcal{O}(1/k)$ according to [Nes05].
In this work, both algorithms are enhanced with an adaptive step-size control and event-triggered communication, maintaining their efficiency (see chapter 2 & 3).

1.1 Outline of this work

In chapter 2, we present an adaptive accelerated distributed gradient scheme with event-triggered communication which is based on Nesterov's optimal first order method from [Nes05].
To this end, section 2.1 contains an introduction to Nesterov's method which is designed for convex optimization problems, where a continuously differentiable convex objective function with Lipschitz continuous gradient is constrained by a closed and convex subset of a real vector space. Nesterov's algorithm basically consists in each iteration of two simple strongly convex subproblems that contain only first order information which yields that the algorithm in all is suitable for a parallel implementation in a multi-agent network under mild assumptions such as the separability of the constraint set (see section 2.2). We show how this distributed version can be enhanced with event-triggered communication by introducing, similarly to [WL10, ZC10], outdated versions of the subblocks of the optimization variable that are used by the agents in the optimization process to compute their subblock of the gradient. The resulting distributed Nesterov-Algorithm with event-triggered communication (DNA-EC) inherits the convergence speed of the order $\mathcal{O}\left(L/k^2\right)$ as well as the simple structure of the subproblems that are slightly modified compared to their origins in order to handle the error due to event-triggered communication in the proof of convergence (see section 2.2).
In a next step, we enhance the DNA-EC by an adaptive step-size control to further accelerate the convergence of the algorithm. To this end, we firstly equip Nesterov's algorithm with an adaptive step-size control which is based on the work of Nesterov in [Nes13] and the observation that the first subproblem in Nesterov's algorithm is a projected gradient step with step size $1/L$ which suggests to use instead step-sizes of the from $1/L_k$ for $L_k \leq L$ in each iteration (see section 2.3).
Secondly, we modify this step-size control in a way that it can be implemented in parallel by using a consensus technique that is widely used in distributed optimization (e.g. [SFJ06, CCW10, DUAH12]) and is described in section 2.4.

Finally, these features are combined in the distributed adaptive Nesterov-Algorithm with event-triggered communication (DANA-EC) that is presented together with the proofs of convergence for different versions of it in section 2.5.

In chapter 3, we enhance the proximal center algorithm (PCA) by Necoara and Suykens [NS08] with event-triggered communication and an adaptive step-size control by using the results of chapter 2.
To this end, section 3.1 contains an introduction to the PCA which is designed for the application to convex problems with a separable convex objective function and nonseparable linear (in)equality constraints as described above.
In section 3.1.1, we improve the convergence result for the PCA from [NS08] (see section 3.1), where the number of iterations that is required to achieve a desired quality of the approximate solution depends, i.a., on the Lipschitz constant of the gradient of the augmented dual function. We show for a certain class of prox-functions that the Lipschitz constant can be analytically minimized with respect to the convexity parameters of these prox-functions in order to reduce the number of iterations and moreover, that it is possible to determine the analytical solutions of the optimal convexity parameters distributedly by the application of the consensus technique described in section 2.4. We further improve the convergence result of the PCA by proposing a scaling technique that provides an additional degree of freedom in the bounds on the quality of the approximate solution.
In section 3.2, we enhance the PCA by applying the developed DANA-EC instead of Nesterov's scheme to maximize the dual augmented function which results in the distributed adaptive proximal center algorithm with event-triggered communication (DAPCA-EC). Finally, we give two convergence results in this section which differ in the choice of prox-functions and the boundedness of the dual feasible set, however, in both cases we maintain the complexity of the PCA with $\mathcal{O}(1/k)$ iterations.

To prepare the application of the DAPCA-EC to distributedly (and with event-triggered communication) solve network problems that arise in a power system network, we establish in detail the models of the nonconvex alternating current optimal power flow (AC-OPF) problem as well as the convex direct current optimal power flow (DC-OPF) problem in chapter 4.
As mentioned above, we refer to the AC-OPF problem as the problem of finding the minimal cost of real power generation subject to constraints such as the power balance equations, real power generation limits, and limits on the power flow at the branches (e.g.

transmission lines) (see section 4.4). As the AC-OPF problem is nonconvex, its simplified linearization, the DC-OPF problem, is considered in practice if only the amount and cost of real power production is of interest (see section 4.5).
For a better understanding of the model that underlies these problems, an introduction to the structure of a power system and some important components is given in section 4.1, where we as well justify why it could be favorable to solve these optimization problems in a distributed manner. Moreover, technical terms such as the phasor representations of current and voltage as well as the definitions of real, reactive, and apparent power are introduced in section 4.2 to prepare the derivation of the power balance equations (see section 4.3) that relate the difference of the produced and consumed power to the power flowing in the network.

In chapter 5, we show how the DAPCA-EC can be applied to solve the DC-OPF problem and the AC-OPF problem distributedly and with event-triggered communication.
In section 5.1, we dually decompose the DC-OPF problem whose separable structure allows to directly apply the DAPCA-EC after the regularization of the real power generation cost function. We determine the Lipschitz constant as well as the partial derivatives of the augmented dual function and explicitly state the DAPCA-EC for this problem as it can be implemented. Moreover, we discuss in detail that the communication of the agents in the optimization process is fully local in the sense that it is sufficient for the agents to use the branches of the power system network for the information exchange with direct neighbors that are placed at the buses (nodes) of the power system. Finally, we show that the subproblems in each iteration of the DAPCA-EC applied to solve the DC-OPF problem have analytical solutions, superseding the need of solvers in the optimization process.
To prepare the application of the DAPCA-EC to solve the AC-OPF problem in parallel and with event-triggered communication, we show in section 5.2 how a semidefinite optimization problem, where a separable and convex objective function is constrained by a linear matrix inequality (LMI), can be solved distributedly by the DAPCA-EC. As the LMI introduces a dual matrix multiplier in dual decomposition that would require global information exchange among the agents, we apply the range-space conversion method from Kim et al. [KKMY11] which relates to semidefinite matrix completion and enables us to restate the LMI in a way that the dually decomposed problem can be solved distributedly by the DAPCA-EC with local communication if the sparsity structure of the LMI is chordal and coincides with the topology of the multi-agent network. However, for the case that this assumption is not satisfied, we show that the information exchange

can be carried out almost locally by finding the minimal chordal extension of the graph that represents the sparsity structure of the LMI. Finally, we determine the Lipschitz constant as well as the partial derivatives of the augmented dual function and explicitly state the DAPCA-EC applied to solve this class of semidefinite optimization problems in parallel and with event-triggered communication. We discuss in detail the communication topology of the agents and give analytical solutions for the subproblems that have to be solved in each iteration of the DAPCA-EC and partially are semidefinite problems themselves.

In section 5.3, we consider the semidefinite dual of the AC-OPF problem as derived by Lavaei and Low in [LL10, LL12]. This dual is an LMI-constrained optimization problem with separable and linear objective function whose optimum coincides with the optimum of the nonconvex AC-OPF problem under assumptions that are usually satisfied in practice as shown by the authors. Specifying the results from section 5.2, we apply the range-space conversion method to restate the LMI and solve the dual AC-OPF problem by the application of the DAPCA-EC which is explicitly stated and can be implemented by the agents with local information exchange if the considered power network is chordal (e.g. distribution network) or with close to local communication if the considered power system network is not chordal (e.g. transmission network) as will be discussed. Last but not least, we proof nontrivial analytical solutions to the subproblems in the DAPCA-EC applied to solve the dual of the AC-OPF problem and show how an approximate solution of the AC-OPF problem can be derived distributedly from the computed approximate solutions of its dual.

Finally, the numerical results for the application of different versions of the DAPCA-EC to the DC-OPF problem and the dual of the AC-OPF problem are presented in chapter 6. For the numerical investigation we used the data of benchmark IEEE test cases that represent portions of the American Electric Power System in the Midwestern US with 14, 30, and 57 buses (nodes).

In section 6.1, we discuss the choice of different parameters for the DAPCA-EC such as the threshold that determines the event-triggered communication and the start values as well as the update parameters for the adaptive step-size control. Moreover, we show how the scaling technique derived in section 3.1.1 is applied.

In sections 6.2 and 6.4, the numerical results for the IEEE 57 bus test case are exemplarily discussed and compactly summarized for the DC-OPF problem and the dual of the AC-OPF problem by the means of several tables that contain exhaustive data obtained by the following way of investigation:

In the first step, the PCA is compared with the DPCA-EC (DAPCA-EC without adaptive step-size control) to find out to what extent the information exchange can be reduced by event-triggered communication for a fixed number of iterations (given by the convergence result for the PCA in section 3.1) to achieve a predetermined accuracy of the approximate solution. We anticipate that for this setting event-triggered communication leads to considerable savings with respect to the information exchange.
Moreover, the comparison is repeated with a stopping-criterion for the primal gap and the constraint violation at the approximate solution in each iteration of the DPCA-EC, to firstly investigate the tightness of the convergence result for the PCA in section 3.1, and to secondly find out if there is a trade-off between the communication savings due to the usage of event-triggered communication and the necessary number of iterations to obtain a certain quality of the approximate solution. Remarkably, the results for this test setting show that the number of iterations and thereby the information exchange can be reduced by the application of the DPCA-EC (compared to the PCA) for a sufficiently tight threshold despite the inaccuracy that is introduced by the usage of event-triggered communication into the optimization process.
In the second step, the same stopping criterion is used to investigate how the adaptive step-size strategy in the DAPCA-EC helps to reduce the number of iterations compared to the DPCA-EC and it appears that the iterations can be reduced by up to four fifths. Furthermore, the impact of event-triggered communication in combination with the adaptive step-size strategy is studied and the results show that the overall communication can be crucially reduced if the threshold is tight enough. Finally, as the consensus technique for the distributed implementation of the adaptive step-size strategy is in all very time and information consuming if it is executed in each iteration of the DAPCA-EC, we propose several heuristics that execute the step-size update (and thereby the consensus technique) only sporadically to reduce the information exchange as well as the computation time, keeping the good results of the DAPCA-EC with respect to the number of iterations and the communication savings.
Similar results for the 14 and 30 bus test cases can be found in the appendix 7.

Parts of this thesis are already published or in preparation for publication:

In [MUA14] (Meinel, Ulbrich, and Albrecht), we presented the distributed Nesterov-Algorithm with event-triggered communication (DNA-EC), as well as the distributed proximal center algorithm with event-triggered communication (DPCA-EC). We gave convergence results for both and improved the accuracy estimates of the proximal center algorithm by the application of a scaling technique as well as optimal convexity parameters (or alternatively optimal scaling parameters). Moreover, we numerically investigated the impact of event-triggered communication by the application of the DPCA-EC to DC-OPF problems.

The preprint [MU14] (Meinel and Ulbrich) is in preparation for publication and contains the enhancement of the DNA-EC and the DPCA-EC with the adaptive step-size control that can be implemented distributedly by a consensus technique, yielding the DANA-EC as well as the DAPCA-EC. Moreover, it contains how the LMI-constrained convex problem with separable structure can be solved distributedly by the D(A)PCA-EC to prepare the application to the dual AC-OPF problem. Finally, the numerical results that show the impact of the adaptive step-size strategy and event-triggered communication are part of this preprint.

Further publications that are closely related to parts of this thesis:

In [DMUH14a, sec. 4] (Deroo, Meinel, Ulbrich, and Hirche) as well as [DMUH15, 3.3] (Deroo, Meinel, Ulbrich, and Hirche), we applied the range-space conversion method in combination with dual decomposition to solve an LMI-constrained strongly convex stability related problem with a distributed version of the proximal center algorithm in parallel (for details see section 5.2).

2 Distributed first order method with event-tiggering

In this chapter, an optimal first order scheme by Nesterov (NA 2.1.3) is presented in section 2.1 which is suitable for parallel implementation in a multi-agent network as shown in section 2.2, where the algorithm is additionally enhanced by event-triggered communication, yielding the distributed Nesterov-Algorithm with event-triggered communication (DNA-EC 2.2.5).
To accelerate the convergence speed of the DNA-EC 2.2.5, the algorithm is modified by an adaptive step-size control in sections 2.3 - 2.5, yielding the distributed adaptive Nesterov-Algorithm with event-triggered communication (DANA-EC 2.5.1). This is done firstly by modifying the Nesterov-Algorithm 2.1.3 in section 2.3, resulting in the adaptive Nesterov-Algorithm (ANA 2.3.2). Secondly, in section 2.4 it is shown how the ANA can be implemented distributedly which we call distributed adaptive Nesterov-Algorithm (DANA) in the following and finally, in section 2.5 the convergence of the DANA with event-triggered communication (DANA-EC 2.5.1) is shown.

2.1 Nesterov's optimal first order method

The content of this section was essentially published in [MUA14, sec. 2.1] (Meinel, Ulbrich, and Albrecht) and is reproduced here in similar form.

Nesterov's optimal first order scheme [Nes05, sec. 3], which we denote in the following by Nesterov-Algorithm (NA) for simplicity, is applicable for convex optimization problems

$$\min_{x \in Q} f(x), \tag{2.1}$$

where the constraint set $Q \subseteq E$ is a closed and convex subset of a real vector space E, and the objective function $f\colon Q \to \mathbb{R}$ is convex and continuously differentiable with Lipschitz continuous gradient, i.e., the gradient ∇f satisfies the following inequality [Nes05]:

$$\|\nabla f(x) - \nabla f(y)\|_{E^*} \leq L \|x - y\|_E \quad \forall x, y \in Q, \tag{2.2}$$

where $L > 0$ is the Lipschitz constant and $\|\cdot\|_{E^*}$ the norm that corresponds to the dual space E^* of E. However, in this work we consider the real vector space $E = \mathbb{R}^m$, i.e., we have $\|\cdot\|_{E^*} = \|\cdot\|_E = \|\cdot\|$ [Nes05], where by $\|\cdot\|$ the Euclidean norm is denoted. It follows that inequality (2.2) becomes

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \forall x, y \in Q. \tag{2.3}$$

Before the Nesterov-Algorithm is stated, we introduce the term prox-function after giving the definition of a strongly convex function by the following theorem.

Theorem 2.1.1. (Strongly convex function) [UU12, Theo. 6.3, 3.]
Let the function $d(x)\colon Q \to \mathbb{R}$ be continuously differentiable on an open environment of the convex set Q. Then $d(x)$ is strongly convex if and only if there exists a parameter $\mu > 0$ such that the following inequality is satisfied for all $x, y \in Q$:

$$d(y) - d(x) \geq \nabla d(x)^T (y - x) + \mu \|x - y\|^2. \tag{2.4}$$

In [Nes05], the parameter $\sigma = \mu/2 > 0$ is called convexity parameter and the strongly convexity property of a continuously differentiable function $d(x)$ is equivalent to [BT89, Prop. A.41]

$$(\nabla d(x) - \nabla d(y))^T (y - x) \geq \sigma \|x - y\|^2 \quad \forall x, y \in Q. \tag{2.5}$$

The following definition of a prox-function and the corresponding center slightly extends [NS08, Def. 2.4] by additionally demanding continuously differentiability.

Definition 2.1.2. (Prox-function) [NS08, based on Def. 2.4]
A continuously differentiable function $d(x)\colon Q \to \mathbb{R}$ is called prox-function if it is strongly convex and satisfies

$$d(x^0) = 0,$$

where

$$x^0 = \operatorname*{arg\,min}_{x \in Q} \; d(x)$$

is called the center of Q.

Finally, the initialization of the Nesterov-Algorithm [Nes05, p. 135] is done by choosing a prox-function $d(x)$ with convexity parameter $\sigma > 0$ whose center (minimum) x^0 serves as the starting point. Moreover, a positive sequence $\{\alpha_k\}_{k \geq 0}$ has to be chosen that occurs in the following algorithm as well as the quantities

$$\tau_k = \frac{\alpha_{k+1}}{A_{k+1}}, \quad \text{where} \quad A_k = \sum_{i=0}^{k} \alpha_i. \tag{2.6}$$

Algorithm 2.1.3. (Nesterov-Algorithm) [MUA14, Algo. 2.1]
For $k \geq 0$ do:

1. *Compute* $\nabla f(x^k)$.
2. *Find* $y^k = \underset{y \in Q}{\arg\min} \left\{ \left\langle \nabla f(x^k), y - x^k \right\rangle + \frac{L}{2} \left\| y - x^k \right\|^2 \right\}$.
3. *Find* $z^k = \underset{z \in Q}{\arg\min} \left\{ \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla f(x^j), z - x^j \right\rangle \right\}$.
4. *Set* $x^{k+1} = \tau_k z^k + (1 - \tau_k) y^k$.

We notice that some (redundant) constant terms are omitted in the argmin-problems of Algorithm 2.1.3 compared to the representation in [Nes05] in order to reveal its parallelizable nature if the prox-function $d(x)$ as well as Q have a suitable structure discussed more detailed in the following section 2.2.
We end this introduction to Nesterov's optimal first order scheme with the following result which merges [Nes05, Lem. 1] with [Nes05, Theo. 2] (with adopted wording).

Theorem 2.1.4. [MUA14, Theo. 2.2]
Let the sequence $\{\alpha_k\}_{k \geq 0}$ satisfy the condition

$$\alpha_0 \in (0,\, 1],\ \alpha_{k+1}^2 \leq A_{k+1},\ \alpha_k > 0,\ k \geq 0. \tag{2.7}$$

Then the relation

$$A_k f(y^k) \leq \Psi^k = \min_{z \in Q} \left\{ \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left(f(x^j) + \left\langle \nabla f(x^j), z - x^j \right\rangle \right) \right\}$$

holds for $k \geq 0$ and therefore

$$f(y^k) - f(x^{opt}) \leq \frac{L d(x^{opt})}{\sigma A_k},$$

where x^{opt} is an optimal solution to problem (2.1).

Proof. The proof is given in [Nes05, proof of Theo 2]. □

Finally, in [Nes05] the following choice of the sequence $\{\alpha_k\}_{k \geq 0}$ is proposed which satisfies conditions (2.7) in Theorem 2.1.4.

Lemma 2.1.5. [Nes05, Lem. 2]
For $k \geq 0$ define $\alpha_k = (k+1)/2$. Then

$$\tau_k = \frac{2}{k+3}, \quad A_k = \frac{(k+1)(k+2)}{4},$$

and conditions (2.7) are satisfied.

Applying Lemma 2.1.5 in Theorem 2.1.4 yields the following estimate of the gap between the function values at the optimal solution x^{opt} and the iterate y^k from step 2 of Algorithm 2.1.3 [Nes05, Theo. 2]:

$$f(y^k) - f(x^{\text{opt}}) \leq \frac{4Ld(x^{\text{opt}})}{\sigma(k+1)(k+2)}.$$

In other words, for a given accuracy $\epsilon > 0$ the gap is less than ϵ if

$$\frac{4Ld(x^{\text{opt}})}{\sigma(k+1)(k+2)} \leq \epsilon$$

which immediately shows the complexity of $\mathcal{O}\left(\sqrt{L/\epsilon}\right)$ iterations for the Nesterov-Algorithm 2.1.3 as given in [Nes05], where for comparison it is also noted that the standard gradient projection method applied to problem (2.1) needs $\mathcal{O}(1/\epsilon)$ iterations.

2.2 Distributed Nesterov-Algorithm with event-triggering

The content of this section was essentially published in [MUA14, sec. 2.2] (Meinel, Ulbrich, and Albrecht) and is reproduced here in similar form.

To be able to formulate Nesterov's Algorithm 2.1.3 in a way that it can be implemented in a distributed manner, we have to define a multi-agent network whose agents each control a different subblock of the optimization variable $x \in Q \subseteq \mathbb{R}^m$ of problem (2.1). To this end, let the multi-agent network consist of $s \leq m$ agents, where $agent_{x_l}$ controls subblock $x_l \in \mathbb{R}^{m_l}$ of $x = (x_1^T, \ldots, x_s^T)^T \in Q \subseteq \mathbb{R}^m$ and $\sum_{l=1}^{s} m_l = m$. For the ease of notation, we will omit the transpose symbols in this work whenever a vector is composed of different subvectors and the dimensions are clear from the context, e.g., $x = (x_1, \ldots, x_s) \in Q$.
It can be seen immediately that step 2 and 3 of Algorithm 2.1.3 are executable in parallel by the agents if the feasible set Q of problem (2.1) as well as the prox-function $d(x)$ are block-separable according to the partitioning of the optimization variable $x = (x_1, \ldots, x_s)$ which yields the following assumptions.

Assumptions 2.2.1. (Separability of Q and $d(x)$) (cf. [MUA14, Ass. 2.5])

1. *The feasible set $Q \in \mathbb{R}^m$ of problem (2.1) is block-separable in the following way:*

$$Q = Q_1 \times \ldots \times Q_s \;\; \textit{with} \;\; x_l \in Q_l \subseteq \mathbb{R}^{m_l} \;\; \textit{and} \;\; \sum_{l=1}^{s} m_l = m.$$

2. *The prox-function $d(x)$ with convexity parameter σ in Algorithm 2.1.3 is separable accordingly, i.e.,*

$$d(x) = \sum_{l=1}^{s} d_l(x_l) \ \ with \ \ x_l \in Q_l \subseteq \mathbb{R}^{m_l},$$

where $d_l \colon Q_l \to \mathbb{R}$ is a prox-function with convexity parameter σ for $l = 1, \ldots, s$.

For example, with Assumptions 2.2.1 step 3 of Algorithm 2.1.3 can be written as

$$\begin{aligned} z^k &= \operatorname*{arg\,min}_{z \in Q} \left\{ \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla f(x^j), z - x^j \right\rangle \right\} \\ &= \operatorname*{arg\,min}_{(z_1,\ldots,z_s) \in Q_1 \times \ldots \times Q_s} \left\{ \frac{L}{\sigma} \sum_{l=1}^{s} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \sum_{l=1}^{s} \left\langle \nabla_l f(x^j), z_l - x_l^j \right\rangle \right\} \\ &= \sum_{l=1}^{s} \operatorname*{arg\,min}_{z_l \in Q_l} \left\{ \frac{L}{\sigma} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla_l f(x^j), z_l - x_l^j \right\rangle \right\}, \end{aligned} \tag{2.8}$$

where $\nabla_l f(x) \in \mathbb{R}^{m_l}$ denotes subblock l of the gradient of $f(x)$. Obviously, the right-hand side in (2.8) can be solved in parallel by the agents, where $agent_{x_l}$ solves subproblem

$$\operatorname*{arg\,min}_{z_l \in Q_l} \left\{ \frac{L}{\sigma} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla_l f(x^j), z_l - x_l^j \right\rangle \right\}, \tag{2.9}$$

corresponding to subblock x_l that he is responsible for.

However, $agent_{x_l}$ needs to communicate with other agents in the multi-agent network in order to be able to compute $\nabla_l f(x^k)$ in each iteration. Generally, and especially for problems arising in large-scale networks, subblock $\nabla_l f(x)$ does not depend on all other subblocks of the optimization variable, i.e., the communication topology of the multi-agent network is usually not complete. In the following, we describe the communication topology of the multi-agent network by a graph called information dependency graph (IDG) which is defined as follows.

Definition 2.2.2. (Information dependency graph)

A graph with s nodes is called information dependency graph (IDG) if node l is connected to node j by an undirected line provided that subblock $\nabla_l f(x)$ of the gradient of the objective function in problem (2.1) depends on subblock $x_j \neq x_l$ of the optimization variable $x = (x_1, \ldots, x_s)$.
Moreover, the set $N_{IDG}(l) = \{j_1, \ldots, j_{\eta_l}\}$ denotes the set of η_l neighbors of node l.

We notice that the IDG is defined as an undirected graph as the above dependencies are mutual. Finally, with Assumptions 2.2.1 it is straight forward to state a distributed version of Algorithm 2.1.3 which is called distributed Nesterov-Algorithm (DNA) in the following.

Algorithm 2.2.3. (DNA)
For $l = 1,\dots,s$ and $k \geq 0$ do in parallel:

1. *Compute* $\nabla_l f(x^k)$.
2. *Find* $y_l^k = \underset{y_l \in Q_l}{\arg\min} \left\{ \left\langle \nabla_l f(x^k), y_l - x_l^k \right\rangle + \frac{L}{2} \left\| y_l - x_l^k \right\|^2 \right\}$.
3. *Find* $z_l^k = \underset{z_l \in Q_l}{\arg\min} \left\{ \frac{L}{\sigma} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla_l f(x^j), z_l - x_l^j \right\rangle \right\}$.
4. *Set* $x_l^{k+1} = \tau_k z_l^k + (1 - \tau_k) y_l^k$.
5. *Send* x_l^{k+1} *to* $agent_{x_j}$ *if* $l \in N_{IDG}(j)$.

In step 5 of the DNA 2.2.3, the iterates have to be exchanged in every iteration which may result in a large communication traffic that is undesirable especially for capacity limited wireless communication networks [WL10, sec. I]. To remedy this drawback, we enhance the DNA 2.2.3 with event-triggered communication similarly to [ZC10, sec. 2 - 3] and [WL10, sec. 4] by defining an outdated vector $x^{l,k}$ that is available to $agent_{x_l}$ in iteration k of the DNA 2.2.3.

Definition 2.2.4. (Outdated subblocks of the optimization variable)
Without restriction and for the ease of notation assume that $0 \in Q$. *For* $l = 1,\dots,s$ *and* $k \geq 0$ *let* $x^{l,k} = \left(x_1^{l,k},\dots,x_s^{l,k}\right) \in Q \subseteq \mathbb{R}^m$ *denote the outdated vector available to* $agent_{x_l}$ *in iteration* k*, whose subblocks satisfy*

$$\begin{cases} \left\| x_j^{l,k} - x_j^k \right\|_1 \leq \Delta_k & \text{if } j \in N_{IDG}(l), \\ x_j^{l,k} = 0 & \text{if } j \notin N_{IDG}(l) \cup \{l\}, \\ x_l^{l,k} = x_l^k & \text{else,} \end{cases} \tag{2.10}$$

for a given threshold $\Delta_k \geq 0$ *with* $\Delta_0 = 0$.

In [ZC10], event-triggered communication is used for distributedly solving an unconstrained problem, where convergence is guaranteed if, i.a., the objective function is continuously differentiable with Lipschitz continuous gradient, whereas in [WL10] event-triggered communication is applied to distributedly solve the DC-OPF problem by minimizing a corresponding unconstrained augmented cost function with a gradient scheme, where convergence is guaranteed for convex, strictly increasing, and differentiable cost functions. In contrast to these approaches, our choice of the threshold Δ_k does not depend on the state of the optimization variable subblocks as will be discussed.

Finally, we propose the following distributed Nesterov-Algorithm with event-triggered communication (DNA-EC) that extends the DNA 2.2.3 by letting the agents use the outdated iterates $x^{l,k}$ from Definition 2.2.4 instead of x^k, and by the addition of the separable term

$$L\Delta_k \eta_l \left\| y_l - x_l^k \right\|_1 \tag{2.11}$$

in step 2 of the DNA which is necessary for the proof of convergence.

Algorithm 2.2.5. (DNA-EC) [MUA14, Algo. 2.7]
For $l = 1, \dots, s$ and $k \geq 0$ do in parallel:

1. *Compute $\nabla_l f(x^{l,k})$.*
2. *Find* $y_l^k = \underset{y_l \in Q_l}{\arg\min} \left\{ \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle + L\Delta_k \eta_l \left\| y_l - x_l^k \right\|_1 + \frac{L}{2} \left\| y_l - x_l^k \right\|^2 \right\}.$
3. *Find* $z_l^k = \underset{z_l \in Q_l}{\arg\min} \left\{ \frac{L}{\sigma} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right\}.$
4. *Set* $x_l^{k+1} = \tau_k z_l^k + (1 - \tau_k) y_l^k.$
5. *Send x_l^{k+1} if necessary: For $j \in N_{IDG}(l)$*
   if $\left\| x_l^{j,k} - x_l^{k+1} \right\|_1 > \Delta_{k+1}$ ***then***
   set $x_l^{j,k+1} = x_l^{k+1}$ and send $x_l^{j,k+1}$ to $agent_{x_j}$.
   else
   set $x_l^{j,k+1} = x_l^{j,k}$ and signal that no data will be sent.

Obviously, the DNA-EC 2.2.5 coincides with the DNA 2.2.3 for the choice $\Delta_k = 0$, $k \geq 0$, as in this case we have $\nabla_l f(x^{l,k}) = \nabla_l f(x^k)$ for $l = 1, \dots, s$ according to Definition 2.2.4.
The larger the threshold Δ_k is chosen, the less exchange of iterates is needed in step 5, and the less computations of $\nabla_l f(x^{l,k})$ in step 1 have to be executed. However, if Δ_k is chosen too large, the DNA-EC may not converge or may need more iterations than the DNA to compute a comparable solution which possibly results in a higher information exchange , i.e., the choice of Δ_k is crucial.
Finally, the following example is given to show that the nice structures of the subproblems in the DNA-EC often allow analytical solutions which is favorable with respect to the computational efficiency.

Example 2.2.6. [MUA14, Ex. 2.8]
For $Q = Q_1 \times Q_2 \times \cdots \times Q_m$ with compact and convex sets $Q_i \subset \mathbb{R}$ and $d(z) = (\sigma/2)\,\|z\|^2$, the subproblem in step 2 of Algorithm 2.2.5 is given by

$$
\begin{aligned}
y_l^k &= \underset{y_l \in Q_l \subset \mathbb{R}}{\arg\min} \left\{ \nabla_l f(x^{l,k}) \left(y_l - x_l^k\right) + L\Delta_k \eta_l \left|y_l - x_l^k\right| + \frac{L}{2}\left(y_l - x_l^k\right)^2 \right\} \\
&= \begin{cases} h_l^+(y_l) & \text{if } y_l - x_l^k \geq 0, \\ h_l^-(y_l) & \text{if } y_l - x_l^k \leq 0, \end{cases}
\end{aligned}
$$

where

$$
\begin{aligned}
h_l^+(y_l) &= \overbrace{\left(\nabla_l f(x^{l,k}) + L\Delta_k \eta_l\right)}^{=\Delta_+} \left(y_l - x_l^k\right) + \frac{L}{2}\left(y_l - x_l^k\right)^2, \\
h_l^-(y_l) &= \underbrace{\left(\nabla_l f(x^{l,k}) - L\Delta_k \eta_l\right)}_{\Delta_-} \left(y_l - x_l^k\right) + \frac{L}{2}\left(y_l - x_l^k\right)^2.
\end{aligned}
$$

We have

$$
\begin{aligned}
h_l^{+\prime}(y_l) &= \Delta_+ + L\left(y_l - x_l^k\right) = 0 \iff y_l^+ = -\frac{\Delta_+}{L} + x_l^k, \\
h_l^{-\prime}(y_l) &= \Delta_- + L\left(y_l - x_l^k\right) = 0 \iff y_l^- = -\frac{\Delta_-}{L} + x_l^k,
\end{aligned}
$$

and with $Q_l = [\underline{Q}_l, \overline{Q}_l]$ and $\Delta_L = L\Delta_k \eta_l$ the minimum y_l^k over Q_l for $l = 1, \ldots, m$ is

$$
y_l^k = \max\left\{\min\left\{y_l^{opt}, \overline{Q}_l\right\}, \underline{Q}_l\right\}, \text{ where } y_l^{opt} = \begin{cases} -\frac{\nabla_l f(x^{l,k}) + \Delta_L}{L} + x_l^k & \text{if } -\frac{\nabla_l f(x^{l,k}) + \Delta_L}{L} \geq 0. \\ -\frac{\nabla_l f(x^{l,k}) - \Delta_L}{L} + x_l^k & \text{if } -\frac{\nabla_l f(x^{l,k}) - \Delta_L}{L} \leq 0. \\ x_l^k & \text{else.} \end{cases}
$$

Moreover, it is straight forward to see that the solution of the subproblem in step 3 of the DNA-EC 2.2.5 is given by

$$
z_l^k = \max\left\{\min\left\{-\frac{\sum_{j=0}^k \alpha_j \nabla_l f(x^{l,j})}{L}, \overline{Q}_l\right\}, \underline{Q}_l\right\}.
$$

We close this section with the proof of convergence of the DNA-EC 2.2.5 in Theorem 2.2.12 which is prepared subsequently.

Lemma 2.2.7. [MUA14, Lem. 2.9]
For y, x^k, $x^{l,k} \in Q$ and $k \geq 0$ the following inequality holds:

$$f(y) \leq f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle + L\Delta_k \sum_{l=1}^{s} \eta_l \left\| y_l - x_l^k \right\|_1 + \frac{L}{2} \left\| y - x^k \right\|^2.$$

Proof. [MUA14, proof of Lem. 2.9]
Similar to [ZC10, proof of Theo. 1] or [BT89, proof of Prop. 5.1, p. 529], we apply the Descent Lemma [BT89, Lem. 2.1, p. 203] which yields that the Lipschitz continuity assumption of $\nabla f(x)$ is equivalent to

$$f(y) \leq f(x) + \langle \nabla f(x),\, y - x \rangle + \frac{L}{2} \|y - x\|^2 \quad \forall x, y \in Q.$$

With the definition of $x^{l,k}$ in (2.10), we have for $x^k, y \in Q$ that

$$\begin{aligned}
f(y) &\leq f(x^k) + \left\langle \nabla f(x^k),\, y - x^k \right\rangle + \frac{L}{2} \left\| y - x^k \right\|^2 \\
&= f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle \\
&\quad + \sum_{l=1}^{s} \left\langle \nabla_l f(x^k) - \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle + \frac{L}{2} \left\| y - x^k \right\|^2 \\
&\leq f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle \\
&\quad + L \sum_{l=1}^{s} \left\| (x_{j_1}^k, \ldots, x_{j_{\eta_l}}^k) - (x_{j_1}^{l,k}, \ldots, x_{j_{\eta_l}}^{l,k}) \right\| \left\| y_l - x_l^k \right\| + \frac{L}{2} \left\| y - x^k \right\|^2 \qquad (2.12) \\
&\leq f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle \\
&\quad + L\Delta_k \sum_{l=1}^{s} \eta_l \left\| y_l - x_l^k \right\|_1 + \frac{L}{2} \left\| y - x^k \right\|^2,
\end{aligned}$$

where (2.12) is obtained by the usage of the Lipschitz continuity of $\nabla f(x)$ similarly to [BT89] or [ZC10], and the Cauchy-Schwarz inequality [Beu14, sec. 10.3]. □

Lemma 2.2.8. [MUA14, Lem. 2.10]
The application of Lemma 2.2.7 to $y^k = (y_1^k, \ldots, y_s^k)$ computed in step 2 of the DNA-EC 2.2.5 yields the following inequality for $k \geq 0$:

$$f(y^k) \leq f(x^k) + \min_{y \in Q} \left\{ \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l - x_l^k \right\rangle + L\Delta_k \sum_{l=1}^{s} \eta_l \left\| y_l - x_l^k \right\|_1 + \frac{L}{2} \left\| y - x^k \right\|^2 \right\}.$$

Proof. Obvious. □

In preparation for the last and main lemma that is needed for the proof of convergence of the DNA-EC 2.2.5, the following assumption is necessary to handle the error caused by the usage of event-triggered communication.

Assumption 2.2.9. (Boundedness of Q)
The closed and convex feasible set Q of problem (2.1) is bounded.

The boundedness of the set Q allows the definition of a diameter C of Q as

$$C = \max_{y,x \in Q} \|x - y\|_1, \tag{2.13}$$

where the 1-norm is chosen due to its separability. Following the notation in [Nes05, p. 133], define for $k \geq 0$ the problem

$$\Psi^k = \quad \min_{z \in Q} \left\{ \rho_k + \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right) \right\}, \tag{2.14}$$

which is related to step 3 of the DNA-EC as

$$z^k = \arg\min_{z \in Q} \left\{ \rho_k + \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right) \right\} \tag{2.15}$$

$$= \arg\min_{z \in Q} \left\{ \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right\}. \tag{2.16}$$

Here,

$$\rho_k = \eta_{\max} L C \sum_{j=0}^{k} \alpha_j \Delta_j, \tag{2.17}$$

where $\eta_{\max}$ is the maximal degree of the IDG, L is the Lipschitz constant of the gradient of $f(x)$ in problem (2.1), and α_j are the a priori chosen positive parameters in the DNA-EC 2.2.5 . Finally, define

$$E_0 = 0 \tag{2.18}$$

and

$$E_k = A_{k-1} \tau_{k-1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^k) - \nabla_l f(x^{l,k}), y_l^{k-1} - z_l^{k-1} \right\rangle \quad \text{for} \quad k \geq 1, \tag{2.19}$$

where the quantities A_{k-1} and τ_{k-1} are given in (2.6) for $k \geq 1$.

The following lemma extends [Nes05, Lem. 1] by additionally considering event-triggered communication.

Lemma 2.2.10. [MUA14, Lem. 2.11]
Let $\{\alpha_k\}_{k\geq 0}$ satisfy

$$\alpha_0 \in (0,\, 1],\, \alpha_{k+1}^2 \leq A_{k+1}, \tag{2.20}$$

and set

$$x^{k+1} = \tau_k z^k + (1-\tau_k) y^k, \tag{2.21}$$

where y^k and z^k are the optimal solutions in step 2 and 3 of the DNA-EC 2.2.5.
Then the following inequality holds for $k \geq 0$:

$$\Psi^k \geq A_k f(y^k) + \sum_{j=0}^{k} E_j. \tag{2.22}$$

Proof. [MUA14, proof of Lem. 2.11]
The proof follows [Nes05, proof of Lem. 1] and extends it by additionally considering event-triggered communication. For $k = 0$ we have

$$d(z) \geq \underbrace{d(x^0)}_{=0} + \underbrace{\nabla d(x^0)^T (z - x^0)}_{\geq 0} + \frac{\sigma}{2} \left\| z - x^0 \right\|^2 \geq \frac{\sigma}{2} \left\| z - x^0 \right\|^2,$$

due to the strongly convexity of $d(z)$ and the fact that x^0 minimizes $d(z)$ (cf. [BT89, Prop. 3.1]). It follows that

$$\begin{aligned}
\Psi^0 &= \min_{z \in Q} \left\{ \underbrace{\rho_0}_{=0} + \frac{L}{\sigma} d(z) + \alpha_0 \left(f(x^0) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,0}), z_l - x_l^0 \right\rangle \right) \right\} \\
&\geq \alpha_0 \min_{z \in Q} \left\{ \frac{L}{2\alpha_0} \left\| z - x^0 \right\|^2 + f(x^0) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,0}), z_l - x_l^0 \right\rangle \right\} \\
&\geq \alpha_0 f(y^0) = A_0 f(y^0) + \underbrace{E_0}_{=0},
\end{aligned}$$

where the last inequality follows with Lemma 2.2.8. Now assume that the relation $\Psi^k \geq A_k f(y^k) + \sum_{j=0}^{k} E_j$ holds for some $k \in \mathbb{N}_0$. As the function

$$h_k(z) = \rho_k + \frac{L}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right)$$

is strongly convex with convexity parameter L, it follows that

$$h_k(z) \geq \Psi^k + \frac{L}{2} \left\| z - z^k \right\|^2.$$

We obtain

$$\begin{aligned}
\Psi^{k+1} &= \min_{z\in Q}\left\{\rho_k + \eta_{\max} LC\alpha_{k+1}\Delta_{k+1} + \frac{L}{\sigma}d(z) + \sum_{j=0}^{k+1}\alpha_j\left(f(x^j) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,j}), z_l - x_l^j\right\rangle\right)\right\}\\
&\geq \min_{z\in Q}\left\{\Psi^k + \frac{L}{2}\left\|z - z^k\right\|^2 + \eta_{\max} LC\alpha_{k+1}\Delta_{k+1}\right.\\
&\quad \left. + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right)\right\}\\
&\geq \min_{z\in Q}\left\{\Psi^k + \frac{L}{2}\left\|z - z^k\right\|^2 + \eta_{\max} L\alpha_{k+1}\Delta_{k+1}\left\|z - z^k\right\|_1\right.\\
&\quad \left. + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right)\right\}.
\end{aligned}$$

Due to the convexity of f, the definition of x^{k+1} in (2.21), and the induction hypothesis, we have

$$\begin{aligned}
&\Psi^k + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right)\\
&\geq A_k f(y^k) + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right) + \sum_{j=0}^{k} E_j\\
&\geq A_k\left(f(x^{k+1}) + \left\langle \nabla f(x^{k+1}),\, y^k - x^{k+1}\right\rangle\right)\\
&\quad + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right) + \sum_{j=0}^{k} E_j\\
&= A_k\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), y_l^k - x_l^{k+1}\right\rangle\right)\\
&\quad + A_k\left(\sum_{l=1}^{s}\left\langle \nabla_l f(x^{k+1}) - \nabla_l f(x^{l,k+1}), y_l^k - x_l^{k+1}\right\rangle\right)\\
&\quad + \alpha_{k+1}\left(f(x^{k+1}) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1}\right\rangle\right) + \sum_{j=0}^{k} E_j\\
&= A_{k+1} f(x^{k+1}) + \alpha_{k+1}\left(\sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k\right\rangle\right)\\
&\quad + \underbrace{A_k\left(\sum_{l=1}^{s}\left\langle \nabla_l f(x^{k+1}) - \nabla_l f(x^{l,k+1}), \tau_k(y_l^k - z_l^k)\right\rangle\right)}_{=\,E_{k+1}} + \sum_{j=0}^{k} E_j,
\end{aligned}$$

where the last equality follows with (2.21) and the fact that $\tau_k = \alpha_{k+1}/A_{k+1}$ as

$$
\begin{aligned}
&= A_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), y_l^k - x_l^{k+1} \right\rangle + \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1} \right\rangle \\
&= -A_k \tau_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l^k - y_l^k \right\rangle \\
&\quad + \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k + z_l^k - x_l^{k+1} \right\rangle \\
&= -A_k \tau_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l^k - y_l^k \right\rangle \\
&\quad + \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k + (1-\tau_k)\left(z_l^k - y_l^k\right) \right\rangle \\
&= -\left(A_k + \alpha_{k+1}\right) \tau_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l^k - y_l^k \right\rangle \\
&\quad + \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k + z_l^k - y_l^k \right\rangle \\
&= \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k \right\rangle .
\end{aligned}
$$

The rest of the proof is almost identical to the final part of [Nes05, proof of Lem. 1]. From condition (2.20) and $\tau_k = \alpha_{k+1}/A_{k+1}$ it follows that ${A_{k+1}}^{-1} \geq {\tau_k}^2$ and we obtain

$$
\begin{aligned}
\Psi^{k+1} \geq\; & A_{k+1} f(x^{k+1}) + \min_{z\in Q} \Bigg\{ \eta_{\max} L \alpha_{k+1} \Delta_{k+1} \left\| z - z^k \right\|_1 + \frac{L}{2} \left\| z - z^k \right\|^2 \\
& + \alpha_{k+1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k \right\rangle \Bigg\} + \sum_{j=0}^{k+1} E_j \\
=\; & A_{k+1} f(x^{k+1}) + A_{k+1} \min_{z\in Q} \Bigg\{ \eta_{\max} L \Delta_{k+1} \tau_k \left\| z - z^k \right\|_1 + \frac{L}{2A_{k+1}} \left\| z - z^k \right\|^2 \\
& + \tau_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k \right\rangle \Bigg\} + \sum_{j=0}^{k+1} E_j \\
\geq\; & A_{k+1} f(x^{k+1}) + A_{k+1} \min_{z\in Q} \Bigg\{ \eta_{\max} L \Delta_{k+1} \tau_k \left\| z - z^k \right\|_1 + \frac{L}{2} {\tau_k}^2 \left\| z - z^k \right\|^2 \\
& + \tau_k \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - z_l^k \right\rangle \Bigg\} + \sum_{j=0}^{k+1} E_j .
\end{aligned} \tag{2.23}
$$

For $z \in Q$ let

$$y = \tau_k z + (1-\tau_k) y^k .$$

As $\tau_k \in [0,1]$, we have $y \in Q$ and with the definition of x^{k+1} in (2.21) we can write

$$y - x^{k+1} = \tau_k (z - z^k).$$

It follows that

$$
\begin{aligned}
&\min_{z\in Q}\left\{\eta_{\max}L\Delta_{k+1}\tau_k\left\|z-z^k\right\|_1+\frac{L}{2}{\tau_k}^2\left\|z-z^k\right\|^2+\tau_k\left(\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,k+1}),z_l-z_l^k\right\rangle\right)\right\}\\
=&\min_{y\in\tau_k Q+(1-\tau_k)y^k}\left\{\eta_{\max}L\Delta_{k+1}\left\|y-x^{k+1}\right\|_1+\frac{L}{2}\left\|y-x^{k+1}\right\|^2+\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,k+1}),y_l-x_l^{k+1}\right\rangle\right\}\\
\geq&\min_{y\in Q}\left\{\eta_{\max}L\Delta_{k+1}\left\|y-x^{k+1}\right\|_1+\frac{L}{2}\left\|y-x^{k+1}\right\|^2+\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,k+1}),y_l-x_l^{k+1}\right\rangle\right\}\\
\geq&\min_{y\in Q}\left\{L\Delta_{k+1}\sum_{l=1}^{s}\eta_l\left\|y_l-x_l^{k+1}\right\|_1+\frac{L}{2}\left\|y-x^{k+1}\right\|^2+\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,k+1}),y_l-x_l^{k+1}\right\rangle\right\}\\
\geq& f(y^{k+1})-f(x^{k+1}),
\end{aligned}
\tag{2.24}
$$

where the last inequality follows with Lemma 2.2.8.
Substituting (2.24) in (2.23) yields $\Psi^{k+1}\geq A_{k+1}f(y^{k+1})+\sum_{j=0}^{k+1}E_j$. □

Remark 2.2.11. (Concave objective function)
A revision of the proof of Lemma 2.2.10 shows that the application of the DNA-EC 2.2.5 to maximize a concave and continuously differentiable function $f\colon Q\to\mathbb{R}$ (on a closed and convex set $Q\subseteq\mathbb{R}^m$ with Lipschitz continuous gradient) yields the following relation after k iterations:

$$
\Psi^k\leq A_k f(y^k)+\sum_{j=0}^{k}E_j, \tag{2.25}
$$

where

$$
\Psi^k=\max_{z\in Q}\left\{-\rho_k-\frac{L}{\sigma}d(z)+\sum_{j=0}^{k}\alpha_j\left(f(x^j)+\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,j}),z_l-x_l^j\right\rangle\right)\right\}.
$$

Finally, the following convergence result for the DNA-EC 2.2.5 can be given given which extends [Nes05, Theo. 2].

Theorem 2.2.12. (Convergence of the DNA-EC 2.2.5) [MUA14, Theo. 2.12]
Let y^k be generated by the DNA-EC 2.2.5 with α_k as in Lemma 2.1.5 and $\Delta_k=\beta\delta^k$, where $\delta\in(0,1)$ and $\beta\in\mathbb{R}_+$. Then for $k\geq 0$ the inequality

$$
f(y^k)-f(x^{opt})<\frac{\sigma 6\eta_{max}\beta LCg'(\delta)+4Ld(x^{opt})}{\sigma(k+1)(k+2)} \tag{2.26}
$$

holds, where x^{opt} is an optimal solution of problem (2.1), C is defined as in (2.13), and

$$
g(\delta)=\sum_{j=0}^{\infty}\delta^j=\frac{1}{1-\delta}\ \textit{for}\ \ \delta\in(0,1).
$$

Proof. [MUA14, proof of Theo. 2.12]
To prove the theorem, we have to derive an upper bound for the left-hand side Ψ^k in inequality (2.22) and a lower bound for $\sum_{j=0}^{k} E_j$ occurring in the right-hand side. We start with

$$\Psi^k = \min_{z\in Q}\left\{\rho_k + \frac{L}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,j}), z_l - x_l^j\right\rangle\right)\right\}$$
$$= \min_{z\in Q}\left\{\rho_k + \frac{L}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \left\langle\nabla f(x^j),\, z - x^j\right\rangle\right)\right.$$
$$\left. + \sum_{j=0}^{k}\alpha_j\left(\sum_{l=1}^{s}\left\langle\nabla_l f(x^{l,j}) - \nabla_l f(x^j), z_l - x_l^j\right\rangle\right)\right\}$$
$$\le \min_{z\in Q}\left\{\rho_k + \frac{L}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \left\langle\nabla f(x^j),\, z - x^j\right\rangle\right)\right.$$
$$\left. + \sum_{j=0}^{k}\alpha_j\left(\sum_{l=1}^{s} L\left\|(x_{j_1}^{l,j},\dots,x_{j_{\eta_l}}^{l,j}) - (x_{j_1}^{j},\dots,x_{j_{\eta_l}}^{j})\right\|\left\|z_l - x_l^j\right\|\right)\right\} \tag{2.27}$$
$$\le \rho_k + \frac{L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}}) + L\eta_{\max}\sum_{j=0}^{k}\alpha_j\Delta_j\left\|x^{\text{opt}} - x^j\right\|_1 \tag{2.28}$$
$$\le 2\eta_{\max}LC\sum_{j=0}^{k}\alpha_j\Delta_j + \frac{L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}})$$
$$= \eta_{\max}\beta LC\sum_{j=1}^{k}(j+1)\delta^j + \frac{L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}})$$
$$< \eta_{\max}\beta LC g'(\delta) + \frac{L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}}),$$

where we used the Lipschitz continuity assumption of the gradient of f to obtain (2.27) and the fact that f is convex (as it was done in [Nes05, proof of Theo. 2]) to obtain (2.28). Similarly, we derive a lower bound for the accumulated error $\sum_{j=0}^{k} E_j$.

$$\sum_{j=0}^{k} E_j \ge -\left|\sum_{j=1}^{k} A_{j-1}\tau_{j-1}\left(\sum_{l=1}^{s}\left\langle\nabla_l f(x^j) - \nabla_l f(x^{l,j}), y_l^{j-1} - z_l^{j-1}\right\rangle\right)\right|$$
$$\ge -\sum_{j=1}^{k} A_{j-1}\tau_{j-1}\left(\sum_{l=1}^{s} L\left\|(x_{j_1}^{j},\dots,x_{j_{\eta_l}}^{j}) - (x_{j_1}^{l,j},\dots,x_{j_{\eta_l}}^{l,j})\right\|\left\|y_l^{j-1} - z_l^{j-1}\right\|\right)$$
$$\ge -\eta_{\max}L\sum_{j=1}^{k} A_{j-1}\tau_{j-1}\Delta_j\left\|y_l^{j-1} - z_l^{j-1}\right\|_1$$
$$\ge -\eta_{\max}LC\sum_{j=1}^{k} A_{j-1}\tau_{j-1}\Delta_j = -\eta_{\max}\beta LC\sum_{j=1}^{k}\frac{j(j+1)}{4}\frac{2}{j+2}\delta^j$$
$$> -\frac{\eta_{\max}\beta LC}{2}g'(\delta).$$

Substituting these bounds in (2.22) results in

$$f(y^k) - f(x^{\text{opt}}) < \frac{4\left(\frac{3}{2}\eta_{\max}\beta LCg'(\delta) + \frac{L}{\sigma}d(x^{\text{opt}})\right)}{(k+1)(k+2)}.$$

□

The maximal degree $\eta_{\max}$ of the IDG in (2.26) is independent of the multi-agent network size if the structure of the objective function $f(x)$ in problem (2.1) is independent of the dimension of the optimization variable which holds in general for network related problems with variable size.
Finally, the choice of the threshold $\Delta_k = \beta\delta^k$ in Theorem 2.2.12 guarantees the convergence of the DNA-EC 2.2.5 for $\delta \in (0,1)$, however, the particular choices of $\beta > 0$ and $\delta \in (0,1)$ decide if the algorithm outputs a solution with less information exchange compared to the same solution obtained without event-triggered communication.

2.3 Adaptive Nesterov-Algorithm

The content of this section is in preparation for publication in [MU14] (Meinel and Ulbrich).

In this section, we enhance Nesterov's Algorithm 2.1.3 with an adaptive step-size control that is based on the work of Nesterov in [Nes13, sec. 3 - 4] and is proposed in [BCG11, sec. 5.3] as well. This step-size control is motivated by the following observation.

Remark 2.3.1. (cf. [MUA14, Rem. 2.4])
The subproblem in step 2 of the NA 2.1.3 is a projected gradient step with step size $1/L$ *as according to [BT89, sec. 3.3.2]*

$$\begin{aligned} y^k &= \operatorname*{arg\,min}_{y\in Q}\left\{\left\langle \nabla f(x^k), y - x^k\right\rangle + \frac{L}{2}\left\|y - x^k\right\|^2\right\} \\ &= \operatorname*{arg\,min}_{y\in Q}\left\{\frac{2}{L}\left\langle \nabla f(x^k), y - x^k\right\rangle + \left\|y - x^k\right\|^2 + \frac{1}{L^2}\left\|\nabla f(x^k)\right\|^2\right\} \\ &= \operatorname*{arg\,min}_{y\in Q}\left\{\left\|y - x^k + \frac{1}{L}\nabla f(x^k)\right\|^2\right\}. \end{aligned}$$

Remark 2.3.1 suggests the application of step-sizes $1/L_k$ with $L_k \leq L$ for $k \geq 0$ in order to reduce the number of iterations of the NA 2.1.3.

To this end, we define similar to [Nes13, (2.7)] for $L_k > 0$ and $x \in Q$ the quantities $F_{L_k}(x)$ and $a_{L_k}(x) \in Q$ by

$$F_{L_k}(x) = \min_{y \in Q} \left\{ f(x) + \langle \nabla f(x), y - x \rangle + \frac{L_k}{2} \|y - x\|^2 \right\},$$
$$a_{L_k}(x) = \arg F_{L_k}(x) = \arg\min_{y \in Q} \left\{ \nabla f(x)^T y + \frac{L_k}{2} \|y - x\|^2 \right\}.$$

Obviously, we have $y^k = a_L(x^k)$ in step 2 of the NA 2.1.3.
The adaptive Nesterov-Algorithm (ANA) is initialized just as the NA 2.1.3 by choosing a prox-function $d(x)$ with convexity parameter $\sigma > 0$ which defines the starting point x^0 as

$$x^0 = \arg\min_{x \in Q} d(x). \tag{2.29}$$

Moreover, choose $\gamma > 1$, $L_{-1} \in (0, L]$, $\{\alpha_k\}_{k \geq 0}$ with $\alpha_k > 0$, $\alpha_0 \in (0, 1]$, and set τ_k as well as A_k as in (2.6) for $k \geq 0$.

Algorithm 2.3.2. (ANA)
For $k \geq 0$ do:

1. *Compute $\nabla f(x^k)$ and set $L_k = L_{k-1}$.*
2. *Find $y^k = a_{L_k}(x^k) = \arg\min_{y \in Q} \left\{ \nabla f(x^k)^T y + \frac{L_k}{2} \left\| y - x^k \right\|^2 \right\}$.*
3. *Compute $\nabla f(y^k)$ and $f(y^k)$.*

   if

$$f(y^k) \leq F_{L_k}(x^k) = \min_{y \in Q} \left\{ f(x^k) + \left\langle \nabla f(x^k), y - x^k \right\rangle + \frac{L_k}{2} \left\| y - x^k \right\|^2 \right\} \tag{2.30}$$

   then

   continue with step 4.

   else

   set $L_k = L_k \gamma$ and go to step 2.

4. *Find $z^k = \arg\min_{z \in Q} \left\{ \frac{L_k}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \nabla f(x^j)^T z \right\}$.*
5. *Set $x^{k+1} = \tau_k z^k + (1 - \tau_k) y^k$.*

In [Nes13, (3.1)] step 3 is applied in a gradient method whose accelerated version [Nes13, (4.9)]) is similar to Algorithm 2.3.2 with the difference, i.a., that there the parameter α_{k+1} is defined by a solution of a quadratic equation that depends on L_k and thus has to be determined during the optimization process. Moreover, the step-size control implemented

by (2.30) is also proposed in [BCG11, sec. 5.3] for several optimal first order methods of similar type (see Remark 2.4.2).
Inequality (2.30) is always satisfied for $y^k = a_{L_k}(x^k)$ with $L_k \geq L$ (cf. [Nes13, Rem. 1]) due to the well known equivalency of the Lipschitz continuity of the gradient of f (2.3) and the following inequality [BT89, Lem. 2.1]:

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 \quad \forall x, y \in Q. \tag{2.31}$$

Alternatively, Lemma 2.2.7 with $\Delta_k = 0$ for $k \geq 0$ can be applied.
It follows immediately that the ANA 2.3.2 coincides with the NA 2.1.3 if $L_{-1} = L$ and in this case step 3 of the ANA 2.3.2 is redundant.
The convergence of the ANA follows immediately by substituting L with L_k in [Nes05, proof of Lem. 1], however, it is additional shown in section 2.5 as the ANA is a special case of the distributed adaptive Nesterov-Algorithm with event-triggered communication (DANA-EC) that is derived in the subsequent sections.

2.4 Distributed adaptive Nesterov-Algorithm

The content of this section is in preparation for publication in [MU14] (Meinel and Ulbrich).

As detailed in section 2.2, the subproblems in step 2 and 4 of the ANA 2.3.2 can be solved in parallel if Assumptions 2.2.1 hold, i.e., if the feasible set Q of problem (2.1) as well as the prox-function $d(x)$ are separable according to the partitioning of the optimization variable $x = (x_1, \dots, x_s)$ into s subblocks, each assigned to a different agent of a multi-agent network. Unfortunately, this is not the case for step 3 of the algorithm as the evaluation of $f(y^k)$ can not be done in parallel generally. To remedy this drawback, we consider instead the following inequality that is also proposed in [BCG11, sec. 5.3] (however, for different reasons (see Remark 2.4.2)) as it implies condition (2.30) which is shown in the following lemma, motivated by the Descent Lemma [BT89, Lem. 2.1, p. 203]:

$$\frac{L_k}{2} \left\| y^k - x^k \right\|^2 \geq \left\langle \nabla f(y^k) - \nabla f(x^k), y^k - x^k \right\rangle, \tag{2.32}$$

Lemma 2.4.1.
If $x^k, y^k \in Q$ with $y^k = a_{L_k}(x^k)$ satisfy (2.32) then the following inequality holds:

$$f(y^k) \leq \min_{y \in Q} \left\{ f(x^k) + \left\langle \nabla f(x^k), y - x^k \right\rangle + \frac{L_k}{2} \left\| y - x^k \right\|^2 \right\}.$$

Proof. It is well known that inequality

$$\langle \nabla f(y), x - y \rangle \leq f(x) - f(y) \tag{2.33}$$

is satisfied for all $x, y \in Q$ if f is a convex function [UU12, Theo. 6.3]. It follows that

$$\begin{aligned} f(y^k) &\leq f(x^k) + \left\langle \nabla f(x^k), y^k - x^k \right\rangle + \left\langle \nabla f(y^k) - \nabla f(x^k), y^k - x^k \right\rangle \\ &\leq f(x^k) + \left\langle \nabla f(x^k), y^k - x^k \right\rangle + \frac{L_k}{2} \left\| y^k - x^k \right\|^2, \end{aligned}$$

where the last inequality follows with (2.32). □

Remark 2.4.2.

Shortly before the completion of this thesis, we learned that in [KCD15] the adaptive step-size control implemented by (2.32) is used in a fast gradient method that is applied for distributed optimization in dual decomposition, however, a central coordinator is proposed to verify inequality (2.32) (see [KCD15, sec. 4.4]). The authors of [KCD15] adopted this step-size control from [BCG11], where the implication from Lemma 2.4.1 is used to prevent cancellation errors due to the application of (2.30) in several optimal first order methods (see [BCG11, sec. 5.3]).

Our motivation to use (2.32) instead of (2.30) is that the agents of the multi-agent network can verify the inequality in parallel and with local communication by using a distributed averaging consensus technique that finds widespread application in distributed optimization (see for instance [DUAH12] and therein [CCW10]). The following description of this consensus technique is taken from [CCW10, sec. 3 B]:
Let $A \in \mathbb{R}^{s \times s}$ be a symmetric, doubly stochastic ($\sum_j A_{lj} = \sum_l A_{lj} = 1$), and nonnegative matrix that has positive diagonal entries and is compatible with the undirected graph that represents the multi-agent network of $agent_{x_1}, \ldots, agent_{x_s}$, i.e., $A_{lj} = A_{jl} > 0 \iff agent_{x_l}$ and $agent_{x_j}$ are neighbors or $l = j$. Moreover, for $t = 0, 1, \ldots$ and $l = 1, \ldots, s$ define the recursion

$$\zeta_l^k(t+1) = A_{ll} \zeta_l^k(t) + \sum_{j \in N(l)} A_{lj} \zeta_j^k(t), \tag{2.34}$$

$$\nu_l^k(t+1) = A_{ll} \nu_l^k(t) + \sum_{j \in N(l)} A_{lj} \nu_j^k(t), \tag{2.35}$$

where $N(l)$ is the index set of $agent_{x_l}$'s neighbors in the multi-agent network and

$$\begin{aligned} \zeta_l^k(0) &= \frac{L_k}{2} \left\| y_l^k - x_l^k \right\|^2, \\ \nu_l^k(0) &= \left\langle \nabla_l f(y^k) - \nabla_l f(x^k), y_l^k - x_l^k \right\rangle. \end{aligned}$$

Then the following theorem holds.

Theorem 2.4.3. [CCW10, Theo. 1]
If the multi-agent network represented by A is connected, then

$$\lim_{t\to\infty} \zeta_l^k(t) = \frac{\sum_{j=1}^{s} \zeta_j^k(0)}{s} \text{ and } \lim_{t\to\infty} \nu_l^k(t) = \frac{\sum_{j=1}^{s} \nu_j^k(0)}{s} \text{ for } l = 1,\ldots,s$$

and it follows that

$$\lim_{t\to\infty} \frac{\zeta_l^k(t)}{\nu_l^k(t)} = \frac{\sum_{j=1}^{s} \zeta_j^k(0)}{\sum_{j=1}^{s} \nu_j^k(0)}. \tag{2.36}$$

Moreover, in [CCW10, proof of Theo. 1] it is shown that $\zeta_l^k(t)$ and $\nu_l^k(t)$ converge geometrically to $\sum_{j=1}^{s} \zeta_j^k(0)/s$ and $\sum_{j=1}^{s} \nu_j^k(0)/s$ with $t \to \infty$. Finally, Theorem 2.4.3 shows that each $agent_{x_l}$ can verify condition (2.32) with local communication by checking if $\zeta_l^k(t)/\nu_l^k(t) \geq 1$ is satisfied for sufficiently large t, and it follows that this modification of the ANA 2.3.2 can be implemented in parallel, resulting in the distributed adaptive Nesterov-Algorithm (DANA) which coincides for $\Delta_k = 0$ with the DANA-EC 2.5.1 presented in the next section. Regarding the choice of the sufficiently large t, the following stopping criterion for the consensus algorithm is proposed in [CCW10, (14)]:
For $l = 1,\ldots,s$ stop consensus if

$$\frac{\left|\zeta_l^k(t) - \zeta_l^k(t-1)\right|}{\left|\zeta_l^k(t-1)\right|} \leq \epsilon_{\text{cons}} \quad \text{and} \quad \frac{\left|\nu_l^k(t) - \nu_l^k(t-1)\right|}{\left|\nu_l^k(t-1)\right|} \leq \epsilon_{\text{cons}}, \tag{2.37}$$

where $\epsilon_{\text{cons}} > 0$ is the desired accuracy.

Remark 2.4.4. (Metropolis rule)
In [CCW10, p. 1149], the following Metropolis rule from [XBL06, sec. 2] is described which allows to build the components of the matrix A, that is used in recursion (2.34) and (2.35), with neighborhood information. Let

$$A_{lj} = \begin{cases} \frac{1}{1+\max(|N(l)|,|N(j)|)} & \text{if } (l,j) \in \mathcal{E},\ l \neq j, \\ 1 - \sum_{j\in N(l)\setminus\{l\}} A_{lj} & \text{if } l = j, \\ 0 & \text{otherwise}, \end{cases} \tag{2.38}$$

where $\mathcal{E}$ is the set of lines of the graph that represents the multi-agent network.

2.5 Distributed adaptive Nesterov-Algorithm with event-triggering

The content of this section is in preparation for publication in [MU14] (Meinel and Ulbrich).

Finally, event-triggered communication can be incorporated into the DANA similarly as in section 2.2, yielding the distributed adaptive Nesterov-Algorithm with event-triggered communication (DANA-EC) .
To this end, let $x^{l,k} \in Q$ denote the outdated vector introduced in Definition 2.2.4 and moreover, let $y^{l,k} \in Q$ be defined accordingly with the difference that for $y^{l,k}$ the threshold $\widetilde{\Delta}_k = (L_k/L)\Delta_k$ is considered.
The initialization of the DANA-EC is done by choosing a starting point x^0 according to (2.29) and Assumptions 2.2.1, i.e., x^0 is the minimum of a prox-function $d(x) = \sum_{l=1}^{s} d_l(x_l)$ with convexity parameter $\sigma > 0$, where $d(x^0) = 0$. Moreover, choose $\gamma > 1$, $L_{-1} \in (0, L]$, $\{\alpha_k\}_{k \geq 0}$ with $\alpha_k > 0$, $\alpha_0 \in (0, 1]$, and set τ_k as well as A_k as in (2.6) for $k \geq 0$. Finally, let $y_j^{l,-1} = x_j^0$ for $l, j = 1, \ldots s$.

Algorithm 2.5.1. (DANA-EC)
For $l = 1,\dots,s$ and $k \geq 0$ do in parallel:

1. *Compute* $\nabla_l f(x^{l,k})$ *and set* $L_k = L_{k-1}$.
2. *Find* $y_l^k = \underset{y_l \in Q_l}{\arg\min} \left\{ \nabla_l f(x^{l,k})^T y_l + L_k \Delta_k \eta_l \left\| y_l - x_l^k \right\|_1 + \frac{L_k}{2} \left\| y_l - x_l^k \right\|^2 \right\}$.
3. ***if*** $L_k < L$ ***then***

   (a) *Send* y_l^k *if necessary: For* $j \in N_{IDG}(l)$
   if $\|y_l^{j,k-1} - y_l^k\|_1 > \frac{L_k}{L}\Delta_k$ ***then***
   set $y_l^{j,k} = y_l^k$ *and send* $y_l^{j,k}$ *to agent*$_{x_j}$.
   else
   set $y_l^{j,k} = y_l^{j,k-1}$ *and signal that no data will be sent.*

   (b) *Compute* $\nabla_l f(y^{l,k})$ *and check with consensus*
   if

$$\frac{L_k}{2} \left\| y^k - x^k \right\|^2 \geq \sum_{l=1}^{s} \left\langle \nabla_l f(y^{l,k}) - \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle \tag{2.39}$$

   then
   continue with step 4.
   else
   set $L_k = L_k \gamma$ *and go to step 2.*

4. *Find* $z_l^k = \underset{z_l \in Q_l}{\arg\min} \left\{ \frac{L_k}{\sigma} d_l(z_l) + \sum_{j=0}^{k} \alpha_j \nabla_l f(x^{l,j})^T z_l \right\}$.
5. *Set* $x_l^{k+1} = \tau_k z_l^k + (1-\tau_k) y_l^k$.
6. *Send* x_l^{k+1} *if necessary: For* $j \in N_{IDG}(l)$
   if $\|x_l^{j,k} - x_l^{k+1}\|_1 > \Delta_{k+1}$ ***then***
   set $x_l^{j,k+1} = x_l^{k+1}$ *and send* $x_l^{j,k+1}$ *to agent*$_{x_j}$.
   else
   set $x_l^{j,k+1} = x_l^{j,k}$ *and signal that no data will be sent.*

Remark 2.5.2.

1. *The DANA-EC 2.5.1 coincides with the DNA-EC 2.2.5 for the choice $L_{-1} = L$ as in this case step 3 of the DANA-EC is not executed.*

2. *Different from step 3 of the ANA 2.3.2, it has to be checked in step 3 of the DANA-EC 2.5.1 if $L_k < L$ which is due to the fact that condition (2.39) might not be satisfied even if $L_k \geq L$. However, we show in the following that the convergence of the DANA-EC is already guaranteed if condition (2.39) is satisfied for $L_k < L$ or if y_l^k in step 2 is computed for some $L_k \geq L$.*

Finally, as described in the previous section 2.4, condition (2.39) can be checked in parallel with local communication by applying the averaging consensus technique with

$$\zeta_l^k(0) = \frac{L_k}{2}\left\|y_l^k - x_l^k\right\|^2,$$
$$\nu_l^k(0) = \left\langle \nabla_l f(y^{l,k}) - \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle.$$

The rest of this section contains the proof of convergence of the DANA-EC 2.5.1. To this end, we define similarly as in section 2.3 the quantities $\widetilde{F}_{L_k}(x), \tilde{a}_{L_k}(x) \in Q$ for $L_k > 0$ and $x, x^{l,k} \in Q$ by

$$\widetilde{F}_{L_k}(x) = \min_{y\in Q}\left\{ f(x) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,k}), y_l - x_l\right\rangle + L_k\Delta_k \sum_{l=1}^{s}\eta_l \left\|y_l - x_l\right\|_1 + \frac{L_k}{2}\left\|y - x\right\|^2 \right\},$$
$$\tilde{a}_{L_k}(x) = \arg\widetilde{F}_{L_k}(x) = \arg\min_{y\in Q}\left\{ \sum_{l=1}^{s}\nabla_l f(x^{l,k})^T y_l + L_k\Delta_k \sum_{l=1}^{s}\eta_l \left\|y_l - x_l\right\|_1 + \frac{L_k}{2}\left\|y - x\right\|^2 \right\}.$$

In the following lemma we show that condition (2.39) implies $f(\tilde{a}_{L_k}(x^k)) \leq \widetilde{F}_{L_k}(x^k)$.

Lemma 2.5.3.
For $y^k = \tilde{a}_{L_k}(x^k) \in Q$ and $x^k, x^{l,k}, y^{l,k} \in Q$, where $y^{l,k}$ satisfies (2.10) with respect to y^k for the threshold $\widetilde{\Delta}_k = (L_k/L)\Delta_k$, inequality

$$\frac{L_k}{2}\left\|y^k - x^k\right\|^2 \geq \sum_{l=1}^{s}\left\langle \nabla_l f(y^{l,k}) - \nabla_l f(x^{l,k}), y_l^k - x_l^k\right\rangle$$

from step 3 of the DANA-EC 2.5.1 implies that

$$f(y^k) \leq \widetilde{F}_{L_k}(x^k). \tag{2.40}$$

Proof. The convexity of f and the Lipschitz continuity of its gradient yield

$$\begin{aligned}
f(y^k) &\leq f(x^k) + \left\langle \nabla f(y^k), y^k - x^k \right\rangle \\
&= f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle + \sum_{l=1}^{s} \left\langle \nabla_l f(y^k) - \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle \\
&= f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle + \sum_{l=1}^{s} \left\langle \nabla_l f(y^k) - \nabla_l f(y^{l,k}), y_l^k - x_l^k \right\rangle \\
&\quad + \sum_{l=1}^{s} \left\langle \nabla_l f(y^{l,k}) - \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle \\
&\leq f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle + \sum_{l=1}^{s} L \left\| (y_{j_1}^k, \ldots, y_{j_{\eta_l}}^k) - (y_{j_1}^{l,k}, \ldots, y_{j_{\eta_l}}^{l,k}) \right\| \left\| y_l^k - x_l^k \right\| \\
&\quad + \sum_{l=1}^{s} \left\langle \nabla_l f(y^{l,k}) - \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle \\
&\leq f(x^k) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k}), y_l^k - x_l^k \right\rangle + L_k \Delta_k \sum_{l=1}^{s} \eta_l \left\| y_l^k - x_l^k \right\|_1 + \frac{L_k}{2} \left\| y^k - x^k \right\|^2.
\end{aligned}$$

□

To proof the convergence of the DANA-EC 2.5.1, we derive a result in the following that extends Lemma 2.2.10. To this end, we have to modify the definition of ρ_k (2.17) in section 2.2 for $k \geq 0$ by

$$\rho_k = \eta_{\max} L_k C \sum_{j=0}^{k} \alpha_j \Delta_j + \sum_{j=0}^{k} \left(L_j - L_{j-1} \right) C, \tag{2.41}$$

where C is the diameter of the set Q which is assumed to be bounded, i.e., Assumption 2.2.9 holds. The other quantities needed for the proof of Lemma 2.2.10 are used here as well and repeated for convenience, where

$$\Psi^k = \min_{z \in Q} \left\{ \rho_k + \frac{L_k}{\sigma} d(z) + \sum_{j=0}^{k} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right) \right\} \tag{2.42}$$

and

$$E_k = A_{k-1} \tau_{k-1} \sum_{l=1}^{s} \left\langle \nabla_l f(x^k) - \nabla_l f(x^{l,k}), y_l^{k-1} - z_l^{k-1} \right\rangle \quad \text{for} \quad k \geq 1, \tag{2.43}$$

where $E_0 = 0$. The following lemma extends Lemma 2.2.10 from section 2.2.

Lemma 2.5.4.
Let $\{\alpha_k\}_{k\geq 0}$ satisfy

$$\alpha_0 \in (0,\, 1],\, \alpha_{k+1}^2 \leq A_{k+1}, \tag{2.44}$$

and set

$$x^{k+1} = \tau_k z^k + (1-\tau_k) y^k, \tag{2.45}$$

where y^k and z^k are the optimal solutions in step 2 and 4 of the DANA-EC 2.5.1.
Then the following inequality holds for $k \geq 0$:

$$\Psi^k \geq A_k f(y^k) + \sum_{j=0}^{k} E_j. \tag{2.46}$$

Proof. The proof follows the proof of Lemma 2.2.10. Due to the strongly convexity of $d(z)$ and the fact that $x_0 = \arg\min_{z\in Q} d(z)$, it follows with the optimality condition for convex problems that

$$d(z) \geq \frac{\sigma}{2} \left\| z - x^0 \right\|^2.$$

Moreover, as $L_0 = L_{-1}$, $\Delta_0 = 0$, and $\alpha_0 \in (0,1]$ we obtain

$$\begin{aligned}
\Psi^0 &= \min_{z\in Q} \left\{ \underbrace{\rho_0}_{=0} + \frac{L_0}{\sigma} d(z) + \alpha_0 \left(f(x^0) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,0}), z_l - x_l^0 \right\rangle \right) \right\} \\
&\geq \alpha_0 \min_{z\in Q} \left\{ \frac{L_0}{2\alpha_0} \left\| z - x^0 \right\|^2 + f(x^0) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,0}), z_l - x_l^0 \right\rangle \right\} \\
&\geq \alpha_0 f(y^0) = A_0 f(y^0) + \underbrace{E_0}_{=0},
\end{aligned}$$

where the last inequality follows with Lemma 2.5.3 if $L_0 < L$ or Lemma 2.2.7 if $L_0 \geq L$.

Assume that $\Psi^k \geq A_k f(y^k)$ holds for some $k \in \mathbb{N}_0$. We obtain

$$\begin{aligned}
\Psi^{k+1} \geq &\min_{z\in Q} \{\rho_k + \eta_{\max} L_{k+1} C\alpha_{k+1}\Delta_{k+1} + (L_{k+1} - L_k)\, C \\
&+\frac{L_{k+1}}{\sigma} d(z) + \sum_{j=0}^{k+1} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right) \Bigg\} \\
\geq &\min_{z\in Q} \{\rho_k + \eta_{\max} L_{k+1} C\alpha_{k+1}\Delta_{k+1} + (L_{k+1} - L_k)\, C \\
&+\frac{L_k}{\sigma} d(z) + \sum_{j=0}^{k+1} \alpha_j \left(f(x^j) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,j}), z_l - x_l^j \right\rangle \right) \Bigg\} \\
\geq &\min_{z\in Q} \Bigg\{ \Psi^k + \frac{L_k}{2} \left\| z - z^k \right\|^2 + \eta_{\max} L_{k+1} C\alpha_{k+1}\Delta_{k+1} + (L_{k+1} - L_k)\, C \\
&+\alpha_{k+1} \left(f(x^{k+1}) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1} \right\rangle \right) \Bigg\} \\
\geq &\min_{z\in Q} \Bigg\{ \Psi^k + \eta_{\max} L_{k+1} \alpha_{k+1}\Delta_{k+1} \left\| z - z^k \right\|_1 + \frac{L_k + L_{k+1} - L_k}{2} \left\| z - z^k \right\|^2 \\
&+\alpha_{k+1} \left(f(x^{k+1}) + \sum_{l=1}^{s} \left\langle \nabla_l f(x^{l,k+1}), z_l - x_l^{k+1} \right\rangle \right) \Bigg\}.
\end{aligned}$$

The rest of the proof works just like the final part of the proof of Lemma 2.2.10 by substituting L with L_{k+1} and using Lemma 2.5.3 if $L_{k+1} < L$ or Lemma 2.2.7 if $L_{k+1} \geq L$. □

Finally, the convergence of the DANA-EC 2.5.1 can be shown, extending Theorem 2.2.12.

Theorem 2.5.5. (Convergence of the DANA-EC 2.5.1)
Let y^k be generated by the DANA-EC 2.5.1 with α_k as in Lemma 2.1.5 and $\Delta_k = \beta\delta^k$, where $\delta \in (0,1)$ and $\beta \in \mathbb{R}_+$. Then for $k \geq 0$ the inequality

$$f(y^k) - f(x^{opt}) < \frac{\sigma\gamma LC\,(6\eta_{max}\beta g'(\delta) + 4) + 4\gamma L d(x^{opt})}{\sigma(k+1)(k+2)} \tag{2.47}$$

holds, where x^{opt} is an optimal solution of problem (2.1) and

$$g(\delta) = \sum_{j=0}^{\infty} \delta^j = \frac{1}{1-\delta} \quad \textit{for} \ \ \delta \in (0,1).$$

Proof. The proof follows the proof of Theorem 2.2.12. We have to derive an upper bound for Ψ^k in (2.46) and a lower bound for $\sum_{j=0}^{k} E_j$. Starting with Ψ^k we obtain

$$\begin{aligned}
\Psi^k &= \min_{z\in Q}\left\{\rho_k + \frac{L_k}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,j}), z_l - x_l^j\right\rangle\right)\right\}\\
&= \min_{z\in Q}\left\{\rho_k + \frac{L_k}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \left\langle \nabla f(x^j),\, z - x^j\right\rangle\right)\right.\\
&\quad\left. + \sum_{j=0}^{k}\alpha_j\left(\sum_{l=1}^{s}\left\langle \nabla_l f(x^{l,j}) - \nabla_l f(x^j), z_l - x_l^j\right\rangle\right)\right\}\\
&\le \min_{z\in Q}\left\{\rho_k + \frac{\gamma L}{\sigma}d(z) + \sum_{j=0}^{k}\alpha_j\left(f(x^j) + \left\langle \nabla f(x^j),\, z - x^j\right\rangle\right)\right.\\
&\quad\left. + \sum_{j=0}^{k}\alpha_j\left(\sum_{l=1}^{s}L\left\|(x_{j_1}^{l,j},\dots,x_{j_{\eta_l}}^{l,j}) - (x_{j_1}^{j},\dots,x_{j_{\eta_l}}^{j})\right\|\left\|z_l - x_l^j\right\|\right)\right\}\\
&\le \rho_k + \frac{\gamma L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}}) + \gamma L\eta_{\max}\sum_{j=0}^{k}\alpha_j\Delta_j\left\|x^{\text{opt}} - x^j\right\|_1\\
&\le \sum_{j=0}^{k}\left(L_j - L_{j-1}\right)C + 2\eta_{\max}\gamma LC\sum_{j=0}^{k}\alpha_j\Delta_j + \frac{\gamma L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}})\\
&= \eta_{\max}\beta\gamma LC\sum_{j=1}^{k}(j+1)\delta^j + (L_k - L_{-1})C + \frac{\gamma L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}})\\
&< \gamma LC\left(\eta_{\max}\beta g'(\delta) + 1\right) + \frac{\gamma L}{\sigma}d(x^{\text{opt}}) + A_k f(x^{\text{opt}}).
\end{aligned}$$

The lower bound for $\sum_{j=0}^{k} E_j$ is given by (see proof of Theorem 2.2.12)

$$\sum_{j=0}^{k} E_j > -\frac{\eta_{\max}\beta\gamma LC}{2}g'(\delta).$$

Substituting these bounds in (2.46) yields:

$$f(y^k) - f(x^{\text{opt}}) < \frac{4\gamma\left(LC\left(\frac{3}{2}\eta_{\max}\beta g'(\delta) + 1\right) + \frac{L}{\sigma}d(x^{\text{opt}})\right)}{(k+1)(k+2)}.$$

□

Remark 2.5.6. (Efficiency estimate for DANA-EC 2.5.1)
From equation (2.47) it can be seen that the DANA-EC 2.5.1 has the same efficiency estimate as the Nesterov-Algorithm 2.1.3 which is of the order $\mathcal{O}\left(\sqrt{L/\epsilon}\right)$ as described in section 2.1.

If no event-triggered communication is applied, i.e., $\Delta_k = 0$ for $k \ge 0$, the convergence of the ANA 2.3.2 and the DANA, respectively, follows immediately as a special case of Theorem 2.5.5 (cf. [Nes05, Theo. 2]).

Theorem 2.5.7. (Convergence of the (D)ANA)
Let y^k be generated by the ANA 2.3.2 with α_k as in Lemma 2.1.5 and $\Delta_k = \beta\delta^k$, where $\delta \in (0,1)$ and $\beta \in \mathbb{R}_+$. Then for $k \geq 0$ the inequality

$$f(y^k) - f(x^{opt}) \leq \frac{4L_k(d(x^{opt}) + C\sigma)}{\sigma(k+1)(k+2)}$$

holds, where x^{opt} is an optimal solution of problem (2.1).

Proof. The proof is identical to the proof of Theorem 2.5.5, considering that with $\Delta_k = 0$ for $k \geq 0$ it follows that $x^{l,k} = x^l$ for $l = 1, \ldots, s$, and the quantities ρ_k in (2.41) and E_k in (2.43) become

$$\rho_k = \sum_{j=0}^{k} \left(L_j - L_{j-1}\right) C$$
$$E_k = 0.$$

□

3 Distributed dual decomposition method with event-triggering

In this chapter, a version of the proximal center algorithm by Necoara and Suykens (PCA 3.1.2) is presented in section 3.1. More precisely, this version of [NS08, Algo. 3.2] is a simplification that does not guarantee monotonicity of the function values, however, for the convergence theory this feature is irrelevant as mentioned in [NS08, sec. II B] (see also [Nes05, sec. 3]), and we therefore neglect it as it hinders a full distributed implementation. However, for simplicity we use the term proximal center algorithm (PCA) in the following.
In section 3.2, the PCA is enhanced by the implementation of the DANA-EC 2.5.1 which yields the distributed adaptive proximal center algorithm with event-triggered communication (DAPCA-EC 3.2.2).

3.1 Proximal center algorithm

The content of this section was essentially published in [MUA14, sec. 3.1] (Meinel, Ulbrich, and Albrecht) and is reproduced here in similar form.

The proximal center algorithm is a dual decomposition method that applies Nesterov's accelerated first order scheme from section 2.1 (NA 2.1.3) and a smoothing technique from [Nes05, sec. 2] to find an approximate solution of a separable convex problem

$$\min_{x \in X} \sum_{i=1}^{n} \Phi_i(x_i) \tag{3.1a}$$

$$\text{s.t.} \sum_{i=1}^{n} A_i x_i = b_A, \tag{3.1b}$$

$$\sum_{i=1}^{n} B_i x_i \leq b_B, \tag{3.1c}$$

where the set $X = X_1 \times \cdots \times X_n$ is separable with compact and convex sets $X_i \in \mathbb{R}^{m_i}$. Moreover, the cost functions $\Phi_i \colon X_i \to \mathbb{R}$ are continuous and convex functions that are not required to be differentiable, and finally, the constraints in (3.1b) and (3.1c) are defined by given matrices $A_i \in \mathbb{R}^{m_A \times m_i}$ and $B_i \in \mathbb{R}^{m_B \times m_i}$, as well as $b_A \in \mathbb{R}^{m_A}$ and $b_B \in \mathbb{R}^{m_B}$.

In the PCA, the separable structure of problem (3.1) is exploited by forming its dual which is then smoothed in a way that preserves the separability of the dual objective function with respect to the subblocks of the primal optimization variable. To state the dual, consider the Lagrangian [GK02, sec. 6.2.1]

$$\mathcal{L}(x,\mu,\lambda) = \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \mu \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \lambda \right\rangle$$

of problem (3.1), where $\mu \in \mathbb{R}^{m_A}$ are dual multipliers related to the equality constraints (3.1b) and $\lambda \in \mathbb{R}^{m_B}$ are dual multipliers related to the inequality constraints (3.1c). The dual problem of (3.1) is then given by [GK02, (6.6)]

$$\max_{(\mu,\lambda)\in\mathbb{R}^{m_A}\times\mathbb{R}_+^{m_B}} f(\mu,\lambda), \tag{3.2}$$

where the dual objective function is defined as [GK02, Def. 6.6]

$$\begin{aligned} f(\mu,\lambda) &= \min_{x\in X} \mathcal{L}(x,\mu,\lambda) \\ &= \min_{x\in X} \left\{ \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \mu \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \lambda \right\rangle \right\}. \end{aligned} \tag{3.3}$$

It is well known from duality theory that the dual function $f(\mu,\lambda)$ is concave [GK02, Lem. 6.11]. Moreover, the optimal value of the dual problem (3.2) coincides with the optimal value of the primal problem (3.1) (i.e., strong duality holds) if the relative interior of the feasible set of (3.1) is not empty and if the optimal function value of (3.1) is finite [GK02, Theo. 6.13] which is assumed in the following. Finally, the primal optimal solution x^{opt} of problem (3.1) can be obtained by the evaluation of $f((\mu,\lambda)^{\text{opt}})$, where $(\mu,\lambda)^{\text{opt}}$ are (not necessarily unique) optimal dual multipliers that solve (3.2).

Obviously, the dual function can be evaluated distributedly by n agents which motivates to solve the dual problem instead of the primal problem in distributed optimization, however, as the primal solution $x(\mu,\lambda)$ of (3.3) is not necessarily unique, the dual objective function may not be smooth and an iterative scheme for convex problems with differentiable objective function, such as the Nesterov-Algorithm 2.1.3, can not be applied to maximize the dual objective function in parallel. To remedy this drawback, the authors of [NS08] propose to smooth the Lagrangian by (strongly convex) prox-functions $d_{x_i}\colon X_i \to \mathbb{R}$ with convexity parameters $\sigma_{x_i} > 0$ for $i = 1,\ldots,n$ that are scaled with a smoothing parameter $c > 0$, yielding the augmented dual function

$$\begin{aligned} f_c(\mu,\lambda) &= \min_{x\in X} \left\{ \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \mu \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \lambda \right\rangle + c\sum_{i=1}^{n} d_{x_i}(x_i) \right\} \\ &= \sum_{i=1}^{n} \min_{x_i\in X_i} \left\{ \Phi_i(x_i) + \langle A_i x_i, \mu\rangle + \langle B_i x_i, \lambda\rangle + c d_{x_i}(x_i) \right\} - \langle b_A, \mu\rangle - \langle b_B, \lambda\rangle \end{aligned} \tag{3.4}$$

which obviously is still evaluable in parallel with respect to the primal subblocks $x_1, \dots, x_n$. Moreover, the augmented dual function $f_c(\mu, \lambda)$ is continuously differentiable and has a Lipschitz continuous gradient as shown in the following theorem which slightly extends [NS08, Theo. 3.1], where only equality constraints are considered for (3.1):

Theorem 3.1.1. (Existence and Lipschitz continuity of ∇f_c)
The augmented dual objective function f_c in (3.4) is continuously differentiable with

$$\nabla f_c(\mu, \lambda) = \begin{pmatrix} \sum_{i=1}^n A_i x_i(\mu, \lambda) - b_A \\ \sum_{i=1}^n B_i x_i(\mu, \lambda) - b_B \end{pmatrix}, \tag{3.5}$$

where $x_i(\mu, \lambda)$ are the unique arguments of the minima in (3.4). Furthermore, ∇f_c is Lipschitz continuous with Lipschitz constant

$$L_c = \sum_{i=1}^n \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{c\sigma_{x_i}}. \tag{3.6}$$

Proof. The continuously differentiability of f_c is shown in [BT89, p.669] which is given as a reference in [NS08, proof of Theo. 3.1].
To prove the Lipschitz continuity of ∇f_c, we follow [NS08, proof of Theo. 3.1] and extend it by additionally considering inequality constraints.
The first order optimal condition for problems with continuously differentiable objective function that are constrained by a convex feasible set [BT89, Prop. 3.1] yields the following inequalities for given Lagrange multipliers (μ, λ) and (ν, γ):

$$\left\langle \sum_{i=1}^n \nabla\Phi_i(x_i(\mu, \lambda)) + \sum_{i=1}^n A_i^T \mu + \sum_{i=1}^n B_i^T \lambda + c \sum_{i=1}^n \nabla d_{x_i}(x_i(\mu, \lambda)), x_i(\nu, \gamma) - x_i(\mu, \lambda) \right\rangle \geq 0,$$
$$\left\langle \sum_{i=1}^n \nabla\Phi_i(x_i(\nu, \gamma)) + \sum_{i=1}^n A_i^T \nu + \sum_{i=1}^n B_i^T \gamma + c \sum_{i=1}^n \nabla d_{x_i}(x_i(\nu, \gamma)), x_i(\mu, \lambda) - x_i(\nu, \gamma) \right\rangle \geq 0.$$

From the convexity of Φ_i it follows that $\Phi_i + d_{x_i}$ is strongly convex with convexity parameter σ_{X_i} and with (2.5) as well as (2.4) we have

$$\begin{aligned} 0 \leq & \sum_{i=1}^n \Phi_i(x_i(\nu, \gamma)) + c \sum_{i=1}^n d_{x_i}(x_i(\nu, \gamma)) - \sum_{i=1}^n \Phi_i(x_i(\mu, \lambda)) - c \sum_{i=1}^n d_{x_i}(x_i(\mu, \lambda)) \\ & - \sum_{i=1}^n \frac{c\sigma_{X_i}}{2} \|x_i(\nu, \gamma) - x_i(\mu, \lambda)\|^2 + \sum_{i=1}^n \left\langle A_i^T \mu + B_i^T \lambda, x_i(\nu, \gamma) - x_i(\mu, \lambda) \right\rangle, \\ 0 \leq & \sum_{i=1}^n \Phi_i(x_i(\mu, \lambda)) + c \sum_{i=1}^n d_{x_i}(x_i(\mu, \lambda)) - \sum_{i=1}^n \Phi_i(x_i(\nu, \gamma)) - c \sum_{i=1}^n d_{x_i}(x_i(\nu, \gamma)) \\ & - \sum_{i=1}^n \frac{c\sigma_{X_i}}{2} \|x_i(\mu, \lambda) - x_i(\nu, \gamma)\|^2 + \sum_{i=1}^n \left\langle A_i^T \nu + B_i^T \gamma, x_i(\mu, \lambda) - x_i(\nu, \gamma) \right\rangle. \end{aligned}$$

Finally, the summation of both inequalities and the application of the Cauchy-Schwarz inequality yields (with $(\mu,\lambda) = (\mu^T,\lambda^T)^T$)

$$\begin{aligned}
&\sum_{i=1}^{n} c\sigma_{X_i} \|x_i(\mu,\lambda) - x_i(\nu,\gamma)\|^2 \\
\leq &\sum_{i=1}^{n} \left\langle A_i^T(\mu-\nu) + B_i^T(\lambda-\gamma), x_i(\nu,\gamma) - x_i(\mu,\lambda)\right\rangle \\
= &\sum_{i=1}^{n} \left\langle \left(A_i^T, B_i^T\right)((\mu,\lambda)-(\nu,\gamma)), x_i(\nu,\gamma) - x_i(\mu,\lambda)\right\rangle \\
= &\sum_{i=1}^{n} \left\langle (\mu,\lambda)-(\nu,\gamma), \left(A_i^T, B_i^T\right)^T x_i(\nu,\gamma) - \left(A_i^T, B_i^T\right)^T x_i(\mu,\lambda)\right\rangle
\end{aligned}$$

which is used to derive

$$\begin{aligned}
&\|\nabla f_c(\mu,\lambda) - \nabla f_c(\nu,\gamma)\|^2 \\
= &\left\| \sum_{i=1}^{n} \left[\left(A_i^T, B_i^T\right)^T x_i(\nu,\gamma) - \left(A_i^T, B_i^T\right)^T x_i(\mu,\lambda) \right] \right\|^2 \\
\leq &\sum_{i=1}^{n} \frac{\left\|\left(A_i^T, B_i^T\right)^T\right\|^2}{c\sigma_{X_i}} c\sigma_{X_i} \|x_i(\nu,\gamma) - x_i(\mu,\lambda)\|^2 \\
\leq &\sum_{i=1}^{n} \frac{\left\|\left(A_i^T, B_i^T\right)^T\right\|^2}{c\sigma_{X_i}} \sum_{i=1}^{n} c\sigma_{X_i} \|x_i(\nu,\gamma) - x_i(\mu,\lambda)\|^2 \\
\leq &\sum_{i=1}^{n} \frac{\left\|\left(A_i^T, B_i^T\right)^T\right\|^2}{c\sigma_{X_i}} \sum_{i=1}^{n} \left\langle (\mu,\lambda)-(\nu,\gamma), \left(A_i^T, B_i^T\right)^T x_i(\nu,\gamma) - \left(A_i^T, B_i^T\right)^T x_i(\mu,\lambda)\right\rangle \\
\leq &\sum_{i=1}^{n} \frac{\left\|\left(A_i^T, B_i^T\right)^T\right\|^2}{c\sigma_{X_i}} \|(\mu,\lambda)-(\nu,\gamma)\| \, \|\nabla f_c(\mu,\lambda) - \nabla f_c(\nu,\gamma)\| .
\end{aligned}$$

Dividing both sides of the above inequality by $\|\nabla f_c(\mu,\lambda) - \nabla f_c(\nu,\gamma)\|$ proves the Lipschitz continuity of ∇f_c with Lipschitz constant (3.6). □

In the PCA, the smoothed dual function $f_c(\mu,\lambda)$ is iteratively maximized with the Nesterov-Algorithm 2.1.3 (cf. Remark 2.2.11) which is possible according to Theorem 3.1.1. To state the PCA, we denote by $Q_A \times Q_B \subseteq \mathbb{R}^{m_A} \times \mathbb{R}_+^{m_B}$ a closed and convex feasible set for the dual multipliers (μ,λ) and consider the following augmented dual problem:

$$\max_{(\mu,\lambda)\in Q_A\times Q_B} f_c(\mu,\lambda). \tag{3.7}$$

The initialization is done accordingly to the initialization of the Nesterov-Algorithm 2.1.3 by choosing a prox-function $d(\mu,\lambda)\colon Q_A \times Q_B \to \mathbb{R}$ (with convexity parameter $\sigma > 0$)

which defines the starting point as $(\mu,\lambda)^0 = \arg\min_{(\mu,\lambda)\in Q_A\times Q_B} d(\mu,\lambda)$. Finally, the following version of [NS08, Algo. 3.2], that applies the NA 2.1.3 to problem (3.7) with $\{\alpha_k\}_{k\geq 0}$ chosen as in Lemma 2.1.5, can be stated as

Algorithm 3.1.2. (PCA) [MUA14, Algo. 3.1]
For $k \geq 0$ do:

1. *Given $(\mu,\lambda)^k \in Q_A \times Q_B$, for $i = 1,\dots,n$ compute*
$$x_i^{k+1} = \arg\min_{x_i\in X_i} \left\{ \Phi_i(x_i) + \left\langle \mu^k, A_i x_i \right\rangle + \left\langle \lambda^k, B_i x_i \right\rangle + c d_{x_i}(x_i) \right\}.$$
2. *Compute* $\nabla f_c((\mu,\lambda)^k) = \begin{pmatrix} \sum_{i=1}^n A_i x_i^{k+1} - b_A \\ \sum_{i=1}^n B_i x_i^{k+1} - b_B \end{pmatrix}$.
3. *Find* $(u,h)^k = \arg\max_{(u,h)\in Q_A\times Q_B} \left\{ \left\langle \nabla f_c((\mu,\lambda)^k), (u,h) - (\mu,\lambda)^k \right\rangle - \frac{L_c}{2} \left\| (u,h) - (\mu,\lambda)^k \right\|^2 \right\}.$
4. *Find* $(v,t)^k = \arg\max_{(v,t)\in Q_A\times Q_B} \left\{ -\frac{L_c}{\sigma} d(v,t) + \sum_{j=0}^k \frac{j+1}{2} \left\langle \nabla f_c((\mu,\lambda)^j), (v,t) - (\mu,\lambda)^j \right\rangle \right\}.$
5. *Set* $(\mu,\lambda)^{k+1} = \frac{2}{k+3}(v,t)^k + \frac{k+1}{k+3}(u,h)^k$.

Obviously, the PCA 3.1.2 can be implemented distributedly if the set $Q_A \times Q_B$ and the prox-function $d(\mu,\lambda)$ are separable according to Assumptions 2.2.1.
To state a convergence result for the PCA 3.1.2, we denote in the following by $M^{\text{opt}} \times \Lambda^{\text{opt}}$ the set of optimal dual multipliers for the dual problem (3.2) and assume that

$$M^{\text{opt}} \times \Lambda^{\text{opt}} \cap Q_A \times Q_B \neq \emptyset.$$

The following lemma slightly extends [NS08, Lem. 3.3 and the conclusions afterwards] by additionally considering inequality constraints and bounds the primal gap, i.e., the distance from the optimal objective function value of problem (3.1):

Lemma 3.1.3. [MUA14, Lem. 3.2]
For every $(\mu,\lambda)^{opt} \in M^{opt} \times \Lambda^{opt}$, $(\mu,\lambda) \in Q_A \times Q_B$, and $x_i \in X_i$ for $i = 1,\dots,n$, the following inequalities hold:

$$-\left\|(\mu,\lambda)^{opt}\right\| \left\| \begin{matrix} \sum_{i=1}^n A_i x_i - b_A \\ \left[\sum_{i=1}^n B_i x_i - b_B\right]^+ \end{matrix} \right\| \leq \sum_{i=1}^n \Phi_i(x_i) - f^{opt} \leq \sum_{i=1}^n \Phi_i(x_i) - f_0(\mu,\lambda), \tag{3.8}$$

where $f^{opt} = f_0((\mu,\lambda)^{opt})$ and $[\cdot]^+$ denotes the projection onto $\mathbb{R}_+^{m_B}$.

Proof. [MUA14, proof of Lem. 3.2]
The proof combines [NS08, Rem. 3.8] with [NS08, proof of Lem. 3.3], where the bounds on the primal gap are given for equality constrained convex problems. The lower bound of (3.8) can be derived as follows:

$$\begin{aligned} f^{\text{opt}} &= \min_{x_i \in X_i (i=1,\dots,n)} \left\{ \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \mu^{\text{opt}} \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \lambda^{\text{opt}} \right\rangle \right\} \\ &\le \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \mu^{\text{opt}} \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \lambda^{\text{opt}} \right\rangle \\ &\le \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \begin{pmatrix} |\sum_{i=1}^{n} A_i x_i - b_A| \\ \sum_{i=1}^{n} B_i x_i - b_B \end{pmatrix}, \begin{pmatrix} |\mu^{\text{opt}}| \\ \lambda^{\text{opt}} \end{pmatrix} \right\rangle \\ &\le \sum_{i=1}^{n} \Phi_i(x_i) + \left\langle \begin{bmatrix} |\sum_{i=1}^{n} A_i x_i - b_A| \\ \sum_{i=1}^{n} B_i x_i - b_B \end{bmatrix}^{+}, \begin{pmatrix} |\mu^{\text{opt}}| \\ \lambda^{\text{opt}} \end{pmatrix} \right\rangle \\ &\le \sum_{i=1}^{n} \Phi_i(x_i) + \left\| \begin{matrix} \sum_{i=1}^{n} A_i x_i - b_A \\ [\sum_{i=1}^{n} B_i x_i - b_B]^{+} \end{matrix} \right\| \left\| (\mu, \lambda)^{\text{opt}} \right\|, \end{aligned}$$

where the last inequality follows by the Cauchy-Schwarz inequality. The upper bound on the primal gap is obvious as $f^{\text{opt}} = f_0((\mu,\lambda)^{\text{opt}}) \ge f_0(\mu,\lambda)$ for all $(\mu,\lambda) \in Q_A \times Q_B$. □

Define by [NS08, sec. III A]

$$D_{X_i} \ge \max_{x_i \in X_i} d_{x_i}(x_i) \tag{3.9}$$

an upper bound on the value range of d_{x_i} over the compact set X_i for $i = 1,\dots,n$. The upper bound on the primal gap given in (3.8) can be expressed in terms of the primal iterates computed with the PCA 3.1.2, such that no dual function evaluation is necessary as shown in the following theorem which slightly extends [NS08, Theo. 3.4] by additionally considering inequality constraints.

Theorem 3.1.4. [MUA14, Theo. 3.9 with $P(\Delta) = 0$]
After k iterations of the PCA 3.1.2, the convex sum of the primal iterates (computed in step 1)

$$\hat{x}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} x_i^{j+1} \quad \textit{for} \quad i = 1,\dots,n$$

satisfies with $(\hat{\mu},\hat{\lambda}) = (u,h)^k$ (computed in step 3) the following upper bound on the primal gap:

$$\begin{aligned} \sum_{i=1}^{n} \Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \le c \sum_{i=1}^{n} D_{X_i} - \max_{(\mu,\lambda) \in Q_A \times Q_B} \Bigg\{ &-\frac{4L_c}{(k+1)^2 \sigma} d(\mu,\lambda) + \left\langle \sum_{i=1}^{n} A_i \hat{x}_i - b_A, \mu \right\rangle \\ &+ \left\langle \sum_{i=1}^{n} B_i \hat{x}_i - b_B, \lambda \right\rangle \Bigg\}. \end{aligned}$$

Proof. The proof is almost identical to [NS08, Theo. 3.4] and follows immediately from the proof of Theorem 3.2.5 (for $\Delta = 0$) given in section 3.2. □

Finally, the following convergence result for the PCA 3.1.2 can be given which slightly extends [NS08, Theo. 3.7] by additionally considering inequality constraints. (For the sake of comprehensibility, formulation is based on [NS08, Theo. 3.6 & Theo. 3.7].)

Theorem 3.1.5. [MUA14, Theo. 3.3]
Assume that $Q_A \times Q_B = \mathbb{R}^{m_A} \times \mathbb{R}_+^{m_B}$ and the prox-function $d(\mu,\lambda) = (\sigma/2)\left\|(\mu,\lambda)\right\|^2$. For the choice $c = \epsilon / \sum_{i=1}^n D_{X_i}$ in (3.4) and

$$k+1 = \left\lceil 2\sqrt{\frac{L_c}{\epsilon}} \right\rceil \tag{3.10}$$

where

$$L_c = \frac{\sum_{i=1}^n D_{X_i}}{\epsilon} \frac{\sum_{i=1}^n \left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}, \tag{3.11}$$

after k iterations of the PCA 3.1.2 the convex sum of the primal iterates (computed in step 1)

$$\hat{x}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} x_i^{j+1} \quad \text{for} \quad i = 1,\ldots,n \tag{3.12}$$

satisfies with $(\hat{\mu},\hat{\lambda}) = (u,h)^k$ (computed in step 3) the following bounds on the primal gap:

$$-\left\|(\mu,\lambda)^{opt}\right\|\left(\left\|(\mu,\lambda)^{opt}\right\| + \sqrt{\|(\mu,\lambda)^{opt}\|^2 + 2}\right)\epsilon \le \sum_{i=1}^n \Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \le \epsilon, \tag{3.13}$$

as well as the following bound on the constraint violation:

$$\left\| \begin{matrix} \sum_{i=1}^n A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^n B_i\hat{x}_i - b_B\right]^+ \end{matrix} \right\| \le \epsilon\left(\left\|(\mu,\lambda)^{opt}\right\| + \sqrt{\|(\mu,\lambda)^{opt}\|^2 + 2}\right). \tag{3.14}$$

Proof. Applying Lemma 3.1.3 and Theorem 3.1.4, the proof is almost identical to [NS08, proof of Theo. 3.7] but will be given here for the sake of completeness as we additionally consider inequality constraints in (3.1).
To obtain the upper bound on the primal gap in (3.13), consider the result of Theorem 3.1.4

$$\begin{aligned}\sum_{i=1}^n \Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \le c\sum_{i=1}^n D_{X_i} - \max_{(\mu,\lambda)\in Q_A\times Q_B} \Bigg\{ -\frac{4L_c}{(k+1)^2\sigma} d(\mu,\lambda) + \left\langle \sum_{i=1}^n A_i\hat{x}_i - b_A, \mu \right\rangle \\ + \left\langle \sum_{i=1}^n B_i\hat{x}_i - b_B, \lambda \right\rangle \Bigg\},\end{aligned}$$

where the maximization part has the solution

$$\max_{(\mu,\lambda)\in Q_A\times Q_B}\left\{-\frac{4L_c}{(k+1)^2\sigma}d(\mu,\lambda)+\left\langle\sum_{i=1}^n A_i\hat{x}_i-b_A,\mu\right\rangle+\left\langle\sum_{i=1}^n B_i\hat{x}_i-b_B,\lambda\right\rangle\right\}$$
$$=\frac{(k+1)^2}{8L_c}\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|^2.$$

It follows that

$$\sum_{i=1}^n\Phi_i(\hat{x}_i)-f_0(\hat{\mu},\hat{\lambda})\le c\sum_{i=1}^n D_{X_i}-\frac{(k+1)^2}{8L_c}\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|^2\le c\sum_{i=1}^n D_{X_i} \tag{3.15}$$

which immediately yields the right-hand side of (3.13) for the choice $c=\epsilon/\sum_{i=1}^n D_{X_i}$. From Lemma 3.1.3, the choice of c, and inequality (3.15) it follows that

$$-\|(\mu,\lambda)^{\text{opt}}\|\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|\le\epsilon-\frac{(k+1)^2}{8L_c}\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|^2,$$

yielding

$$\frac{(k+1)^2}{8L_c}\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|^2-\|(\mu,\lambda)^{\text{opt}}\|\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i-b_A\\ [\sum_{i=1}^n B_i\hat{x}_i-b_B]^+\end{matrix}\right\|-\epsilon\le 0.$$

In other words, the constraint violation is less than the largest root of

$$\frac{(k+1)^2}{8L_c}y^2-\|(\mu,\lambda)^{\text{opt}}\|\,y-\epsilon$$

which is given by the quadratic formula

$$y=\left(\|(\mu,\lambda)^{\text{opt}}\|+\sqrt{\|(\mu,\lambda)^{\text{opt}}\|^2+\frac{(k+1)^2}{2L_c}\epsilon}\right)\frac{4L_c}{(k+1)^2}.$$

With $k+1=\lceil 2\sqrt{L_c/\epsilon}\rceil$, one obtains

$$y\le\left(\|(\mu,\lambda)^{\text{opt}}\|+\sqrt{\|(\mu,\lambda)^{\text{opt}}\|^2+\frac{\left(2\sqrt{L_c/\epsilon}\right)^2}{2L_c}\epsilon}\right)\frac{4L_c}{\left(2\sqrt{L_c/\epsilon}\right)^2}$$
$$=\left(\|(\mu,\lambda)^{\text{opt}}\|+\sqrt{\|(\mu,\lambda)^{\text{opt}}\|^2+2}\right)\epsilon$$

which gives the bound on the constraint violation in (3.14) and with Lemma 3.1.3 the lower bound on the primal gap in (3.13). □

3.1.1 Minimal Lipschitz constant and scaling technique

The content of this section was essentially published in [MUA14, sec. 3.2 & 3.4] (Meinel, Ulbrich, and Albrecht) and is reproduced here in similar form.

In this section, the convergence result from Theorem 3.1.5 is improved for the following choice of prox-functions

$$d_{x_i}(x_i) = \frac{\sigma_{x_i}}{2} \|x_i\|^2 \quad \text{for} \quad i = 1, \dots, n \tag{3.16}$$

in the augmented dual function (3.4). Firstly, it is shown how the number of iterations in Theorem 3.1.5 can be reduced by minimizing the Lipschitz constant (3.11) with respect to the convexity parameter σ_{x_i} for $i = 1, \dots, n$.
Secondly, a scaling technique is presented to compensate for a large value of $\|(\mu, \lambda)^{\text{opt}}\|$ in the bounds on the primal gap given in Theorem 3.1.5.
Firstly, for the above choice of prox-functions, we obtain the upper bounds

$$D_{X_i} = \frac{\sigma_{x_i}}{2} \max_{x_i \in X_i} \|x_i\|^2$$

and it follows that the necessary number of iterations in Theorem 3.1.5 is given by

$$k = \left\lceil 2\sqrt{\frac{L_c}{\epsilon}} \right\rceil - 1,$$

where

$$\begin{aligned} L_c(\sigma_X) &= \frac{\sum_{i=1}^n D_{X_i}}{\epsilon} \frac{\sum_{i=1}^n \left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}} \\ &= \frac{\sum_{i=1}^n \sigma_{x_i} \max_{x_i \in X_i} \|x_i\|^2}{2\epsilon} \frac{\sum_{i=1}^n \left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}} \\ &= \frac{\sigma_X^T d_X}{2\epsilon} \sum_{i=1}^n \frac{v_i}{\sigma_{x_i}} \end{aligned}$$

with $v = (\left\|(A_1^T, B_1^T)^T\right\|^2, \dots, \left\|(A_n^T, B_n^T)^T\right\|^2)^T$, $d_X = (\max_{x_1 \in X_1} \|x_1\|^2, \dots, \max_{x_n \in X_n} \|x_n\|^2)^T$, and $\sigma_X = (\sigma_{x_1}, \dots, \sigma_{x_n})^T$. In other words, the minimization of $L_c(\sigma_X)$ results in a minimization of the necessary number of iterations.
Let σ_X^{opt} be a minimum of $L_c(\sigma_X)$. Then ${\sigma_X^{\text{opt}}}^T d_X = \zeta$ for some $\zeta > 0$ and it follows that $\sigma_X^{\text{opt}}/\zeta$ minimizes $L_c(\sigma_X)$ as well. In other words, it is sufficient to solve

$$\operatorname*{arg\,min}_{\sigma_X > 0} \sum_{i=1}^n \frac{v_i}{\sigma_{x_i}} \tag{3.17a}$$

$$\text{s.t. } d_X^T \sigma_X = 1. \tag{3.17b}$$

From the KKT-conditions [UU12, Theo. 16.14] it follows that an optimal solution of (3.17) satisfies

$$\left(\frac{-v_1}{\sigma_{x_1}^{\mathrm{opt}2}} + \mu d_{X_1}, \ldots, \frac{-v_n}{\sigma_{x_1}^{\mathrm{opt}2}} + \mu d_{X_n}\right)^T = 0,$$
$$d_X^T \sigma_X^{\mathrm{opt}} = 1,$$
$$\left(\sigma_{x_1}^{\mathrm{opt}}, \ldots, \sigma_{x_n}^{\mathrm{opt}}\right)^T > 0$$

which yields

$$\sigma_{x_i}^{\mathrm{opt}} = \frac{1}{\sum_{j=1}^{n} \sqrt{v_j d_{X_j}}} \sqrt{\frac{v_i}{d_{x_i}}} \quad \text{for} \quad i = 1, \ldots, n. \tag{3.18}$$

Secondly, to obtain an additional degree of freedom in the bounds of the primal gap (3.13) which allows to compensate for a large value of $\|(\mu, \lambda)^{\mathrm{opt}}\|$, consider the following scaled version of the primal problem (3.1) with scaling factor $s > 1$:

$$\min_{x(s) \in X(s)} \sum_{i=1}^{n} \Phi_i \left(\frac{x_i(s)}{s}\right) \tag{3.19a}$$
$$\text{s.t.} \ \sum_{i=1}^{n} A_i x_i(s) = b_A(s), \tag{3.19b}$$
$$\sum_{i=1}^{n} B_i x_i(s) \leq b_B(s), \tag{3.19c}$$

where $x(s) = sx$, $X(s) = sX$, $b_A(s) = sb_A$, and $b_B(s) = sb_B$ for $i = 1, \ldots, n$.
Obviously, an optimal solution x^{opt} of problem (3.1) yields an optimal solution $x^{\mathrm{opt}}(s) = sx^{\mathrm{opt}}$ of the scaled problem (3.19), and it follows that the maximum of the corresponding dual problem

$$f(\mu(s), \lambda(s)) = \min_{x(s) \in X(s)} \left\{ \sum_{i=1}^{n} \Phi_i \left(\frac{x_i(s)}{s}\right) + \left\langle \sum_{i=1}^{n} A_i x_i(s) - b_A(s), \mu(s) \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i(s) - b_B(s), \lambda(s) \right\rangle \right\}$$

is obtained at $(\mu(s), \lambda(s))^{\mathrm{opt}} = (\mu, \lambda)^{\mathrm{opt}}/s$ as

$$\begin{aligned}
&\sum_{i=1}^{n} \Phi_i \left(\frac{x_i^{\mathrm{opt}}(s)}{s}\right) + \left\langle \sum_{i=1}^{n} A_i x_i(s) - b_A(s), \frac{\mu^{\mathrm{opt}}}{s} \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i(s) - b_B(s), \frac{\lambda^{\mathrm{opt}}}{s} \right\rangle \\
&= \sum_{i=1}^{n} \Phi_i \left(\frac{s x_i^{\mathrm{opt}}}{s}\right) + \left\langle s \sum_{i=1}^{n} A_i x_i - s b_A, \frac{\mu^{\mathrm{opt}}}{s} \right\rangle + \left\langle s \sum_{i=1}^{n} B_i x_i - s b_B, \frac{\lambda^{\mathrm{opt}}}{s} \right\rangle \\
&= \sum_{i=1}^{n} \Phi_i \left(x_i^{\mathrm{opt}}\right) + \left\langle \sum_{i=1}^{n} A_i x_i^{\mathrm{opt}} - b_A, \mu^{\mathrm{opt}} \right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i^{\mathrm{opt}} - b_B, \lambda^{\mathrm{opt}} \right\rangle .
\end{aligned}$$

For the choice

$$d_{x_i}(x_i) = \frac{\sigma_{x_i}}{2} \|x_i\|^2 \quad \text{for} \quad i = 1, \dots, n, \tag{3.20}$$

we obtain $D_{X_i}(s) = s^2 D_{X_i}$ and with Theorem 3.1.5 it follows that after

$$\begin{aligned} k &= \left\lceil 2\sqrt{\frac{\sum_{i=1}^n D_{X_i}(s) \sum_{i=1}^n \|(A_i^T, B_i^T)^T\|^2}{\epsilon^2 \sigma_{x_i}}} \right\rceil - 1 \\ &= \left\lceil 2\frac{s}{\epsilon}\sqrt{\frac{\sum_{i=1}^n D_{X_i} \sum_{i=1}^n \|(A_i^T, B_i^T)^T\|^2}{\sigma_{x_i}}} \right\rceil - 1 \end{aligned}$$

iterations of the PCA 3.1.2, the following bounds on the primal gap hold:

$$-\frac{1}{s}\|(\mu,\lambda)^{\text{opt}}\| \left(\frac{1}{s}\|(\mu,\lambda)^{\text{opt}}\| + \sqrt{\frac{1}{s^2}\|(\mu,\lambda)^{\text{opt}}\|^2 + 2} \right) \epsilon \le \sum_{i=1}^n \Phi_i(\hat{x}_i) - f_0(\hat{\mu}, \hat{\lambda}) \le \epsilon, \tag{3.21}$$

as well as the following bound on the constraint violation:

$$s \left\| \begin{matrix} \sum_{i=1}^n A_i \hat{x}_i - b_A \\ [\sum_{i=1}^n B_i \hat{x}_i - b_B]^+ \end{matrix} \right\| \le \epsilon \left(\frac{1}{s}\|(\mu,\lambda)^{\text{opt}}\| + \sqrt{\frac{1}{s^2}\|(\mu,\lambda)^{\text{opt}}\|^2 + 2} \right), \tag{3.22}$$

where $\hat{x}$ is defined by (3.12) and $(\hat{\mu}, \hat{\lambda}) = (u, h)^k$ (computed in step 3). Obviously, the increase of s tightens the lower bound in (3.21) and the upper bound in (3.22) more than the decrease of ϵ does, however, the impact on the number of iterations is the same.

3.2 Distributed adaptive proximal center algorithm with event-triggering

Parts of the content of this section were essentially published in [MUA14, sec. 3.3] (Meinel, Ulbrich, and Albrecht) to establish the DPCA-EC and are used in this section to develop the DAPCA-EC which is in preparation for publication in [MU14] (Meinel and Ulbrich).

In this section, the distributed adaptive proximal center algorithm with event-triggered communication (DAPCA-EC) is presented which enhances the PCA 3.1.2 by the application of the DANA-EC 2.5.1 instead of the NA 2.1.3 to maximize the augmented dual objective function (3.4) in order to find an approximate solution to problem (3.1).
To be able to apply the DANA-EC 2.5.1 to the augmented dual problem (3.7), we assume the set $Q_A \times Q_B$ to be separable according to a given partition of the dual multipliers into s subblocks as described in section 2.2 (Assumptions 2.2.1), where each subblock $(\mu, \lambda)_l$ is controlled by an agent that is referred to as dual agent and denoted by $agent_{(\mu,\lambda)_l}$ in the following.

Moreover, a globally outdated version of the iterate $(\mu,\lambda)^k$ has to be defined such that the (primal) agents $agent_{x_1},\dots,agent_{x_n}$ need to update their primal subblocks $x_1,\dots,x_n$ only once per iteration as will become clear when the DAPCA-EC is stated.

Definition 3.2.1. (Globally outdated vector)
Denote by $(\bar{\mu},\bar{\lambda})^k = ((\bar{\mu},\bar{\lambda})^k_1,\dots,(\bar{\mu},\bar{\lambda})^k_s)$ the globally outdated vector whose subblocks are available to the dual agents in iteration $k \geq 0$ and satisfy

$$\left\|(\bar{\mu},\bar{\lambda})^k_l - (\mu,\lambda)^k_l\right\|_1 \leq \Delta_k \tag{3.23}$$

for $l = 1,\dots,s$ and a given threshold $\Delta_k \geq 0$ with $\Delta_0 = 0$.

The globally outdated vector is a special case of Definition 2.2.4, where the outdated vector $(\mu,\lambda)^{l,k} = \left((\mu,\lambda)^{l,k}_1,\dots,(\mu,\lambda)^{l,k}_s\right) \in Q_A \times Q_B \subseteq \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}_+$ (available to $agent_{(\mu,\lambda)_l}$ in iteration k) is defined by

$$(\mu,\lambda)^{l,k}_j = \begin{cases} (\bar{\mu},\bar{\lambda})^k_j & \text{if } j \in N_{IDG}(l) \cup \{l\}, \\ 0 & \text{else,} \end{cases} \tag{3.24}$$

for $l,j = 1,\dots,s$. In other words, the dual agents use the same outdated information for the computation of the subblocks $\nabla_l f_c((\mu,\lambda)^{l,k}) = \nabla_l f_c((\bar{\mu},\bar{\lambda})^k)$.
To initialize the following DAPCA-EC 3.2.2, choose $\gamma > 1$, $L_{-1} \in (0, L_c]$, and the starting point $(\bar{\mu},\bar{\lambda})^0 = (\mu,\lambda)^0$ according to (2.29) and Assumptions 2.2.1 as the minimum of a separable prox-function $d(\mu,\lambda) = \sum_{l=1}^s d_l((\mu,\lambda)_l)$ with convexity parameter $\sigma > 0$, where $d((\bar{\mu},\bar{\lambda})^0) = 0$. Moreover, set $(\bar{u},\bar{h})^{-1} = (\bar{\mu},\bar{\lambda})^0$. Finally, the DAPCA-EC can be stated as follows, where its depiction is based on our DPCA-EC [MUA14, Algo. 3.6].

Algorithm 3.2.2. (DAPCA-EC) *For $k \geq 0$ do in parallel:*
For $i = 1,\dots,n$, given the required subblocks of $(\bar{\mu},\bar{\lambda})^k$, $agent_{x_i}$

1. *computes*

$$x_i^{k+1} = \operatorname*{arg\,min}_{x_i \in X_i} \left\{ \Phi_i(x_i) + \left\langle \bar{\mu}^k, A_i x_i \right\rangle + \left\langle \bar{\lambda}^k, B_i x_i \right\rangle + c d_{x_i}(x_i) \right\}$$

and sends x_i^{k+1} to the dual agents that require it.

For $l = 1,\dots,s$, given the blocks x_i^{k+1} that are necessary for the computation of $\nabla_l f_c((\bar{\mu},\bar{\lambda})^k) = \nabla_l f_c((\mu,\lambda)^{l,k})$, $agent_{(\mu,\lambda)_l}$

2. *computes* $\nabla_l f_c((\bar{\mu},\bar{\lambda})^k) = \begin{pmatrix} \sum_{i=1}^n A_i x_i^{k+1} - b_A \\ \sum_{i=1}^n B_i x_i^{k+1} - b_B \end{pmatrix}_l$ *and sets $L_k = L_{k-1}$,*

3. *finds*

$$(u,h)_l^k = \underset{(u,h)_l \in (Q_A \times Q_B)_l}{\arg\max} \left\{ \left\langle \nabla_l f_c((\bar{\mu},\bar{\lambda})^k), (u,h)_l \right\rangle - L_k \Delta_k (\eta_l + 1) \left\| (u,h)_l - (\mu,\lambda)_l^k \right\|_1 - \frac{L_k}{2} \left\| (u,h)_l - (\mu,\lambda)_l^k \right\|^2 \right\}.$$

4. ***if*** $L_k < L_c$ ***then***

(a) *agent*$_{(\mu,\lambda)_l}$ *sends* $(u,h)_l^k$ *to the primal agents that require it if necessary:*
if $\left\| (\bar{u},\bar{h})_l^{k-1} - (u,h)_l^k \right\|_1 > \frac{L_k}{L_c} \Delta_k$ ***then***
agent$_{(\mu,\lambda)_l}$ *sets* $(\bar{u},\bar{h})_l^k = (u,h)_l^k$ *and sends* $(\bar{u},\bar{h})_l^k$.
else
agent$_{(\mu,\lambda)_l}$ *sets* $(\bar{u},\bar{h})_l^k = (\bar{u},\bar{h})_l^{k-1}$ *and signals that no data will be sent.*

For $i = 1,\ldots,n$, *given the required subblocks of* $(\bar{u},\bar{h})^k$, *agent*$_{x_i}$

(b) *computes*

$$y_i^{k+1} = \underset{x_i \in X_i}{\arg\min} \left\{ \Phi_i(x_i) + \left\langle \bar{u}^k, A_i x_i \right\rangle + \left\langle \bar{h}^k, B_i x_i \right\rangle + c d_{x_i}(x_i) \right\},$$

and sends y_i^{k+1} *to the dual agents that require it.*

For $l = 1,\ldots,s$, *given the blocks* y_i^{k+1} *that are necessary for the computation of* $\nabla_l f_c((\bar{u},\bar{h})^k) = \nabla_l f_c((u,h)^{l,k})$, *agent*$_{(\mu,\lambda)_l}$

(c) *computes* $\nabla_l f_c((\bar{u},\bar{h})^k) = \begin{pmatrix} \sum_{i=1}^n A_i y_i^{k+1} - b_A \\ \sum_{i=1}^n B_i y_i^{k+1} - b_B \end{pmatrix}_l$ *and checks with consensus*
if

$$-\frac{L_k}{2} \left\| (u,h)^k - (\mu,\lambda)^k \right\|^2 \le \sum_{l=1}^{s} \left\langle \nabla_l f((\bar{u},\bar{h})^k) - \nabla_l f((\bar{\mu},\bar{\lambda})^k), (u,h)_l^k - (\mu,\lambda)_l^k \right\rangle$$

then
continues with step 5,
else
sets $L_k = L_k \gamma$ *and goes to step 3,*

5. *finds* $(v,t)_l^k = \underset{(v,t)_l \in (Q_A \times Q_B)_l}{\arg\max} \left\{ -\frac{L_k}{\sigma} d_l((v,t)_l) + \sum_{j=0}^{k} \frac{j+1}{2} \left\langle \nabla_l f_c((\bar{\mu},\bar{\lambda})^j), (v,t)_l \right\rangle \right\},$

6. *sets* $(\mu,\lambda)_l^{k+1} = \dfrac{2}{k+3}(v,t)_l^k + \dfrac{k+1}{k+3}(u,h)_l^k,$

7. *and sends $(\mu,\lambda)_l^{k+1}$ to the primal agents that require it if necessary:*
   if $\left\|(\bar{\mu},\bar{\lambda})_l^k - (\mu,\lambda)_l^{k+1}\right\|_1 > \Delta_{k+1}$ ***then***
   $agent_{(\mu,\lambda)_l}$ sets $(\bar{\mu},\bar{\lambda})_l^{k+1} = (\mu,\lambda)_l^{k+1}$ and sends $(\bar{\mu},\bar{\lambda})_l^{k+1}$.
   else
   $agent_{(\mu,\lambda)_l}$ sets $(\bar{\mu},\bar{\lambda})_l^{k+1} = (\bar{\mu},\bar{\lambda})_l^k$ and signals that no data will be sent.

Remark 3.2.3.

1. *Step 3 of the DAPCA-EC 3.2.2 differs from step 2 of the DANA-EC 2.5.1 by containing $\eta_l + 1$ instead of η_l which follows from the definition of the globally outdated vector in (3.23) and (3.24). The convergence result for this slightly modified version of the DANA-EC is obtained exactly as in Theorem 2.5.5 with $\eta_{max} + 1$ instead of η_{max}.*

2. *For the choice $L_{-1} = L_c$, the above algorithm implements a slightly modified version of the DNA-EC 2.2.5 and in this case Algorithm 3.2.2 is referred to as DPCA-EC ([MUA14, Algo. 3.6]) in the following.*

3. *As we mentioned in [DMUH15, sec. 3], the Lipschitz constant L_c (3.6) of ∇f_c can be computed in parallel with local communication by the application of the consensus technique, provided that the number of agents in the multi-agent network is known to the agents. The same holds for the optimal convexity parameters $\sigma_{x_i}^{opt}$ (3.18) of the prox-functions that are used to smooth the dual function.*

As already stated in Remark 2.4.2, the adaptive step-size condition in step $4c)$ of the DAPCA-EC is used as well in a distributed dual decomposition method [KCD15, Algo. 4] that applies a fast gradient scheme, where, however, the condition is verified centrally. If the matrices A_i and B_i in (3.4) have a sparse structure, $agent_{x_i}$ possibly does not need all the subblocks of $(\bar{\mu},\bar{\lambda})^k$ in iteration k in order to compute his subblock x_i^{k+1}. To emphasize this, we denote in the following by $(\bar{\mu},\bar{\lambda})^{x_i,k} \in Q_A \times Q_B$ the vector whose subblocks coincide with the outdated subblocks of $(\bar{\mu},\bar{\lambda})^k$ if they are necessary to update x_i^{k+1} and whose subblocks are zero if not. It follows that the event-triggered communication is not only related to the exchange of dual iterates in step $4a)$ and step 7, but also to the exchange of primal iterates as $x_i^{k+1} = x_i^k$ if $(\mu,\lambda)^{x_i,k} = (\mu,\lambda)^{x_i,k-1}$ for $k \geq 1$, i.e., in this case x_i^{k+1} does not need to be send again to the requesting dual agents.
To be able to proof the convergence of the DAPCA-EC 3.2.2, we have to assume that the set $Q_A \times Q_B$ is bounded as done in Assumptions 2.2.9.

Firstly, we extend Lemma 2.5.4 in the following by bounding the error that occurs due to event-triggered communication with the following constant

$$P(\Delta) = 2\gamma L_c C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right) \tag{3.25}$$

that contains the update parameter $\gamma > 0$, the Lipschitz constant L_c given in (3.6), the diameter C of $Q_A \times Q_B$ defined according to (2.13), the threshold parameters $\delta \in (0,1)$ and $\beta > 0$, and finally the derivative of the function $g(\delta)$ defined in Theorem 2.5.5.

Lemma 3.2.4. [MUA14, Lem. 3.8]
The following inequality holds for $(u,h)^k$ (computed in step 3) of the DAPCA-EC 3.2.2 and $(\bar{\mu},\bar{\lambda})^k$ defined according to (3.23) for $k \geq 0$:

$$\begin{aligned}\frac{(k+1)(k+2)}{4} f_c((u,h)^k) \geq \max_{(u,h)\in Q_A\times Q_B} \Bigg\{ -\frac{\gamma L_c}{\sigma} d(u,h) + \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\bar{\mu},\bar{\lambda})^j) \\ + \left\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \right\rangle \Big) \Bigg\} - P(\Delta).\end{aligned} \tag{3.26}$$

Proof. [MUA14, proof of Lem. 3.8]
For the choice $\alpha_k = (k+1)/2$ in Lemma 2.5.4 and with Remark 2.2.11 (which holds for Lemma 2.5.4 as well), we obtain

$$\begin{aligned}\frac{(k+1)(k+2)}{4} f_c((u,h)^k) &\geq \max_{(u,h)\in Q_A\times Q_B} \Bigg\{ -\rho^k - \frac{L_k}{\sigma} d(u,h) + \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) \\ &\qquad + \left\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\mu,\lambda)^j \right\rangle \Big) \Bigg\} - \sum_{j=1}^{k} E_j \\ &\geq \max_{(u,h)\in Q_A\times Q_B} \Bigg\{ -\rho^k - \frac{\gamma L_c}{\sigma} d(u,h) + \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) \\ &\qquad + \left\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\mu,\lambda)^j \right\rangle \Big) \Bigg\} - \sum_{j=1}^{k} E_j,\end{aligned}$$

where ρ^k is defined by (2.41) combined with Remark 3.2.3, i.e.,

$$\rho_k = (\eta_{\max}+1) L_k C \sum_{j=0}^{k} \frac{j+1}{2} \beta \delta^j + \sum_{j=0}^{k} \left(L_j - L_{j-1}\right) C.$$

It can easily be shown (cf. with the proof of Theorem 2.5.5) that

$$\rho^k \leq \frac{P(\Delta)}{4} \quad \text{and} \quad \sum_{j=1}^{k} E_j \leq \frac{P(\Delta)}{4}$$

which yields

$$\frac{(k+1)(k+2)}{4} f_c((u,h)^k) \geq \max_{(u,h)\in Q_A\times Q_B} \Bigg\{ -\frac{\gamma L_c}{\sigma} d(u,h) + \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) + \Big\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\mu,\lambda)^j \Big\rangle \Big) \Bigg\} - \frac{P(\Delta)}{2}.$$

With (3.24), we have

$$\begin{aligned}
&\Big\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\mu,\lambda)^j \Big\rangle \\
&= \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j \Big\rangle + \sum_{l=1}^{s} \Big\langle \nabla_l f_c((\mu,\lambda)^{l,j}) - \nabla_l f_c((\mu,\lambda)^j), (u,h)_l - (\mu,\lambda)^j_l \Big\rangle \\
&\geq \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j \Big\rangle - L_c \sum_{l=1}^{s} \underbrace{\Big\| (\mu,\lambda)^{l,j} - (\mu,\lambda)^j \Big\|}_{\leq (\eta_l+1)\Delta_j} \Big\| (u,h)_l - (\mu,\lambda)^j_l \Big\| \\
&\geq \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j \Big\rangle - (\eta_{\max}+1)\, L_c \Delta_j \underbrace{\Big\| (u,h) - (\mu,\lambda)^j \Big\|_1}_{\leq C}
\end{aligned}$$

and accordingly

$$\Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \Big\rangle \geq \Big\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \Big\rangle - (\eta_{\max}+1)\, L_c C \Delta_j.$$

Finally, we obtain for all $(u,h) \in Q_A \times Q_B$ with the concavity of f_c that

$$\begin{aligned}
&\sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) + \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j \Big\rangle - (\eta_{\max}+1)\, L_c C \Delta_j \Big) - \frac{P(\Delta)}{2} \\
&\geq \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) + \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j \Big\rangle \Big) - \frac{3P(\Delta)}{4} \\
&= \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\mu,\lambda)^j) + \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\mu,\lambda)^j + (\bar{\mu},\bar{\lambda})^j - (\bar{\mu},\bar{\lambda})^j \Big\rangle \Big) - \frac{3P(\Delta)}{4} \\
&\geq \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\bar{\mu},\bar{\lambda})^j) + \Big\langle \nabla f_c((\mu,\lambda)^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \Big\rangle \Big) - \frac{3P(\Delta)}{4} \\
&\geq \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\bar{\mu},\bar{\lambda})^j) + \Big\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \Big\rangle - (\eta_{\max}+1)\, L_c C \Delta_j \Big) - \frac{3P(\Delta)}{4} \\
&\geq \sum_{j=0}^{k} \frac{j+1}{2} \Big(f_c((\bar{\mu},\bar{\lambda})^j) + \Big\langle \nabla f_c((\bar{\mu},\bar{\lambda})^j), (u,h) - (\bar{\mu},\bar{\lambda})^j \Big\rangle \Big) - P(\Delta).
\end{aligned}$$

□

The following lemma provides an upper bound on the primal gap and extends Theorem 3.1.4 by additionally considering the error due to event-triggered communication.

Theorem 3.2.5. [MUA14, Theo. 3.9]
After k iterations of the DAPCA-EC 3.2.2 , the convex sum of the primal iterates (computed in step 1)

$$\hat{x}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} x_i^{j+1} \quad \textit{for} \quad i = 1, \ldots, n$$

satisfies with $(\hat{\mu}, \hat{\lambda}) = (u, h)^k$ (computed in step 3) the following upper bound on the primal gap:

$$\begin{aligned}\sum_{i=1}^{n} \Phi_i(\hat{x}_i) - f_0(\hat{\mu}, \hat{\lambda}) \leq c \sum_{i=1}^{n} D_{X_i} - \max_{(\mu,\lambda) \in Q_A \times Q_B} \Bigg\{ &- \frac{4\gamma L_c}{(k+1)^2 \sigma} d(\mu, \lambda) \\ &+ \left\langle \sum_{i=1}^{n} A_i \hat{x}_i - b_A, \mu \right\rangle + \left\langle \sum_{i=1}^{n} B_i \hat{x}_i - b_B, \lambda \right\rangle \Bigg\} + \frac{4P(\Delta)}{(k+1)^2}.\end{aligned}$$

Proof. The proof is almost identical to [NS08, proof of Theo. 3.4], however, extends it by additionally considering inequality constraints in (3.1) as well as the error that occurs due to event-triggered communication, and will be given here for the sake of completeness: From inequality (3.26) in Lemma 3.2.4 it follows for any $k \geq 0$ that

$$\begin{aligned}f_c(\hat{\mu}, \hat{\lambda}) \geq \max_{(\mu,\lambda) \in Q_A \times Q_B} \Bigg\{ &- \frac{4\gamma L_c}{(k+1)^2 \sigma} d(\mu, \lambda) + \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} \Big(f_c((\bar{\mu}, \bar{\lambda})^j) \\ &+ \left\langle \nabla f_c((\bar{\mu}, \bar{\lambda})^j), (\mu, \lambda) - (\bar{\mu}, \bar{\lambda})^j \right\rangle \Big) \Bigg\} - \frac{4P(\Delta)}{(k+1)^2}, \qquad (3.27)\end{aligned}$$

where $(\bar{\mu},\bar{\lambda})^k$ is defined by (3.2.1). Moreover,

$$\sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}\left(f_c((\bar{\mu},\bar{\lambda})^j)+\left\langle\nabla f_c((\bar{\mu},\bar{\lambda})^j),(\mu,\lambda)-(\bar{\mu},\bar{\lambda})^j\right\rangle\right) \tag{3.28}$$

$$=\sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}\left(\sum_{i=1}^{n}\Phi_i(x_i^{j+1})+\left\langle\sum_{i=1}^{n}A_ix_i^{j+1}-b_A,\bar{\mu}\right\rangle+\left\langle\sum_{i=1}^{n}B_ix_i^{j+1}-b_B,\bar{\lambda}\right\rangle\right.$$
$$\left.+c\sum_{i=1}^{n}d_{x_i}(x_i^{j+1})+\left\langle\begin{pmatrix}\sum_{i=1}^{n}A_ix_i^{j+1}-b_A\\ \sum_{i=1}^{n}B_ix_i^{j+1}-b_B\end{pmatrix},(\mu,\lambda)-(\bar{\mu},\bar{\lambda})^j\right\rangle\right) \tag{3.29}$$

$$=\sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}\left(\sum_{i=1}^{n}\Phi_i(x_i^{j+1})+\left\langle\sum_{i=1}^{n}A_ix_i^{j+1}-b_A,\bar{\mu}\right\rangle+\left\langle\sum_{i=1}^{n}B_ix_i^{j+1}-b_B,\bar{\lambda}\right\rangle\right.$$
$$\left.+c\sum_{i=1}^{n}d_{x_i}(x_i^{j+1})+\left\langle\sum_{i=1}^{n}A_ix_i^{j+1}-b_A,\mu-\bar{\mu}\right\rangle+\left\langle\sum_{i=1}^{n}B_ix_i^{j+1}-b_B,\lambda-\bar{\lambda}\right\rangle\right)$$

$$=\sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}\left(\sum_{i=1}^{n}\Phi_i(x_i^{j+1})+c\sum_{i=1}^{n}d_{x_i}(x_i^{j+1})\right.$$
$$\left.+\left\langle\sum_{i=1}^{n}A_ix_i^{j+1}-b_A,\mu\right\rangle+\left\langle\sum_{i=1}^{n}B_ix_i^{j+1}-b_B,\lambda\right\rangle\right)$$

$$\geq\sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}\left(\sum_{i=1}^{n}\Phi_i(x_i^{j+1})+\left\langle\sum_{i=1}^{n}A_ix_i^{j+1}-b_A,\mu\right\rangle\right.$$
$$\left.+\left\langle\sum_{i=1}^{n}B_ix_i^{j+1}-b_B,\lambda\right\rangle\right) \tag{3.30}$$

$$\geq\sum_{i=1}^{n}\Phi_i(\hat{x}_i)+\left\langle\sum_{i=1}^{n}A_i\hat{x}_i-b_A,\mu\right\rangle+\left\langle\sum_{i=1}^{n}B_i\hat{x}_i-b_B,\lambda\right\rangle, \tag{3.31}$$

where (3.29) follows from step 1 of the DAPCA-EC 3.2.2 , inequality (3.30) follows from the nonnegativity of the prox-functions d_{x_i}, and inequality (3.31) is obtained by making use of the convexity of Φ_i and the definition of $\hat{x}_i$ for $i=1,\ldots,n$. Finally, replacing (3.31) with (3.28) in (3.27) yields

$$f_c(\hat{\mu},\hat{\lambda})-\sum_{i=1}^{n}\Phi_i(\hat{x}_i)\geq\max_{(\mu,\lambda)\in Q_A\times Q_B}\left\{-\frac{4\gamma L_c}{(k+1)^2\sigma}d(\mu,\lambda)+\left\langle\sum_{i=1}^{n}A_i\hat{x}_i-b_A,\mu\right\rangle\right.$$
$$\left.+\left\langle\sum_{i=1}^{n}B_i\hat{x}_i-b_B,\lambda\right\rangle\right\}-\frac{4P(\Delta)}{(k+1)^2}$$

and the claim follows immediately as

$$
\begin{aligned}
f_c(\hat{\mu},\hat{\lambda}) &= \min_{x\in X}\left\{\sum_{i=1}^{n}\Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \hat{\mu}\right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \hat{\lambda}\right\rangle + c\sum_{i=1}^{n} d_{x_i}(x_i)\right\} \\
&\leq \min_{x\in X}\left\{\sum_{i=1}^{n}\Phi_i(x_i) + \left\langle \sum_{i=1}^{n} A_i x_i - b_A, \hat{\mu}\right\rangle + \left\langle \sum_{i=1}^{n} B_i x_i - b_B, \hat{\lambda}\right\rangle\right\} + c\sum_{i=1}^{n} D_{X_i} \\
&= f_0(\hat{\mu},\hat{\lambda}) + c\sum_{i=1}^{n} D_{X_i}.
\end{aligned}
$$

□

We close this section with two convergence results for the DAPCA-EC 3.2.2 that both provide an efficiency estimate of the order $\mathcal{O}(1/\epsilon)$, i.e., of the same order as the efficiency estimate for the PCA given in Theorem 3.1.5.
The following theorem is based on [NS08, Theo. 3.6] and extends it by additionally considering inequality constraints and event-triggered communication.

Theorem 3.2.6. [MUA14, Theo. 3.10]
Assume that $Q_A \times Q_B = \{(\mu,\lambda) \in \mathbb{R}^{m_A} \times \mathbb{R}_+^{m_B} : \|\mu\|_{\max} \leq R,\ \|\lambda\|_{max} \leq R\}$ for some $R > 0$ such that $(\mu,\lambda)^{opt} \in M^{opt} \times \Lambda^{opt} \cap Q_A \times Q_B$ with $||(\mu,\lambda)^{opt}|| < R$. Denote by D a finite constant with $D \geq \max_{(\mu,\lambda)\in Q_A\times Q_B} d(\mu,\lambda)$. For the choice $c = \epsilon/(2\sum_{i=1}^{n} D_{X_i})$ with $\epsilon > 0$ in (3.4) and

$$
k+1 = \left\lceil 2\sqrt{\frac{L_c E(\Delta)}{\epsilon}}\right\rceil,
$$

where

$$
L_c = \frac{2\sum_{i=1}^{n} D_{X_i}}{\epsilon}\frac{\sum_{i=1}^{n}\left\|(A_i^T,B_i^T)^T\right\|^2}{\sigma_{x_i}}
$$

and

$$
E(\Delta) = \gamma\frac{2D + \sigma 4C\left((\eta_{max}+1)\,\beta g'(\delta)+2\right)}{\sigma},
$$

after k iterations of the DAPCA-EC 3.2.2, the convex sum of primal iterates (computed in step 1)

$$
\hat{x}_i = \sum_{j=0}^{k}\frac{2(j+1)}{(k+1)(k+2)}x_i^{j+1} \quad \text{for} \quad i = 1,\ldots,n
$$

satisfies with $(\hat{\mu},\hat{\lambda}) = (u,h)^k$ (computed in step 3) the following bounds on the primal gap:

$$
-\frac{\|(\mu,\lambda)^{opt}\|}{R - \|(\mu,\lambda)^{opt}\|}\epsilon \leq \sum_{i=1}^{n}\Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \leq \epsilon, \tag{3.32}
$$

as well as the following bound on the constraint violation:

$$
\left\|\begin{array}{c}\sum_{i=1}^{n} A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+\end{array}\right\| \leq \frac{\epsilon}{R - \|(\mu,\lambda)^{opt}\|}. \tag{3.33}
$$

Proof. [MUA14, proof of Theo. 3.10]
The proof follows [NS08, proof of Theo. 3.6] and extends it by additionally considering inequality constraints and event-triggered communication. If we have a look at the result of Theorem 3.2.5

$$\sum_{i=1}^{n}\Phi_i(\hat{x}_i)-f_0(\hat{\mu},\hat{\lambda})\le c\sum_{i=1}^{n}D_{X_i}-\max_{(\mu,\lambda)\in Q_A\times Q_B}\Bigg\{-\frac{4\gamma L_c}{(k+1)^2\sigma}d(\mu,\lambda)\\ +\left\langle\sum_{i=1}^{n}A_i\hat{x}_i-b_A,\mu\right\rangle+\left\langle\sum_{i=1}^{n}B_i\hat{x}_i-b_B,\lambda\right\rangle\Bigg\}+\frac{4P(\Delta)}{(k+1)^2},\tag{3.34}$$

the task is to minimize the right-hand side of the inequality with respect to c.
For the maximization part we obtain with the definition of D and
$Q_A\times Q_B=\left\{(\mu,\lambda)\in\mathbb{R}^{m_A}\times\mathbb{R}_+^{m_B}:\|\mu\|_{\max}\le R,\ \|\lambda\|_{\max}\le R\right\}$ that

$$\begin{aligned}\max_{(\mu,\lambda)\in Q_A\times Q_B}&\left\{-\frac{4\gamma L_c}{(k+1)^2\sigma}d(\mu,\lambda)+\left\langle\sum_{i=1}^{n}A_i\hat{x}_i-b_A,\mu\right\rangle+\left\langle\sum_{i=1}^{n}B_i\hat{x}_i-b_B,\lambda\right\rangle\right\}\\ &\ge-\frac{4\gamma L_cD}{(k+1)^2\sigma}+\max_{\mu\in Q_A}\left\langle\sum_{i=1}^{n}A_i\hat{x}_i-b_A,\mu\right\rangle+\max_{\lambda\in Q_B}\left\langle\sum_{i=1}^{n}B_i\hat{x}_i-b_B,\lambda\right\rangle\\ &=-\frac{4\gamma L_cD}{(k+1)^2\sigma}+R\left\|\sum_{i=1}^{n}A_i\hat{x}_i-b_A\right\|_1+R\left\|\left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\right\|_1\\ &\ge-\frac{4\gamma L_cD}{(k+1)^2\sigma}+R\left\|\begin{matrix}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{matrix}\right\|\end{aligned}$$

and for inequality (3.34) we obtain

$$\sum_{i=1}^{n}\Phi_i(\hat{x}_i)-f_0(\hat{\mu},\hat{\lambda})\le c\sum_{i=1}^{n}D_{X_i}+\frac{4\gamma L_cD}{(k+1)^2\sigma}-R\left\|\begin{matrix}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{matrix}\right\|+\frac{4P(\Delta)}{(k+1)^2}\tag{3.35}$$

$$\le c\sum_{i=1}^{n}D_{X_i}+\frac{4\gamma L_cD}{(k+1)^2\sigma}+\frac{4P(\Delta)}{(k+1)^2}.\tag{3.36}$$

With

$$L_c=\sum_{i=1}^{n}\frac{\left\|(A_i^T,B_i^T)^T\right\|^2}{c\sigma_{x_i}}\quad\text{and}\quad P(\Delta)=2\gamma L_cC\left((\eta_{\max}+1)\beta g'(\delta)+2\right),$$

we can express the right-hand side of (3.36) as a function $h(c)$ with

$$\begin{aligned}h(c)&=c\sum_{i=1}^{n}D_{x_i}+\gamma L_c\left(\frac{4D}{(k+1)^2\sigma}+\frac{8C\left((\eta_{\max}+1)\beta g'(\delta)+2\right)}{(k+1)^2}\right)\\ &=c\sum_{i=1}^{n}D_{x_i}+\frac{\gamma}{c}\left(\sum_{i=1}^{n}\frac{\left\|(A_i^T,B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{4D+\sigma 8C\left((\eta_{\max}+1)\beta g'(\delta)+2\right)}{(k+1)^2\sigma}.\end{aligned}$$

To get the minimum of h we have to solve

$$h'(c) = \sum_{i=1}^{n} D_{x_i} - \frac{\gamma}{c^2}\left(\sum_{i=1}^{n} \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{4D + \sigma 8C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right)}{(k+1)^2\sigma} = 0$$

$$\Longleftrightarrow\ c_{1,2}^{\text{opt}} = \pm\sqrt{\gamma\left(\sum_{i=1}^{n} \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{4D + \sigma 8C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right)}{(k+1)^2\sigma\sum_{i=1}^{n} D_{x_i}}}.$$

As c in (3.4) has to be positive, we choose

$$c^{\text{opt}} = \frac{1}{k+1}\sqrt{\gamma\left(\sum_{i=1}^{n} \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{4D + \sigma 8C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right)}{\sigma\sum_{i=1}^{n} D_{x_i}}}. \tag{3.37}$$

Finally, we get

$$h(c^{\text{opt}}) = \frac{2}{k+1}\sqrt{\gamma\left(\sum_{i=1}^{n} \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{\left(4D + \sigma 8C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right)\right)\sum_{i=1}^{n} D_{x_i}}{\sigma}}$$

and with

$$k+1 = \frac{2}{\epsilon}\sqrt{\gamma\left(\sum_{i=1}^{n} \frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\right)\frac{\left(4D + \sigma 8C\left((\eta_{\max}+1)\,\beta g'(\delta)+2\right)\right)\sum_{i=1}^{n} D_{x_i}}{\sigma}},$$

we obtain the right-hand side of inequality (3.32) and the value for $c = c^{\text{opt}}$, yielding

$$k+1 = 2\sqrt{\frac{L_c E(\Delta)}{\epsilon}}.$$

With inequality (3.8) and inequality (3.35), we get

$$\left(R - \left\|(\mu,\lambda)^{\text{opt}}\right\|\right)\left\|\begin{matrix}\sum_{i=1}^{n} A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+\end{matrix}\right\| \le c\sum_{i=1}^{n} D_{X_i} + \frac{4\gamma L_c D}{(k+1)^2\sigma} + \frac{4P(\Delta)}{(k+1)^2} = h(c).$$

The bound on the constraint violation (3.33) follows immediately by replacing c with c^{opt}. Finally, applying inequality (3.8) yields the lower bound on the primal gap. □

For the choice $d(\mu,\lambda) = (\sigma/2)\left\|(\mu,\lambda)\right\|^2$, the following theorem states the convergence of the DAPCA-EC 3.2.2, extending the result in Theorem 3.1.5 by considering event-triggered communication.

Theorem 3.2.7.
Assume that $d(\mu,\lambda) = (\sigma/2)\left\|(\mu,\lambda)\right\|^2$ with arbitrary $\sigma > 0$, and that the convex and compact set $Q_A \times Q_B \subseteq \mathbb{R}^{m_A} \times \mathbb{R}_+^{m_B}$ contains a $(\mu,\lambda)^{opt} \in M^{opt} \times \Lambda^{opt}$ as well as the vector $(\mu,\lambda)^+$ with

$$\mu^+ = \frac{(k+1)^2\sigma}{4L_c}\left(\sum_{i=1}^{n} A_i\hat{x}_i - b_A\right) \quad \text{and} \quad \lambda^+ = \frac{(k+1)^2\sigma}{4L_c}\left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+. \tag{3.38}$$

For the choice $c = \epsilon/(2\sum_{i=1}^n D_{X_i})$ with $\epsilon > 0$ in (3.4) and

$$k+1 = \left\lceil 2\sqrt{\frac{L_c 2R(\Delta)}{\epsilon}}\right\rceil,$$

where

$$L_c = \frac{2\sum_{i=1}^n D_{X_i}}{\epsilon}\frac{\sum_{i=1}^n \left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}$$

and

$$R(\Delta) = \gamma 2C\left((\eta_{max}+1)\,\beta g'(\delta) + 2\right),$$

after k iterations of the DAPCA-EC 3.2.2, the convex sum of primal iterates (computed in step 1)

$$\hat{x}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} x_i^{j+1} \quad \text{for} \quad i = 1,\ldots,n$$

satisfies with $(\hat{\mu},\hat{\lambda}) = (u,h)^k$ (computed in step 3) the following bounds on the primal gap:

$$-\frac{\gamma\epsilon\left\|(\mu,\lambda)^{opt}\right\|}{\sqrt{2R(\Delta)}}\left(\frac{\left\|(\mu,\lambda)^{opt}\right\|}{\sqrt{2R(\Delta)}} + \sqrt{\frac{\left\|(\mu,\lambda)^{opt}\right\|^2}{2R(\Delta)} + \frac{2}{\gamma}}\right) \le \sum_{i=1}^{n}\Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \le \epsilon, \tag{3.39}$$

as well as the following bound on the constraint violation:

$$\left\|\begin{matrix}\sum_{i=1}^n A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^n B_i\hat{x}_i - b_B\right]^+\end{matrix}\right\| \le \frac{\gamma\epsilon}{\sqrt{2R(\Delta)}}\left(\frac{\left\|(\mu,\lambda)^{opt}\right\|}{\sqrt{2R(\Delta)}} + \sqrt{\frac{\left\|(\mu,\lambda)^{opt}\right\|^2}{2R(\Delta)} + \frac{2}{\gamma}}\right). \tag{3.40}$$

Proof. The proof combines [NS08, proof of Theo. 3.6] and [NS08, proof Theo. 3.7] and extends them by additionally considering inequality constraints and event-triggered communication.

As done in the proof of Theorem 3.2.6, we minimize the right-hand side of the following inequality (shown in Theorem 3.2.5) with respect to c:

$$\sum_{i=1}^{n}\Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) \le c\sum_{i=1}^{n} D_{X_i} - \max_{(\mu,\lambda)\in Q_A\times Q_B}\left\{-\frac{4\gamma L_c}{(k+1)^2\sigma}d(\mu,\lambda) + \left\langle\sum_{i=1}^{n} A_i\hat{x}_i - b_A, \mu\right\rangle + \left\langle\sum_{i=1}^{n} B_i\hat{x}_i - b_B, \lambda\right\rangle\right\} + \frac{4P(\Delta)}{(k+1)^2}.$$

For the choice $d(\mu,\lambda) = (\sigma/2)\,\|(\mu,\lambda)\|^2$, the assumption $(\mu,\lambda)^+ \in Q_A \times Q_B$ yields that

$$\begin{aligned}
&\max_{\mu\in Q_A}\left\{-\frac{2\gamma L_c}{(k+1)^2}\|\mu\|^2 + \left\langle \sum_{i=1}^{n} A_i\hat{x}_i - b_A, \mu \right\rangle\right\} \\
&+\max_{\lambda\in Q_B}\left\{-\frac{2\gamma L_c}{(k+1)^2}\|\lambda\|^2 + \left\langle \sum_{i=1}^{n} B_i\hat{x}_i - b_B, \lambda \right\rangle\right\} \\
&=\frac{(k+1)^2}{8\gamma L_c}\left\|\sum_{i=1}^{n} A_i\hat{x}_i - b_A\right\|^2 + \frac{(k+1)^2}{8\gamma L_c}\left\|\left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+\right\|^2 \\
&\geq \frac{(k+1)^2}{8\gamma L_c}\left\|\begin{matrix}\sum_{i=1}^{n} A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+\end{matrix}\right\|^2,
\end{aligned}$$

and we obtain

$$\begin{aligned}
\sum_{i=1}^{n}\Phi_i(\hat{x}_i) - f_0(\hat{\mu},\hat{\lambda}) &\leq c\sum_{i=1}^{n} D_{X_i} - \frac{(k+1)^2}{8\gamma L_c}\left\|\begin{matrix}\sum_{i=1}^{n} A_i\hat{x}_i - b_A \\ \left[\sum_{i=1}^{n} B_i\hat{x}_i - b_B\right]^+\end{matrix}\right\|^2 + \frac{4P(\Delta)}{(k+1)^2} &\quad (3.41)\\
&\leq c\sum_{i=1}^{n} D_{X_i} + \frac{4P(\Delta)}{(k+1)^2}. &\quad (3.42)
\end{aligned}$$

With

$$L_c = \sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{c\sigma_{x_i}} \quad\text{and}\quad P(\Delta) = 2\gamma L_c C\left((\eta_{\max}+1)\,\beta g'(\delta) + 2\right) = L_c R(\Delta),$$

the right-hand side of (3.42) can be expressed as a function $h(c)$ with

$$\begin{aligned}
h(c) &= c\sum_{i=1}^{n} D_{X_i} + L_c\frac{4R(\Delta)}{(k+1)^2} \\
&= c\sum_{i=1}^{n} D_{X_i} + \frac{1}{c}\sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\frac{4R(\Delta)}{(k+1)^2}.
\end{aligned}$$

The positive minimizer c^{opt} of $h(c)$ is given by

$$\begin{aligned}
&h'(c) = \sum_{i=1}^{n} D_{X_i} - \frac{1}{c^2}\sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\frac{4R(\Delta)}{(k+1)^2} = 0 \\
\iff\ &c^{\text{opt}} = \sqrt{\sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\frac{4R(\Delta)}{(k+1)^2\sum_{i=1}^{n} D_{X_i}}},
\end{aligned}$$

yielding

$$\begin{aligned}
h(c^{\text{opt}}) &= 2\sqrt{\sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}\frac{4R(\Delta)}{(k+1)^2}\sum_{i=1}^{n} D_{X_i}} \\
&= \frac{2}{k+1}\sqrt{\sum_{i=1}^{n}\frac{\left\|(A_i^T, B_i^T)^T\right\|^2}{\sigma_{x_i}}4R(\Delta)\sum_{i=1}^{n} D_{X_i}}.
\end{aligned}$$

With

$$k+1=\frac{2}{\epsilon}\sqrt{\sum_{i=1}^{n}\frac{\left\|(A_i^T,B_i^T)^T\right\|^2}{\sigma_{x_i}}4R(\Delta)\sum_{i=1}^{n}D_{X_i}},$$

the upper bound on the primal gap in (3.39) is obtained and the value for $c=c^{\text{opt}}$. It follows that

$$k+1=2\sqrt{\frac{L_c 2R(\Delta)}{\epsilon}}.$$

For the choice of $c=c^{\text{opt}}$, inequality (3.41) can be written as

$$\sum_{i=1}^{n}\Phi_i(\hat{x}_i)-f_0(\hat{\mu},\hat{\lambda})\leq-\frac{(k+1)^2}{8\gamma L_c}\left\|\begin{array}{c}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{array}\right\|^2+\epsilon,$$

and with inequality (3.8), we obtain

$$\frac{(k+1)^2}{8\gamma L_c}\left\|\begin{array}{c}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{array}\right\|^2-\left\|(\mu,\lambda)^{\text{opt}}\right\|\left\|\begin{array}{c}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{array}\right\|-\epsilon\leq 0.$$

It follows that the constraint violation has to be smaller than the largest root of

$$\frac{(k+1)^2}{8\gamma L_c}y^2-\left\|(\mu,\lambda)^{\text{opt}}\right\|y-\epsilon,$$

yielding

$$\left\|\begin{array}{c}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{array}\right\|\leq\frac{4\gamma L_c}{(k+1)^2}\left(\left\|(\mu,\lambda)^{\text{opt}}\right\|+\sqrt{\left\|(\mu,\lambda)^{\text{opt}}\right\|^2+\frac{\epsilon(k+1)^2}{2\gamma L_c}}\right) \tag{3.43}$$

as detailed in the proof of Theorem 3.1.5. With $k+1=2\sqrt{L_c 2R(\Delta)/\epsilon}$, inequality (3.43) can be rewritten as

$$\left\|\begin{array}{c}\sum_{i=1}^{n}A_i\hat{x}_i-b_A\\ \left[\sum_{i=1}^{n}B_i\hat{x}_i-b_B\right]^+\end{array}\right\|\leq\frac{\gamma\epsilon}{2R(\Delta)}\left(\left\|(\mu,\lambda)^{\text{opt}}\right\|+\sqrt{\left\|(\mu,\lambda)^{\text{opt}}\right\|^2+\frac{4R(\Delta)}{\gamma}}\right),$$

yielding the bounds on the constraint violation (3.40) and with (3.8) the lower bound on the primal gap (3.39). □

4 Model of the AC/DC optimal power flow problem

In this chapter, the nonconvex alternating current optimal power flow (AC-OPF) problem as well as the direct current optimal power flow (DC-OPF) problem are described.
To this end, a short introduction to the structure of a power system and some important components is given in section 4.1. In section 4.2, the phasor representations of current and voltage as well as the definitions of real, reactive, and apparent power are presented for a better understanding of the derivation of the power balance equations in section 4.3. Finally, the AC/DC-OPF problems are given in sections 4.4 and 4.5.

4.1 Structure of a power system

The following introduction to the structure of a power system and some important components follows [BV00], [Blu08] and [Cra12].
Broadly speaking, a power system can be divided into three areas, namely the power generating units, where the power is generated, the loads, where the power is consumed, and the transmission/distribution network, where the power is transferred from the generating units to the loads [BV00, sec. 1.0]. In Figure 4.1 the one-line diagram of a portion of a power system is shown.

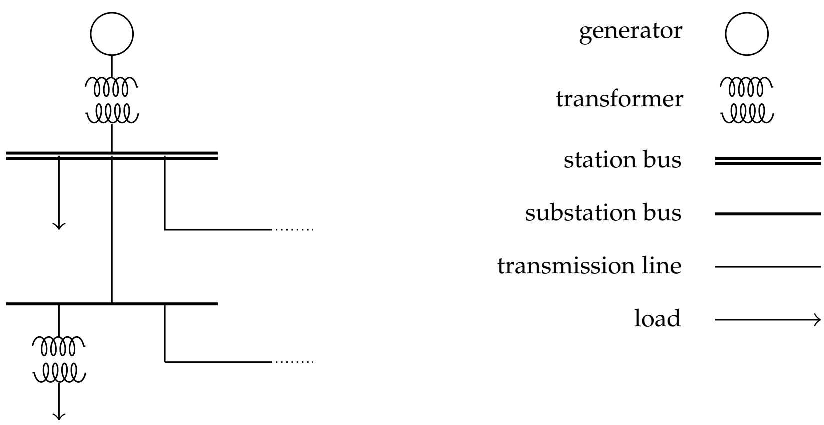

Figure 4.1: One-line diagram (follows [BV00, Fig. 1.10]).

At a generating unit electrical energy is provided by the conversion of either fossil energy stored in fuels such as coal, gas, and oil, or nuclear energy, or renewable energy such as geothermal, solar, and wind energy. Independent of its kind, the source energy is used at a power plant to run generators which consist of a cylinder, called stator, that has three single-phase windings symmetrically placed (120 degrees to each other) at its boundary. Within the stator an electromagnet is rotated, inducing an alternating voltage on each winding [Blu08, chap. 2].

Usually, a generator is capable to produce a voltage on the three single-phase windings that is between 11 and 30 kV [BV00, sec. 1.6]. If a running generator is connected to a closed three-line circuit, the generated voltage produces in each line an alternating current flow whose strength depends on the admittance of the conductor [Blu08, chap. 1-2]. Due to the symmetrical shift (by 120 degrees) of each current flow in each of the three conductors, the analysis of a three-phase system can be done with an equivalent single-phase circuit diagram which simplifies the computations. We therefore consider only single-phase components in the following [Cra12, sec. 2.3].

The advantage of using alternating current (AC), i.e., a current that is flowing back and forth instead of direct current (DC), flowing only in one direction, is that the usage of AC (current and voltage) provides the possibility of using high-voltage power lines (up to 765 kV) in the transmission network to transport electrical power from the generating units over long distances to the distribution networks that connect the loads with the system. As the power loss in a transmission line depends quadratically on the current, the generated voltage is raised with step-up transformers, placed at station buses near the power plants (Figure 4.1), such that the transmitted power is kept constant with the effect that the current is decreased and high losses are avoided [Blu08, chap. 3], [BV00, sec. 1.6].

Basically, a single-phase transformer can be described as follows [Blu08, chap. 4]:

The alternating current flowing into the transformer produces a changing magnetic field around a winding "A". This changing magnetic field induces an alternating voltage at a winding "B" which is separated from winding "A". If the winding number of "B" is smaller than winding number of "A", i.e., the turns ratio of the transformer is bigger than 1, than the voltage on winding "B" is (proportionally to the turns ratio) bigger than the voltage on winding "A". On the other side, if the turns ratio is smaller than one, the corresponding transformer is called step-down transformer.

These step-down transformers are placed at substation buses that connect the transmission network with the distribution network (0.12 to 34.5 kV), where the power is distributed to industrial or domestic consumers which are not capable of transforming a high voltage themselves. However, as seen in Figure 4.1, there is a load directly connected to the station bus which refers to an industrial consumer that is capable of consuming high voltage power or transforming the voltage himself [BV00, sec. 1.6], [Blu08, chap. 4].

Finally, the loads in a power system can be divided into three categories, namely into inductive, capacitive, and resistive loads [Blu08, chap. 1]:

Inductive loads such as motors contain windings, where magnetic fields are established when current is flowing through the windings which is therefore said to lag behind the voltage. The power consumed by an inductive load to establish a magnetic field is called reactive power Q and the power that does the motor's task is called real power P. The extent of the delay or phase shift between the applied voltage and current is denoted by the power factor angle Φ which is detailed in section 4.2.

For capacitive loads such as televisions, the current is leading the voltage and beside the real power, negative reactive power is consumed, i.e., reactive power is provided. It is favorable to balance inductive and capacitive loads in a power system which is done by the installation of phase-shifting transformers, shunt reactors, and shunt capacitors [Blu08, chap. 4] that are used to control the real and reactive power flow.

Finally, resistive loads such as lightbulbs consume only real power.

Since the nineties the electric power industry in the United States is in the process of deregulation [BV00, sec. 1.7]: Before the nineties the generation as well as the transmission and distribution of power was done by only one company in a certain part of the country. To allow competition in the power supply market this monopoly was started to be dissolved by selling the different portions of a power system to private companies.

Regarding this process, a central determination of the optimal power generation (which is done by solving the AC-OPF problem with a central entity) may not be favorable if the power generating companies want to keep certain information private such as the cost or the amount of power generation. We later show that this information, i.a., does not need to be exchanged, when the DAPCA-EC 3.2.2 is applied to solve the AC-OPF and the DC-OPF problem in a distributed manner.

4.2 Real, reactive, and apparent power

On the basis of [BV00, Cra12], we introduce the phasor representations of voltage and current as well as the definitions of real, reactive, and apparent power which help to better understand the notion of complex power and complex power flow, occurring in the power balance equations derived in section 4.3.
All physical quantities in this and the following sections are regarded as per unit values, i.e., SI units such as watt, volt-ampere reactive etc. are neglected [BV00, sec. 5.5].
Following [BV00, chap. 2], the alternating voltage and current at a bus of a power system network can be expressed as

$$\begin{aligned} v(t) &= V_{\max}\cos(\omega t + \theta_V) = \mathrm{Re}\left\{V_{\max} e^{\mathrm{j}\theta_V} e^{\mathrm{j}\omega t}\right\}, \\ i(t) &= I_{\max}\cos(\omega t + \theta_I) = \mathrm{Re}\left\{I_{\max} e^{\mathrm{j}\theta_I} e^{\mathrm{j}\omega t}\right\}, \end{aligned}$$

where $V_{\max}$ and $I_{\max}$ are the amplitudes of the oscillations, ω is their frequency and θ_V, θ_I are their phase angles. Moreover, the instantaneous power is [BV00, p. 23 - 24]

$$p(t) = v(t)\, i(t) = V_{\max} I_{\max} \cos(\omega t + \theta_V)\cos(\omega t + \theta_I) \text{ for } t \in \mathbb{R}.$$

Let $\Phi = \theta_V - \theta_I \in [-\pi/2, \pi/2]$ be the power factor angle [BV00, p. 24] which indicates if the current is leading the voltage ($\Phi < 0$) or lagging the voltage ($\Phi > 0$) or in phase with the voltage ($\Phi = 0$) [BV00, sec. 2.2]. To combine the representations in [BV00] and [Cra12], let $\tilde{t} = t - \theta_V/\omega$ for $t \in \mathbb{R}$. The instantaneous power can then be written as

$$\begin{aligned} p(\tilde{t}) = v(\tilde{t})\, i(\tilde{t}) &= V_{\max} I_{\max} \cos(\omega t)\cos(\omega t - \Phi) \\ &= V_{\max} I_{\max} \cos(\omega t)\left(\cos(\omega t)\cos(\Phi) + \sin(\omega t)\sin(\Phi)\right) \\ &= V_{\max} I_{\max} \cos(\omega t)^2 \cos(\Phi) + V_{\max} I_{\max} \cos(\omega t)\sin(\omega t)\sin(\Phi) \\ &= \frac{V_{\max} I_{\max}}{2}\cos(\Phi)\left(1 + \cos(2\omega t)\right) + \frac{V_{\max} I_{\max}}{2}\sin(2\omega t)\sin(\Phi) \qquad (4.1) \\ &= \frac{V_{\max} I_{\max}}{2}\cos(\Phi) + \frac{V_{\max} I_{\max}}{2}\cos(\Phi)\cos(2\omega t) + \frac{V_{\max} I_{\max}}{2}\sin(\Phi)\sin(2\omega t) \\ &= |V||I|\cos(\Phi) + |V||I|\cos(\Phi)\cos(2\omega t) + |V||I|\sin(\Phi)\sin(2\omega t), \qquad (4.2) \end{aligned}$$

where the identities $\cos(\omega t)^2 = (1 + \cos(2\omega t))/2$ and $\sin(\omega t)\cos(\omega t) = \sin(2\omega t)/2$ were used to obtain (4.1) according to [Cra12, p. 21]. Moreover, V and I in (4.2) are the effective phasor representations of the voltage $v(t)$ and the current $i(t)$ denoted by [BV00, p. 23]

$$V = \frac{V_{\max}}{\sqrt{2}} e^{\mathrm{j}\theta_V} \quad \text{and} \quad I = \frac{I_{\max}}{\sqrt{2}} e^{\mathrm{j}\theta_I}. \qquad (4.3)$$

The real power P is defined as the integral of $p(\tilde{t})$ over $[0, T = 2\pi/\omega]$ [BV00, p. 24 - 25]

$$\begin{aligned} P =& \frac{1}{T}\int_0^T p\left(\tilde{t}\right) d\tilde{t} = \frac{1}{T}\int_0^T p\left(t - \frac{\theta_V}{\omega}\right) dt = \frac{1}{T}\int_0^T |V||I|\cos\left(\Phi\right) dt \\ &+ \underbrace{\int_0^T |V||I|\cos\left(\Phi\right)\cos\left(2\omega t\right) dt}_{=0} + \underbrace{\frac{1}{T}\int_0^T |V||I|\sin\left(\Phi\right)\sin\left(2\omega t\right) dt}_{=0} \\ =& |V||I|\cos\left(\Phi\right). \end{aligned}$$

The reactive power Q is defined as the amplitude of the third term in (4.2) [Cra12, p. 21], i.e.,

$$Q = |V||I|\sin\left(\Phi\right).$$

The complex power S is given by [BV00, p. 26]

$$S = P + \mathrm{j}Q = |V||I|\left(\cos\left(\Phi\right) + \mathrm{j}\sin\left(\Phi\right)\right) = |V||I|e^{\mathrm{j}(\theta_V - \theta_I)} = |V|e^{\mathrm{j}\theta_V}|I|e^{-\mathrm{j}\theta_I} = VI^*,$$

where V and I are the effective phasor representations of $v(t)$ and $i(t)$ in (4.3).
Finally, the apparent power [BV00, p. 28] is defined by $|S| = \sqrt{P^2 + Q^2} = |V||I|$ which is the amplitude of the instantaneous power $p(\tilde{t})$ as [Cra12, p. 21]

$$\begin{aligned} p\left(\tilde{t}\right) &= |V||I|\cos\left(\Phi\right) + |V||I|\cos\left(\Phi\right)\cos\left(2\omega t\right) + |V||I|\sin\left(\Phi\right)\sin\left(2\omega t\right) \\ &= |V||I|\cos\left(\Phi\right) + |V||I|\cos\left(2\omega t - \Phi\right). \end{aligned}$$

For a visualization of the instantaneous power, the reactive power, and the real power see [Cra12, Fig. 2.3].

4.3 Nodal network equations

In this section, we follow [ZMS11, GS94] to derive the nodal network equations that are constraints of the AC-OPF problem.
To this end, we identify the power system network with a graph that has n_b nodes, representing the buses in the power system network, and n_l edges that represent the branches of the system, where a branch is a transmission line or a transformer that connects two buses. Moreover, let $\mathcal{N}_\mathrm{b}$ be the set of buses and $\mathcal{N}_\mathrm{g}$ be the set of generators, where each generator is identified with the bus that it is connected to, i.e., $\mathcal{N}_\mathrm{g} \subseteq \mathcal{N}_\mathrm{b}$. (To be consistent with Figure 4.1 imagine there a bus between the generator and the transformer and a bus between the load and the transformer.) Finally, let $\mathcal{N}_\mathrm{l} \subseteq \mathcal{N}_\mathrm{b} \times \mathcal{N}_\mathrm{b}$ be the set of branches of the power system network and denote by $n_\mathrm{b} = |\mathcal{N}_\mathrm{b}|$, $n_\mathrm{g} = |\mathcal{N}_\mathrm{g}|$, and $n_\mathrm{l} = |\mathcal{N}_\mathrm{l}|$ the number of buses, generators, and branches following, the notation in [ZMS11, sec. 3.1].

The model of either a transmission line or a transformer, i.e., of an element $(i, j) \in \mathcal{N}_\mathrm{l}$, is given by the combined branch model shown in Figure 4.2 which merges the representations from [ZMS11, chap. 3] and [GS94, chap. 6 & sec. 9.6] (where the details for the following definitions can be found) and shows two buses $i, j \in \mathcal{N}_\mathrm{b}$ with voltages V_i and V_j causing the currents I_{ij} and I_{ji}. Depending on the choice of the tap ratio

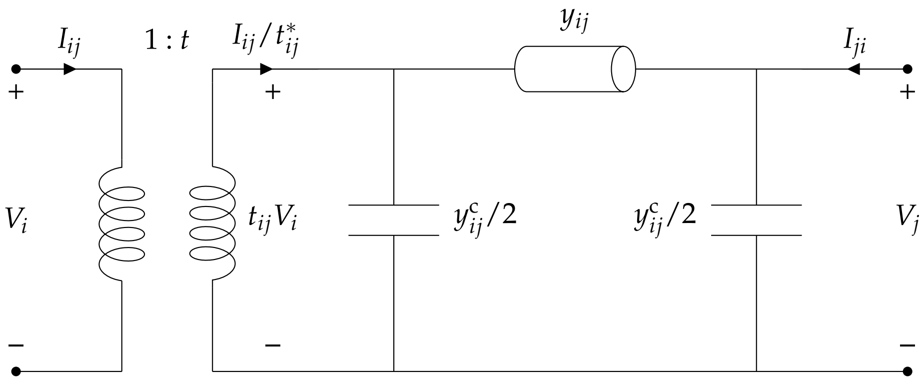

Figure 4.2: Combined branch model[1] corresponding to branch $l = (i, j) \in \mathcal{N}_\mathrm{l}$ (follows [ZMS11, Fig. 3.1] and [GS94, Fig. 6.7/9.7]).

t_{ij} and the charging capacitance $y^\mathrm{c}_{ij} = 1/\mathrm{j}x^\mathrm{c}_{ij}$, where x^c_{ij} is the capacitive reactance of the branch, the buses are connected either by a transmission line or a transformer with admittance $y_{ij} = 1/(r_{ij} + \mathrm{j}x_{ij})$, where r_{ij} is the resistance and x_{ij} is the inductive reactance of the branch. A transformer is modeled by setting the charging capacitance $y^\mathrm{c}_{ij} = 0$ and depending on the transformer type, the tap ratio t_{ij} is either real or complex. If $t_{ij} = \tau_{ij} \in \mathbb{R}_+$, where τ_{ij} is denoted as the turns ratio, the corresponding transformer changes only the magnitude of the voltage V_i and the current I_{ij} as shown in Figure 4.2 (and described in section 4.1). If $t_{ij} = \tau_{ij} \exp(\mathrm{j}\theta^\mathrm{shift}_{ij})$ with $\tau_{ij} \in \mathbb{R}_+$, the corresponding transformer changes the magnitude and shifts the phase angles of the voltage V_i and the current I_{ij} by the phase shift angle $\theta^\mathrm{shift}_{ij}$. On the other side a transmission line is modeled by setting the tap ratio $t_{ij} = 1$.

If connected by a transformer the current flows I_{ji} and I_{ij} between bus i and bus j are the following [GS94, sec. 9.6]:

$$I_{ji} = \left(V_j - t_{ij}V_i\right) y_{ij} = -y_{ij}t_{ij}V_i + y_{ij}V_j, \tag{4.4}$$

$$I_{ij} = -t^*_{ij} I_j = y_{ij}\tau^2_{ij}V_i - y_{ij}t^*_{ij}V_j. \tag{4.5}$$

[1]Created with the LaTeX package circuitikz which provides standardized circuit components.

Accordingly, the current flow equations between bus i and bus j connected by a transmission line are (cf. [ZMS11, sec. 3.2])

$$I_{ji} = \left(V_j - V_i\right) y_{ij} + V_j \frac{y^{\mathrm{c}}_{ij}}{2}, \tag{4.6}$$

$$I_{ij} = \left(V_i - V_j\right) y_{ij} + V_i \frac{y^{\mathrm{c}}_{ij}}{2}. \tag{4.7}$$

Combining (4.4) with (4.6) and (4.5) with (4.7) for a compact representation of the current flow in the combined branch model yields

$$I_{ji} = \left(y_{ij} + \frac{y^{\mathrm{c}}_{ij}}{2}\right) V_j - y_{ij} t_{ij} V_i,$$

$$I_{ij} = \left(y_{ij} + \frac{y^{\mathrm{c}}_{ij}}{2}\right) \tau^2_{ij} V_i - y_{ij} t^*_{ij} V_j,$$

which can be expressed in terms of the branch admittance matrix Y^{br}_l with $l = (i, j) \in \mathcal{N}_{\mathrm{l}}$ by [ZMS11, sec. 3.2]

$$\begin{pmatrix} I_{ij} \\ I_{ji} \end{pmatrix} = Y^{\mathrm{br}}_l \begin{pmatrix} V_i \\ V_j \end{pmatrix},$$

where

$$Y^{\mathrm{br}}_l = \begin{pmatrix} \left(y_{ij} + \frac{y^{\mathrm{c}}_{ij}}{2}\right) \tau^2_{ij} & -y_{ij} t^*_{ij} \\ -y_{ij} t_{ij} & y_{ij} + \frac{y^{\mathrm{c}}_{ij}}{2} \end{pmatrix}.$$

According to [ZMS11, sec. 3.3], a generator placed at bus i is modeled as:

$$S^{\mathrm{g}}_i = P^{\mathrm{g}}_i + \mathrm{j} Q^{\mathrm{g}}_i,$$

where P^{g}_i is the generated real and Q^{g}_i the generated reactive power.
A load at bus i is described as the complex power demand [ZMS11, sec. 3.4]

$$S^{\mathrm{d}}_i = P^{\mathrm{d}}_i + \mathrm{j} Q^{\mathrm{d}}_i,$$

where P^{d}_i is the real power demand and Q^{d}_i the reactive power demand.
Finally, a shunt element (capacitor or reactor) placed at bus i is modeled by the admittance [ZMS11, sec. 3.5]

$$y^{\mathrm{sh}}_i = g^{\mathrm{sh}}_i + \mathrm{j} b^{\mathrm{sh}}_i, \tag{4.8}$$

where g^{sh}_i is the conductance and b^{sh}_i the susceptance of the shunt element.

Incorporating the tap ratios of the transformers as well as the admittances of the shunt elements, transmission lines, and transformers into the bus admittance matrix $Y^{\text{bus}} \in \mathbb{R}^{n_\text{b} \times n_\text{b}}$, the relation between the current injection I_i at a bus i and the voltages V_j at buses j that are connected to bus i can be expressed as [ZMS11, sec. 3.6]

$$I_i = \sum_{j=1}^{n_\text{b}} Y_{ij}^{\text{bus}} V_j \ \text{ for } \ i = 1, \ldots, n_\text{b}, \tag{4.9}$$

where Y^{bus} is defined as follows:
Let $F, T \in \mathbb{R}^{n_\text{l} \times n_\text{b}}$ be the connection matrices given by [ZMS11, p. 17]

$$\begin{aligned} F_{li} &= \begin{cases} 1 & \text{if } l = (i,k) \in \mathcal{N}_\text{l} \text{ for some } k \in \mathcal{N}_\text{b}, \\ 0 & \text{else }, \end{cases} \\ T_{li} &= \begin{cases} 1 & \text{if } l = (k,i) \in \mathcal{N}_\text{l} \text{ for some } k \in \mathcal{N}_\text{b}, \\ 0 & \text{else }. \end{cases} \end{aligned} \tag{4.10}$$

Then the bus admittance matrix Y^{bus} is defined as [ZMS11, sec. 3.6]

$$\begin{aligned} Y^{\text{bus}} = & F^T \operatorname{diag}\left(\left(Y_1^{\text{br}}\right)_{11}, \ldots, \left(Y_{n_\text{l}}^{\text{br}}\right)_{11}\right) F + F^T \operatorname{diag}\left(\left(Y_1^{\text{br}}\right)_{12}, \ldots, \left(Y_{n_\text{l}}^{\text{br}}\right)_{12}\right) T \\ & + T^T \operatorname{diag}\left(\left(Y_1^{\text{br}}\right)_{21}, \ldots, \left(Y_{n_\text{l}}^{\text{br}}\right)_{21}\right) F + T^T \operatorname{diag}\left(\left(Y_1^{\text{br}}\right)_{22}, \ldots, \left(Y_{n_\text{l}}^{\text{br}}\right)_{22}\right) T \\ & + \operatorname{diag}\left(y_1^{\text{sh}}, \ldots, y_{n_\text{b}}^{\text{sh}}\right). \end{aligned}$$

Let $F(i) = \{j \in \mathcal{N}_\text{b} \ : \ (i,j) \in \mathcal{N}_\text{l}\}$ and $T(i) = \{j \in \mathcal{N}_\text{b} \ : \ (j,i) \in \mathcal{N}_\text{l}\}$, then it follows for $i, j \in \mathcal{N}_\text{b}$ that

$$\begin{aligned} Y_{ij}^{\text{bus}} = & \sum_{l=1}^{n_\text{l}} \left[F_{li} \left(Y_l^{\text{br}}\right)_{11} F_{lj} + F_{li} \left(Y_l^{\text{br}}\right)_{12} T_{lj} + T_{li} \left(Y_l^{\text{br}}\right)_{21} F_{lj} + T_{li} \left(Y_l^{\text{br}}\right)_{22} T_{lj} \right] \\ & + \operatorname{diag}\left(y_1^{\text{sh}}, \ldots, y_{n_\text{b}}^{\text{sh}}\right)_{ij} \\ = & \begin{cases} \sum_{k \in F(i)} \left(y_{ik} + \frac{y_{ik}^\text{c}}{2}\right) \tau_{ik}^2 + \sum_{k \in T(i)} \left(y_{ki} + \frac{y_{ki}^\text{c}}{2}\right) + y_i^{\text{sh}} & \text{if } i = j, \\ -y_{ij} t_{ij}^* & \text{if } (i,j) \in \mathcal{N}_\text{l}, \\ -y_{ji} t_{ji} & \text{if } (j,i) \in \mathcal{N}_\text{l}, \\ 0 & \text{else }. \end{cases} \end{aligned} \tag{4.11}$$

With (4.11) the current injection (4.9) at bus i is explicitly given by

$$\begin{aligned} I_i &= \left(\sum_{k \in F(i)} \left(y_{ik} + \frac{y_{ik}^\text{c}}{2}\right) \tau_{ik}^2 + \sum_{k \in T(i)} \left(y_{ki} + \frac{y_{ki}^\text{c}}{2}\right) + y_i^{\text{sh}} \right) V_i - \sum_{j \in F(i)} y_{ij} t_{ij}^* V_j - \sum_{j \in T(i)} y_{ji} t_{ji} V_j \\ &= \sum_{\sum_{j \in F(i)}} I_{ij} + \sum_{\sum_{j \in T(i)}} I_{ij} + y_i^{\text{sh}} V_i. \end{aligned}$$

As described in section 4.2, the complex power injection $S_i = P_i + \mathrm{j}Q_i$ at bus i is [ZMS11, sec. 3.6]

$$\begin{aligned} S_i = V_i I_i^* = V_i \sum_{j=1}^{n_\mathrm{b}} Y_{ij}^{\mathrm{bus}^*} V_j^* &= \sum_{\sum_{j\in F(i)}} V_i I_{ij}^* + \sum_{\sum_{j\in T(i)}} V_i I_{ij}^* + V_i y_i^{\mathrm{sh}^*} V_i^* \\ &= \sum_{\sum_{j\in F(i)}} S_{ij} + \sum_{\sum_{j\in T(i)}} S_{ij} + S_i^{\mathrm{sh}}, \end{aligned}$$

where $S_i^{\mathrm{sh}} = V_i y_i^{\mathrm{sh}^*} V_i^*$ denotes the complex power injected into bus i by the shunt element (capacitor or reactor) and $S_{ij} = V_i I_{ij}^*$ denotes the complex power flow from bus i to bus j connected by a branch [ZMS11, sec. 3.6]:

$$S_{ij} = V_i I_{ij}^* = \begin{cases} V_i \left(\left(y_{ij} + \frac{y_{ij}^{\mathrm{c}}}{2} \right) \tau_{ij}^2 V_i - y_{ij} t_{ij}^* V_j \right) & \text{if } (i,j) \in \mathcal{N}_\mathrm{l}, \\ V_i \left(\left(y_{ji} + \frac{y_{ji}^{\mathrm{c}}}{2} \right) V_i - y_{ji} t_{ji} V_j \right) & \text{if } (j,i) \in \mathcal{N}_\mathrm{l}. \end{cases} \tag{4.12}$$

Finally, for all $i \in \mathcal{N}_\mathrm{b}$ the AC power balance equations that relate the difference of the produced and consumed power to the power flowing in the network are [ZMS11, sec. 3.6]

$$S_i = \begin{cases} S_i^{\mathrm{g}} - S_i^{\mathrm{d}} & \text{if } i \in \mathcal{N}_\mathrm{g}, \\ -S_i^{\mathrm{d}} & \text{if } i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}. \end{cases} \tag{4.13}$$

4.4 AC optimal power flow problem

After having established the nodal network equations in section 4.3, we are now ready to state the AC-OPF problem. The AC-OPF problem actually comprises all sorts of optimization problems that arise in power system networks and depending on the objective function and the constraints it finds application in the real-time control or operational planning of a power system [Mom01, chap. 11 I]. In this work, however, we focus on the fuel cost minimization problem which is of the latter type and is applied to reduce the cost of real power generation subject to constraints such as the power balance equations, real power generating limits, and limits on the power flow at the branches [Mom01, chap. 11 II].

We follow the notation of [LL12, sec. II] and model the real power generation cost as a quadratic function

$$C_i \left(P_i^{\mathrm{g}} \right) = a_{i2} {P_i^{\mathrm{g}}}^2 + a_{i1} P_i^{\mathrm{g}} + a_{i0}, \tag{4.14}$$

where $a_{i2}, a_{i1}, a_{i0} \geq 0$ for $i \in \mathcal{N}_\mathrm{g}$. This quadratic cost model coincides with the model in the benchmark IEEE systems [Uni] that are used in this work for numerical results. Additionally, the cost of reactive power production could be considered too, but according

to [PMVDB05, sec. I], the real power cost is of peculiar interest in the competition on customers as the reactive power production is not explicitly charged.
Following [LL10, sec. II A] and [LL12, sec. II - III], the AC-OPF problem that we consider is given by

$$\min_{P^{\mathrm{g}},Q^{\mathrm{g}},V} \sum_{i\in\mathcal{N}_{\mathrm{g}}} C_i\left(P_i^{\mathrm{g}}\right) \tag{4.15a}$$

$$\text{s.t.}\quad V_i I_i^* = \left(P_i^{\mathrm{g}} - P_i^{\mathrm{d}}\right) + \mathrm{j}\left(Q_i^{\mathrm{g}} - Q_i^{\mathrm{d}}\right) \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{4.15b}$$

$$V_i I_i^* = -P_i^{\mathrm{d}} - \mathrm{j}Q_i^{\mathrm{d}} \qquad \forall i \in \mathcal{N}_{\mathrm{b}} \setminus \mathcal{N}_{\mathrm{g}}, \tag{4.15c}$$

$$|S_{ij}| \leq S_{ij}^{\max} \qquad \forall (i,j) \in \mathcal{N}_{\mathrm{l}}, \tag{4.15d}$$

$$|S_{ji}| \leq S_{ij}^{\max} \qquad \forall (i,j) \in \mathcal{N}_{\mathrm{l}}, \tag{4.15e}$$

$$P_i^{\min} \leq P_i^{\mathrm{g}} \leq P_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{4.15f}$$

$$Q_i^{\min} \leq Q_i^{\mathrm{g}} \leq Q_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{4.15g}$$

$$V_i^{\min} \leq |V_i| \leq V_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{b}}, \tag{4.15h}$$

where (4.15b) and (4.15c) are the power balance equations (4.13). The apparent power flow capacities of each branch $(i,j) \in \mathcal{N}_{\mathrm{l}}$ are observed by the constraints (4.15d) and (4.15e), where S_{ij} and S_{ji} are given in (4.12) for $(i,j) \in \mathcal{N}_{\mathrm{l}}$. Alternatively, limits on the magnitude of the real power flow or the current flow at each branch can be considered, but usually constraints (4.15d) and (4.15e) are used [ZMS11, sec. 5.1].
Finally, constraints (4.15f) and (4.15g) express the limits on the power generation, whereas (4.15h) limits the voltage magnitude at each bus, preventing from overvoltage.

4.5 DC optimal power flow problem

The AC-OPF problem is a nonlinear and nonconvex problem whose optimal solution provides voltage magnitudes and voltage phase angles as well as the optimal real and reactive power generation for a cost-efficient operation of the power system, regarding the real power cost. However, if only the amount and cost of real power production is of interest, the DC-OPF problem can solved which is a linear and convex optimization problem, where only the real power and the voltage phase angles are considered [ZMS11, sec. 5.2] [PMVDB05].
In fact, the DC-OPF problem is a linearized version of the AC-OPF problem obtained under the following assumptions.

Assumptions 4.5.1. [[ZMS11, sec. 3.7], [PMVDB05, sec. III]]

1. *The branch resistances r_{ij} can be ignored as $r_{ij} << x_{ij}$, yielding*

$$y_{ij} = \frac{1}{r_{ij} + jx_{ij}} = \frac{r_{ij}}{r_{ij}^2 + x_{ij}^2} - \frac{jx_{ij}}{r_{ij}^2 + x_{ij}^2} \approx -\frac{j}{x_{ij}}.$$

   Moreover, the charging capacitances y_{ij}^c can be ignored too.

2. *The voltage amplitudes are in the vicinity of* 1 *p.u., i.e.,*

$$V_i \approx e^{j\theta_i}.$$

3. *The voltage angles are so close together that*

$$\sin\left(\theta_i - \theta_j + \theta_{ij}^{shift}\right) \approx \theta_i - \theta_j + \theta_{ij}^{shift}.$$

With these assumptions the approximations of the real part of the power flow balance equations (4.15b), (4.15c) and real part of the power flow constraints (4.15d),(4.15e) can be derived similar to [ZMS11, sec. 3.7]: Applying Assumptions 4.5.1 to the bus admittance matrix Y^{bus} defined by (4.11) yields

$$Y_{ij}^{\text{bus}} \approx \begin{cases} -\sum_{k \in F(i)} \frac{\mathrm{j}}{x_{ik}} \tau_{ik}^2 - \sum_{k \in T(i)} \frac{\mathrm{j}}{x_{ki}} + y_i^{\text{sh}} & \text{if } i = j, \\ \frac{\mathrm{j}}{x_{ij}} t_{ij}^* & \text{if } (i,j) \in \mathcal{N}_1, \\ \frac{\mathrm{j}}{x_{ji}} t_{ji} & \text{if } (j,i) \in \mathcal{N}_1, \\ 0 & \text{else }, \end{cases}$$

and for the approximate current injection at bus i we obtain

$$I_i = \sum_{j=1}^{n_\text{b}} Y_{ij}^{\text{bus}} V_j \approx \left(-\sum_{k \in F(i)} \frac{\mathrm{j}}{x_{ik}} \tau_{ik}^2 - \sum_{k \in T(i)} \frac{\mathrm{j}}{x_{ki}} + y_i^{\text{sh}} \right) e^{\mathrm{j}\theta_i} + \sum_{j \in F(i)} \frac{\mathrm{j}}{x_{ij}} t_{ij}^* e^{\mathrm{j}\theta_j} + \sum_{j \in T(i)} \frac{\mathrm{j}}{x_{ji}} t_{ji} e^{\mathrm{j}\theta_j}.$$

It follows that the real power injection at bus i can be approximated by

$$\begin{aligned}
P_i = \mathrm{Re}\left\{S_i\right\} &\approx \mathrm{Re}\left\{ e^{\mathrm{j}\theta_i} \sum_{j=1}^{n_\mathrm{b}} Y_{ij}^{\mathrm{bus}*} e^{-\mathrm{j}\theta_j} \right\} \\
&\approx \mathrm{Re}\left\{ e^{\mathrm{j}\theta_i} \left(\sum_{k\in F(i)} \frac{\mathrm{j}}{x_{ik}} \tau_{ik}^2 + \sum_{k\in T(i)} \frac{\mathrm{j}}{x_{ki}} + y_i^{\mathrm{sh}*} \right) e^{-\mathrm{j}\theta_i} \right. \qquad (4.16) \\
&\qquad \left. - e^{\mathrm{j}\theta_i} \sum_{j\in F(i)} \frac{\mathrm{j}}{x_{ij}} t_{ij} e^{-\mathrm{j}\theta_j} - e^{\mathrm{j}\theta_i} \sum_{j\in T(i)} \frac{\mathrm{j}}{x_{ji}} t_{ji}^* e^{-\mathrm{j}\theta_j} \right\} \\
&= \mathrm{Re}\left\{ y_i^{\mathrm{sh}*} - e^{\mathrm{j}\theta_i} \sum_{j\in F(i)} \frac{\mathrm{j}}{x_{ij}} \tau_{ij} e^{\mathrm{j}\theta_{ij}^{\mathrm{shift}}} e^{-\mathrm{j}\theta_j} - e^{\mathrm{j}\theta_i} \sum_{j\in T(i)} \frac{\mathrm{j}}{x_{ji}} \tau_{ji} e^{-\mathrm{j}\theta_{ji}^{\mathrm{shift}}} e^{-\mathrm{j}\theta_j} \right\} \\
&= \mathrm{Re}\left\{ y_i^{\mathrm{sh}*} - \sum_{j\in F(i)} \frac{\mathrm{j}}{x_{ij}} \tau_{ij} e^{\mathrm{j}\left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right)} - \sum_{j\in T(i)} \frac{\mathrm{j}}{x_{ji}} \tau_{ji} e^{\mathrm{j}\left(\theta_i - \theta_{ji}^{\mathrm{shift}} - \theta_j\right)} \right\} \\
&= g_i^{\mathrm{sh}} + \sum_{j\in F(i)} \frac{\tau_{ij}}{x_{ij}} \sin\left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right) + \sum_{j\in T(i)} \frac{\tau_{ji}}{x_{ji}} \sin\left(\theta_i - \theta_{ji}^{\mathrm{shift}} - \theta_j\right) \\
&\approx g_i^{\mathrm{sh}} + \sum_{j\in F(i)} \frac{\tau_{ij}}{x_{ij}} \left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right) + \sum_{j\in T(i)} \frac{\tau_{ji}}{x_{ji}} \left(\theta_i - \theta_{ji}^{\mathrm{shift}} - \theta_j\right) \\
&= g_i^{\mathrm{sh}} + \left(\sum_{j\in F(i)} \frac{\tau_{ij}}{x_{ij}} + \sum_{j\in T(i)} \frac{\tau_{ji}}{x_{ji}} \right) \theta_i - \sum_{j\in F(i)} \frac{\tau_{ij}}{x_{ij}} \theta_j - \sum_{j\in T(i)} \frac{\tau_{ji}}{x_{ji}} \theta_j \\
&\qquad + \sum_{j\in F(i)} \frac{\tau_{ij}}{x_{ij}} \theta_{ij}^{\mathrm{shift}} - \sum_{j\in T(i)} \frac{\tau_{ji}}{x_{ji}} \theta_{ji}^{\mathrm{shift}}. \qquad (4.17)
\end{aligned}$$

The approximated real power flows $P_{ij} = \mathrm{Re}\{S_{ij}\}$ and $P_{ji} = \mathrm{Re}\{S_{ji}\}$ between bus i and bus j, that are connected by branch $(i,j) \in \mathcal{N}_1$, are accordingly [ZMS11, sec. 3.7]

$$\begin{aligned}
P_{ij} = \mathrm{Re}\left\{S_{ij}\right\} = \mathrm{Re}\left\{V_i I_{ij}^*\right\} &= \mathrm{Re}\left\{ V_i \left(\left(y_{ij} + \frac{y_{ij}^{\mathrm{c}}}{2} \right) \tau_{ij}^2 V_i - y_{ij} t_{ij}^* V_j \right)^* \right\} \\
&\approx \mathrm{Re}\left\{ e^{\mathrm{j}\theta_i} \left(-\frac{\mathrm{j}}{x_{ij}} \tau_{ij}^2 e^{\mathrm{j}\theta_i} + \frac{\mathrm{j}}{x_{ij}} t_{ij}^* e^{\mathrm{j}\theta_j} \right)^* \right\} = \mathrm{Re}\left\{ e^{\mathrm{j}\theta_i} \left(\frac{\mathrm{j}}{x_{ij}} \tau_{ij}^2 e^{-\mathrm{j}\theta_i} - \frac{\mathrm{j}}{x_{ij}} t_{ij} e^{-\mathrm{j}\theta_j} \right) \right\} \\
&= \mathrm{Re}\left\{ -e^{\mathrm{j}\theta_i} \frac{\mathrm{j}}{x_{ij}} \tau_{ij} e^{\mathrm{j}\theta_{ij}^{\mathrm{shift}}} e^{-\mathrm{j}\theta_j} \right\} = \mathrm{Re}\left\{ -\frac{\mathrm{j}}{x_{ij}} \tau_{ij} e^{\mathrm{j}\left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right)} \right\} \\
&= \frac{\tau_{ij}}{x_{ij}} \sin\left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right) \approx \frac{\tau_{ij}}{x_{ij}} \left(\theta_i + \theta_{ij}^{\mathrm{shift}} - \theta_j\right) \qquad (4.18)
\end{aligned}$$

and

$$\begin{aligned}
P_{ji} &= \mathrm{Re}\left\{S_{ji}\right\} = \mathrm{Re}\left\{V_j I_{ji}^*\right\} = \mathrm{Re}\left\{V_j\left(\left(\left(y_{ij}+\frac{y_{ij}^{\mathrm{c}}}{2}\right)V_j - y_{ij}t_{ij}V_i\right)^*\right)\right\} \\
&\approx \mathrm{Re}\left\{e^{\mathrm{j}\theta_j}\left(-\frac{\mathrm{j}}{x_{ij}}e^{\mathrm{j}\theta_j}+\frac{\mathrm{j}}{x_{ij}}t_{ij}e^{\mathrm{j}\theta_i}\right)^*\right\} = \mathrm{Re}\left\{e^{\mathrm{j}\theta_j}\left(\frac{\mathrm{j}}{x_{ij}}e^{-\mathrm{j}\theta_j}-\frac{\mathrm{j}}{x_{ij}}t_{ij}^*e^{-\mathrm{j}\theta_i}\right)\right\} \\
&= \mathrm{Re}\left\{-e^{\mathrm{j}\theta_j}\frac{\mathrm{j}}{x_{ij}}\tau_{ij}e^{-\mathrm{j}\theta_{ij}^{\mathrm{shift}}}e^{-\mathrm{j}\theta_i}\right\} = \mathrm{Re}\left\{-\frac{\mathrm{j}}{x_{ij}}\tau_{ij}e^{\mathrm{j}\left(\theta_j-\theta_{ij}^{\mathrm{shift}}-\theta_i\right)}\right\} \\
&= \frac{\tau_{ij}}{x_{ij}}\sin\left(\theta_j-\theta_{ij}^{\mathrm{shift}}-\theta_i\right) \approx \frac{\tau_{ij}}{x_{ij}}\left(\theta_j-\theta_{ij}^{\mathrm{shift}}-\theta_i\right).
\end{aligned} \tag{4.19}$$

To write the approximated real power flow and the approximated real power flow equations in a compact form, let $I^{\mathrm{inc}} \in \mathbb{R}^{n_1\times n_\mathrm{b}}$ denote the network incidence matrix defined by [GS94, sec. 7.5]

$$I^{\mathrm{inc}} = F - T,$$

where F and T are the connection matrices (4.10). Moreover, we define the weighted network incidence matrix similar to [WL10, sec. II] by

$$W^{\mathrm{inc}} = \mathrm{diag}\left(\frac{\tau_1}{x_1},\dots,\frac{\tau_{n_1}}{x_{n_1}}\right)I^{\mathrm{inc}},$$

where $\tau_l/x_l = \tau_{ij}/x_{ij}$ for $l=(i,j)\in\mathcal{N}_1$. Defining [ZMS11, sec. 3.7]

$$B^{\mathrm{bus}} = {I^{\mathrm{inc}}}^T W^{\mathrm{inc}} = {I^{\mathrm{inc}}}^T \mathrm{diag}\left(\frac{\tau_1}{x_1},\dots,\frac{\tau_{n_1}}{x_{n_1}}\right)I^{\mathrm{inc}},$$

it follows for $i,j\in\mathcal{N}_\mathrm{b}$ that

$$B_{ij}^{\mathrm{bus}} = \sum_{l=1}^{n_1} I_{li}^{\mathrm{inc}}\frac{\tau_l}{x_l}I_{lj}^{\mathrm{inc}} = \begin{cases} \sum_{k\in F(i)}\frac{\tau_{ik}}{x_{ik}}+\sum_{k\in T(i)}\frac{\tau_{ki}}{x_{ki}} & \text{if } i=j, \\ -\frac{\tau_{ij}}{x_{ij}} & \text{if } (i,j)\in\mathcal{N}_1, \\ -\frac{\tau_{ji}}{x_{ji}} & \text{if } (j,i)\in\mathcal{N}_1, \\ 0 & \text{else}. \end{cases}$$

The approximated real power injection at bus i can be expressed in a compact manner as [ZMS11, sec. 3.7]

$$(4.17) = \sum_{j=1}^{n_\mathrm{b}} B_{ij}^{\mathrm{bus}}\theta_j + \sum_{l=1}^{n_1} W_{li}^{\mathrm{inc}}\theta_l^{\mathrm{shift}} + g_i^{\mathrm{sh}} = \left(B^{\mathrm{bus}}\theta\right)_i + \left({W^{\mathrm{inc}}}^T\theta^{\mathrm{shift}}\right)_i + g_i^{\mathrm{sh}},$$

where $\theta\in\mathbb{R}^{n_\mathrm{b}}$ is the vector containing the voltage phase angles and $\theta^{\mathrm{shift}}\in\mathbb{R}^{n_1}$ is the vector of phase shift angles with $\theta_l^{\mathrm{shift}}=\theta_{ij}^{\mathrm{shift}}$ for $l=(i,j)\in\mathcal{N}_1$. For all $i\in\mathcal{N}_\mathrm{b}$, the approximated real power balance equations are then given by [ZMS11, sec. 3.7]

$$\left(B^{\mathrm{bus}}\theta\right)_i + \left({W^{\mathrm{inc}}}^T\theta^{\mathrm{shift}}\right)_i + g_i^{\mathrm{sh}} = \begin{cases} P_i^{\mathrm{g}} - P_i^{\mathrm{d}} & \text{if } i\in\mathcal{N}_\mathrm{g}, \\ -P_i^{\mathrm{d}} & \text{if } i\in\mathcal{N}_\mathrm{b}\setminus\mathcal{N}_\mathrm{g}. \end{cases}$$

Accordingly, the approximated real power flows from bus i and from bus j at branch $l = (i, j) \in \mathcal{N}_\mathrm{l}$ are compactly written as [ZMS11, sec. 3.7]

$$(4.18) = \left(W^{\mathrm{inc}}\theta\right)_l + \theta_{ij}^{\mathrm{shift}} \quad \text{and} \quad (4.19) = -\left(W^{\mathrm{inc}}\theta\right)_l - \theta_{ij}^{\mathrm{shift}}.$$

In sum, neglecting the reactive power and applying Assumptions 4.5.1 to the AC-OPF problem (4.15) yields the DC-OPF problem (cf. [ZMS11, sec. 5.2])

$$\min_{P^\mathrm{g},\theta} \sum_{i\in\mathcal{N}_\mathrm{g}} C_i\left(P_i^\mathrm{g}\right) \tag{4.20a}$$

$$\text{s.t.} \quad \left(B^{\mathrm{bus}}\theta\right)_i + \left({W^{\mathrm{inc}}}^T\theta^{\mathrm{shift}}\right)_i + g_i^{\mathrm{sh}} = P_i^\mathrm{g} - P_i^\mathrm{d} \qquad \forall i \in \mathcal{N}_\mathrm{g}, \tag{4.20b}$$

$$\left(B^{\mathrm{bus}}\theta\right)_i + \left({W^{\mathrm{inc}}}^T\theta^{\mathrm{shift}}\right)_i + g_i^{\mathrm{sh}} = -P_i^\mathrm{d} \qquad \forall i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}, \tag{4.20c}$$

$$\left|\left(W^{\mathrm{inc}}\theta\right)_l + \theta_{ij}^{\mathrm{shift}}\right| \leq S_{ij}^{\max} \qquad \forall l = (i,j) \in \mathcal{N}_\mathrm{l}, \tag{4.20d}$$

$$\left|-\left(W^{\mathrm{inc}}\theta\right)_l - \theta_{ij}^{\mathrm{shift}}\right| \leq S_{ij}^{\max} \qquad \forall l = (i,j) \in \mathcal{N}_\mathrm{l}, \tag{4.20e}$$

$$P_i^{\min} \leq P_i^\mathrm{g} \leq P_i^{\max} \qquad \forall i \in \mathcal{N}_\mathrm{g}, \tag{4.20f}$$

$$\theta_i^{\min} \leq \theta_i \leq \theta_i^{\max} \qquad \forall i \in \mathcal{N}_\mathrm{b}, \tag{4.20g}$$

where the constraints are the approximated versions of the constraints of the AC-OPF problem (4.15) in the same order. For a compact representation of (4.20) note that constraint (4.20e) is redundant and let $I^\mathrm{g} \in \mathbb{R}^{n_\mathrm{b}\times n_\mathrm{g}}$ be the matrix defined by [ZMS11, sec. 3.3]

$$I_{ij}^\mathrm{g} = \begin{cases} 1 & \text{if generator } j \text{ is connected to bus } i, \\ 0 & \text{else .} \end{cases}$$

Moreover, define $\widetilde{P^\mathrm{d}} = P^\mathrm{d} + {W^{\mathrm{inc}}}^T\theta^{\mathrm{shift}} + g^{\mathrm{sh}}$, $F^{\max} = S^{\max} - \theta^{\mathrm{shift}}$, and $F^{\min} = -S^{\max} - \theta^{\mathrm{shift}}$, where $S^{\max} \in \mathbb{R}^{n_\mathrm{l}}$ is the vector containing $S_l^{\max}$ for $l = (i,j) \in \mathcal{N}_\mathrm{l}$.
Then the DC-OPF problem (4.20) can be written as (cf. [ZMS11, sec. 5.2])

$$\min_{P^\mathrm{g},\theta} \sum_{i\in\mathcal{N}_\mathrm{g}} C_i\left(P_i^\mathrm{g}\right) \tag{4.21a}$$

$$\text{s.t.} \quad B^{\mathrm{bus}}\theta = I^\mathrm{g}P^\mathrm{g} - \widetilde{P^\mathrm{d}}, \tag{4.21b}$$

$$F^{\min} \leq W^{\mathrm{inc}}\theta \leq F^{\max}, \tag{4.21c}$$

$$P_i^{\min} \leq P_i^\mathrm{g} \leq P_i^{\max} \qquad \forall i \in \mathcal{N}_\mathrm{g}, \tag{4.21d}$$

$$\theta_i^{\min} \leq \theta_i \leq \theta_i^{\max} \qquad \forall i \in \mathcal{N}_\mathrm{b}, \tag{4.21e}$$

which is similar to the representation in [WL10, sec. 2] that we used in [MUA14, sec. 4], however, a more exact model is considered here.

5 Distributedly solving the AC/DC optimal power flow problem

In this chapter, the DC-OPF problem (4.21) and the dual of the AC-OPF problem (4.15) are decomposed by dual decomposition to be able to apply the DAPCA-EC 3.2.2.
In section 5.1, it is shown how the DAPCA-EC can be applied to maximize the augmented dual of the DC-OPF problem and in section 5.3, the application of the DAPCA-EC to solve the AC-OPF problem is stated.
Moreover, for both applications of the DAPCA-EC the communication topology is discussed, confirming that the communication is local with respect to the power system network topology for the DC-OPF problem, and the same holds for the AC-OPF problem if the network representing the power system is chordal.

5.1 Application of the DAPCA-EC to the DC-OPF problem

Parts of the content of this section were essentially published in [MUA14, sec. 4] (Meinel, Ulbrich, and Albrecht) for the application of the DPCA-EC to the DC-OPF problem and are used in this section for the application of the DAPCA-EC to the DC-OPF problem.

To be able to apply the DAPCA-EC 3.2.2 to solve the DC-OPF problem

$$\min_{P^{\mathrm{g}},\theta} \sum_{i\in\mathcal{N}_{\mathrm{g}}} C_i\left(P_i^{\mathrm{g}}\right) \tag{5.1a}$$

$$\text{s.t. } B^{\mathrm{bus}}\theta = I^{\mathrm{g}}P^{\mathrm{g}} - \widetilde{P^{\mathrm{d}}}, \tag{5.1b}$$

$$F^{\min} \leq W^{\mathrm{inc}}\theta \leq F^{\max}, \tag{5.1c}$$

$$P_i^{\min} \leq P_i^{\mathrm{g}} \leq P_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{5.1d}$$

$$\theta_i^{\min} \leq \theta_i \leq \theta_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{b}}. \tag{5.1e}$$

in parallel and with local communication, the problem needs to be dually decomposed after defining the compact sets

$$\mathcal{P}_i = \left[P_i^{\min}, P_i^{\max}\right] \text{ for } i \in \mathcal{N}_{\mathrm{g}} \text{ and } \Theta_i = \left[\theta_i^{\min}, \theta_i^{\max}\right] \text{ for } i \in \mathcal{N}_{\mathrm{g}}$$

which allow the application of the convergence results derived in the previous sections.

As described in section 3.1, the Lagrangian of (5.1) with respect to the constraints (5.1b) and (5.1c) is given by

$$\mathcal{L}\left(P^{\mathrm{g}},\theta,\mu,\lambda\right) = \sum_{i\in\mathcal{N}_{\mathrm{g}}} C_i\left(P_i^{\mathrm{g}}\right) + \sum_{l=1}^{n_{\mathrm{l}}} \lambda_l \left(\left(W^{\mathrm{inc}}\theta\right)_l - F_l^{\max}\right) + \sum_{l=1}^{n_{\mathrm{l}}} \lambda_{l+n_{\mathrm{l}}} \left(\left(-W^{\mathrm{inc}}\theta\right)_l + F_l^{\min}\right)$$
$$+ \sum_{i=1}^{n_{\mathrm{b}}} \mu_i \left(\left(I^{\mathrm{g}}P^{\mathrm{g}}\right)_i - \left(B^{\mathrm{bus}}\theta\right)_i - \widetilde{P_i^{\mathrm{d}}}\right). \tag{5.2}$$

Smoothing the Lagrangian by the prox-functions $d_i(x_i) = (\sigma_i/2)x_i^2$ with $\sigma_i > 0$ for $i = 1,\ldots,n_{\mathrm{g}} + n_{\mathrm{b}}$ yields the following dual augmented function:

$$f_c\left(\mu,\lambda\right) = \min_{P_i^{\mathrm{g}}\in\mathcal{P}_i,\ \theta_i\in\Theta_i} \left\{ \sum_{i\in\mathcal{N}_{\mathrm{g}}} C_i\left(P_i^{\mathrm{g}}\right) + \sum_{l=1}^{n_{\mathrm{l}}} \lambda_l \left(\left(W^{\mathrm{inc}}\theta\right)_l - F_l^{\max}\right) \right. \tag{5.3}$$
$$+ \sum_{l=1}^{n_{\mathrm{l}}} \lambda_{l+n_{\mathrm{l}}} \left(\left(-W^{\mathrm{inc}}\theta\right)_l + F_l^{\min}\right)$$
$$\left. + \sum_{i=1}^{n_{\mathrm{b}}} \mu_i \left(\left(I^{\mathrm{g}}P^{\mathrm{g}}\right)_i - \left(B^{\mathrm{bus}}\theta\right)_i - \widetilde{P_i^{\mathrm{d}}}\right) + \sum_{i\in\mathcal{N}_{\mathrm{g}}} \frac{c\sigma_i}{2} {P_i^{\mathrm{g}}}^2 + \sum_{i=1}^{n_{\mathrm{b}}} \frac{c\sigma_{i+n_{\mathrm{g}}}}{2} \theta_i^2 \right\}$$
$$= \sum_{i\in\mathcal{N}_{\mathrm{g}}} \min_{P_i^{\mathrm{g}}\in\mathcal{P}_i} \left\{ C_i\left(P_i^{\mathrm{g}}\right) + \mu_i P_i^{\mathrm{g}} + \frac{c\sigma_i}{2} {P_i^{\mathrm{g}}}^2 \right\}$$
$$+ \sum_{i=1}^{n_{\mathrm{b}}} \min_{\theta_i\in\Theta_i} \left\{ \left(\sum_{l\in L(i)} \left(\lambda_l - \lambda_{l+n_{\mathrm{l}}}\right) W_{li}^{\mathrm{inc}} - \sum_{j\in N(i)\cup\{i\}} \mu_j B_{ji}^{\mathrm{bus}} \right) \theta_i + \frac{c\sigma_{i+n_{\mathrm{g}}}}{2} \theta_i^2 \right\}$$
$$+ \sum_{l=1}^{n_{\mathrm{l}}} \left(\lambda_{l+n_{\mathrm{l}}} F_l^{\min} - \lambda_l F_l^{\max}\right) - \sum_{i=1}^{n_{\mathrm{b}}} \mu_i \widetilde{P_i^{\mathrm{d}}}, \tag{5.4}$$

where for $i \in \mathcal{N}_{\mathrm{b}}$ the set $N(i) = \{j \in \mathcal{N}_{\mathrm{b}} \ : \ (i,j) \in \mathcal{N}_{\mathrm{l}} \vee (j,i) \in \mathcal{N}_{\mathrm{l}}\}$ denotes the set of buses that are connected to bus i by a branch, and $L(i) = \{l \in \mathcal{N}_{\mathrm{l}} \ : \ l = (i,j) \vee l = (j,i) \text{ for some } j \in \mathcal{N}_{\mathrm{b}}\}$ denotes the set of branches l that connect bus i with the power system network. Both sets are similarly defined as in [WL10, sec. II] and indicate the suitable structure of the decomposed DC-OPF problem with respect to local communication exchange for the determination of the primal variables.

According to Theorem 3.1.1, the partial derivatives of f_c are given by

$$\frac{\partial f_c(\mu,\lambda)}{\partial \mu_i} = (I^{\mathrm{g}} P^{\mathrm{g}}(\mu,\lambda))_i - \left(B^{\mathrm{bus}}\theta(\mu,\lambda)\right)_i - \widetilde{P_i^{\mathrm{d}}}$$

$$= \begin{cases} P_i^{\mathrm{g}} - \sum_{j\in N(i)\cup\{i\}} B_{ij}^{\mathrm{bus}}\theta_j(\mu,\lambda) - \widetilde{P_i^{\mathrm{d}}} & \text{if } i \in \mathcal{N}_{\mathrm{g}}, \\ -\sum_{j\in N(i)\cup\{i\}} B_{ij}^{\mathrm{bus}}\theta_j(\mu,\lambda) - \widetilde{P_i^{\mathrm{d}}} & \text{if } i \in \mathcal{N}_{\mathrm{b}} \setminus \mathcal{N}_{\mathrm{g}}, \end{cases} \tag{5.5}$$

$$\frac{\partial f_c(\mu,\lambda)}{\partial \lambda_l} = \left(W^{\mathrm{inc}}\theta(\mu,\lambda)\right)_l - F_l^{\max}$$

$$= \sum_{i\in G(l)} W_{li}^{\mathrm{inc}}\theta_i(\mu,\lambda) - F_l^{\max}, \tag{5.6}$$

$$\frac{\partial f_c(\mu,\lambda)}{\partial \lambda_{l+n_{\mathrm{l}}}} = -\left(W^{\mathrm{inc}}\theta_i(\mu,\lambda)\right) + F_l^{\min}$$

$$= \sum_{i\in G(l)} -W_{li}^{\mathrm{inc}}\theta_i(\mu,\lambda) + F_l^{\min}, \tag{5.7}$$

where the set $G(l) = \{i,j \;:\; l = (i,j)\}$ contains the indices of buses that are connected by branch l for $l \in \mathcal{N}_{\mathrm{l}}$. Using the sets $N(i)$ and $G(l)$ in the above representation, where $G(l)$ is similarly defined as in [LL99, WL09a] (related to network utility maximization), indicates the suitable structure of the partial derivatives of the augmented dual function with respect to a local communication exchange.

Accordingly, with Theorem 3.1.1 the Lipschitz constant L_c of ∇f_c can be determined after detecting the coefficient matrices of the primal variables P_i^{g} and θ_i in the constraints (5.1b) and (5.1c) which are

$$\left({B_i^{\mathrm{bus}}}^T, {W_i^{\mathrm{inc}}}^T, -{W_i^{\mathrm{inc}}}^T\right)^T \in \mathbb{R}^{n_{\mathrm{b}}+2n_{\mathrm{l}}} \text{ for } \theta_i \text{ and } -\left(e_i^T, 0^T, 0^T\right)^T \in \mathbb{R}^{n_{\mathrm{b}}+2n_{\mathrm{l}}} \text{ for } P_i^{\mathrm{g}},$$

where $e_i \in \mathbb{R}^{n_{\mathrm{b}}}$ denotes the unit vector and B_i^{bus} and W_i^{inc} are the ith columns of the corresponding matrices. With Theorem 3.1.1 it follows immediately that the Lipschitz constant is given by

$$L_c = \sum_{i=1}^{n_{\mathrm{b}}} \frac{\left\| \left({B_i^{\mathrm{bus}}}^T, {W_i^{\mathrm{inc}}}^T, -{W_i^{\mathrm{inc}}}^T\right)^T \right\|^2}{c\sigma_{i+n_{\mathrm{g}}}} + \sum_{i\in\mathcal{N}_{\mathrm{g}}} \frac{\left\| -\left(e_i^T, 0^T, 0^T\right)^T \right\|^2}{c\sigma_i}$$

$$= \sum_{i=1}^{n_{\mathrm{b}}} \frac{\sum_{j\in N(i)\cup\{i\}} {B_{ji}^{\mathrm{bus}}}^2 + 2\sum_{l\in L(i)} {W_{li}^{\mathrm{inc}}}^2}{c\sigma_{i+n_{\mathrm{g}}}} + \sum_{i\in\mathcal{N}_{\mathrm{g}}} \frac{1}{c\sigma_i}. \tag{5.8}$$

The Lagrangian (5.2) of the DC-OPF problem does not need to be smoothen with respect to the primal variables P_i^{g} for $i \in \mathcal{N}_{\mathrm{g}}$ to obtain a continuously differentiable dual function if the leading coefficients a_{i2} of the quadratic cost functions

$$C_i\left(P_i^{\mathrm{g}}\right) = a_{i2} {P_i^{\mathrm{g}}}^2 + a_{i1} P_i^{\mathrm{g}} + a_{i0}$$

are positive. In this case it holds for all $x, y \in \mathbb{R}$ that

$$\left(\nabla C_i(x) - \nabla C_i(y)\right)^T (x-y) = 2a_{i2}(x-y)^2,$$

showing the strongly convexity of C_i with convexity parameter $2a_{i2}$ according to (2.5). Defining $\widetilde{C}_i(P_i^{\mathrm{g}}) = a_{i1}P_i^{\mathrm{g}} + a_{i0}$ it can easily be seen that the Lipschitz constant of the gradient of the augmented dual function

$$\begin{aligned} f_c(\mu,\lambda) = & \sum_{i\in\mathcal{N}_{\mathrm{g}}} \min_{P_i^{\mathrm{g}} \in \mathcal{P}_i} \left\{ \widetilde{C}_i\left(P_i^{\mathrm{g}}\right) + \mu_i P_i^{\mathrm{g}} + \frac{2a_{i2}}{2} {P_i^{\mathrm{g}}}^2 \right\} \\ & + \sum_{i=1}^{n_{\mathrm{b}}} \min_{\theta_i \in \Theta_i} \left\{ \left(\sum_{l\in L(i)} \left(\lambda_l - \lambda_{l+n_{\mathrm{l}}}\right) W_{li}^{\mathrm{inc}} - \sum_{j \in N(i)\cup\{i\}} \mu_j B_{ji}^{\mathrm{bus}} \right) \theta_i + \frac{c\sigma_{i+n_{\mathrm{g}}}}{2} \theta_i^2 \right\} \\ & + \sum_{l=1}^{n_{\mathrm{l}}} \left(\lambda_{l+n_{\mathrm{l}}} F_l^{\min} - \lambda_l F_l^{\max} \right) - \sum_{i=1}^{n_{\mathrm{b}}} \mu_i \widetilde{P_i^{\mathrm{d}}}, \end{aligned}$$

whose partial derivatives are given by (5.5), (5.6) and (5.7) as well, is obtained by simply exchanging $c\sigma_i$ with $2a_{i2}$ in (5.8), yielding

$$L_c = \sum_{i=1}^{n_{\mathrm{b}}} \frac{\sum_{j\in N(i)\cup\{i\}} {B_{ji}^{\mathrm{bus}}}^2 + 2\sum_{l\in L(i)} {W_{li}^{\mathrm{inc}}}^2}{c\sigma_{i+n_{\mathrm{g}}}} + \sum_{i\in\mathcal{N}_{\mathrm{g}}} \frac{1}{2a_{i2}}. \tag{5.9}$$

Consider the following multi-agent network whose topology shall coincide with the power system network: For $i \in \mathcal{N}_{\mathrm{b}}$, $agent_i$ is responsible for updating the primal variables θ_i and P_i^{g} (if $i \in \mathcal{N}_{\mathrm{g}}$) and is placed at bus i. For $i \in \mathcal{N}_{\mathrm{b}}$, $agent_{\mu_i}$ controls the variable μ_i and is identified with $agent_i$. For $l = (i,j) \in \mathcal{N}_{\mathrm{l}}$, $agent_{\lambda_l}$ controls the variable λ_l and is identified with $agent_i$ as well. Accordingly, $agent_{\lambda_{l+n_{\mathrm{l}}}}$ controls the variable $\lambda_{l+n_{\mathrm{l}}}$ and is identified with $agent_j$.

The initialization of the DAPCA-EC to solve

$$\max_{(\mu,\lambda)\in Q_\mu \times Q_\lambda} f_c(\mu,\lambda) \tag{5.10}$$

in parallel and with event-triggered communication, where $Q_\mu \subset \mathbb{R}^{n_{\mathrm{b}}}$ and $Q_\lambda \subset \mathbb{R}_+^{2n_{\mathrm{l}}}$ are compact and convex sets that are assumed to contain an optimal dual multiplier $(\mu,\lambda)^{\mathrm{opt}}$, is done in the following by choosing $\gamma > 1$, $L_{-1} \in (0, L_c]$, and the starting point $(\bar{\mu},\bar{\lambda})^0 = (\mu,\lambda)^0$ as the minimum of the separable prox-function $d(\mu,\lambda) = (\sigma/2)\left\|(\mu,\lambda)\right\|^2$ with convexity parameter $\sigma > 0$ according to Assumptions 2.2.1. Moreover, let $(\bar{u},\bar{h})^{-1} = (\bar{\mu},\bar{\lambda})^0$.

Algorithm 5.1.1. (DAPCA-EC to solve the DC-OPF problem) *For $k \geq 0$ do in parallel:*
For $i \in \mathcal{N}_b$, given $\bar{\mu}_j^k$ if $j \in N(i) \cup \{i\}$ and $\bar{\lambda}_l^k, \bar{\lambda}_{l+n_l}^k$ if $l \in L(i)$, agent$_i$

1. *computes*

$$P_i^{g,k+1} = \underset{P_i^g \in \mathcal{P}_i}{\arg\min} \left\{ C_i(P_i^g) + \bar{\mu}_i^k P_i^g + \frac{c\sigma_i}{2} P_i^{g\,2} \right\} \ \ \text{if} \ \ i \in \mathcal{N}_g,$$

$$\theta_i^{k+1} = \underset{\theta_i \in \Theta_i}{\arg\min} \left\{ \left(\sum_{l \in L(i)} (\bar{\lambda}_l^k - \bar{\lambda}_{l+n_l}^k) W_{li}^{inc} - \sum_{j \in N(i) \cup \{i\}} \bar{\mu}_j^k B_{ji}^{bus} \right) \theta_i + \frac{c\sigma_{i+n_g}}{2} {\theta_i}^2 \right\},$$

and sends θ_i^{k+1} to agent$_{\mu_j}$, agent$_{\lambda_l}$, and agent$_{\lambda_{l+n_l}}$ if $j \in N(i)$ and $l \in L(i)$.

For $i \in \mathcal{N}_b$ and $l \in \mathcal{N}_l$, given the iterates $P_i^{g,k+1}$ and θ_i^{k+1} that are necessary for the computation of the partial derivatives of $f_c((\bar{\mu}, \bar{\lambda})^k)$, agent$_{\mu_i}$, agent$_{\lambda_l}$, and agent$_{\lambda_{l+n_l}}$

2. *compute*

$$\nabla_{\mu_i}^k = \frac{\partial f_c((\bar{\mu}, \bar{\lambda})^k)}{\partial \mu_i} = \left(I^g P^{g,k+1} \right)_i - \sum_{j \in N(i) \cup \{i\}} B_{ij}^{bus} \theta_j^{k+1} - \widetilde{P_i^d},$$

$$\nabla_{\lambda_l}^k = \frac{\partial f_c((\bar{\mu}, \bar{\lambda})^k)}{\partial \lambda_l} = \sum_{i \in G(l)} W_{li}^{inc} \theta_i^{k+1} - F_l^{max},$$

$$\nabla_{\lambda_{l+n_l}}^k = \frac{\partial f_c((\bar{\mu}, \bar{\lambda})^k)}{\partial \lambda_{l+n_l}} = \sum_{i \in G(l)} -W_{li}^{inc} \theta_i^{k+1} + F_l^{min},$$

and set $L_k = L_{k-1}$,

3. *find*

$$u_i^k = \underset{\mu \in Q_{\mu_i} \subset \mathbb{R}}{\arg\max} \left\{ \mu \nabla_{\mu_i}^k - (\eta_{\mu_i} + 1) L_k \Delta_k \left| \mu - \mu_i^k \right| - \frac{L_k}{2} (\mu - \mu_i^k)^2 \right\},$$

$$h_l^k = \underset{\lambda \in Q_{\lambda_l} \subset \mathbb{R}_+}{\arg\max} \left\{ \lambda \nabla_{\lambda_l}^k - (\eta_{\lambda_l} + 1) L_k \Delta_k \left| \lambda - \lambda_l^k \right| - \frac{L_k}{2} (\lambda - \lambda_l^k)^2 \right\},$$

$$h_{l+n_l}^k = \underset{\lambda \in Q_{\lambda_{l+n_l}} \subset \mathbb{R}_+}{\arg\max} \left\{ \lambda \nabla_{\lambda_{l+n_l}}^k - \left(\eta_{\lambda_{l+n_l}} + 1 \right) L_k \Delta_k \left| \lambda - \lambda_{l+n_l}^k \right| - \frac{L_k}{2} (\lambda - \lambda_{l+n_l}^k)^2 \right\}.$$

4. ***if*** $L_k < L_c$ ***then***

(a) *agent*$_{\mu_i}$, *agent*$_{\lambda_l}$, *and agent*$_{\lambda_{l+n_l}}$ *exchange information if necessary:*

if $\left|\bar{u}_i^{k-1} - u_i^k\right| > \frac{L_k}{L_c}\Delta_k$ ***then***

agent$_{\mu_i}$ *sets* $\bar{u}_i^k = u_i^k$ *and sends* $\bar{u}_i^k$ *to agent*$_j$ *if* $j \in N(i)$.

else

agent$_i$ *sets* $\bar{u}_i^k = \bar{u}_i^{k-1}$ *and signals that no data will be sent.*

if $\left|\bar{h}_l^{k-1} - h_l^k\right| > \frac{L_k}{L_c}\Delta_k$ ***then***

agent$_{\lambda_l}$ *sets* $\bar{h}_l^k = h_l^k$ *and sends* $\bar{h}_l^k$ *to agent*$_j$ *if* $l \in L(j)$.

else

agent$_{\lambda_l}$ *sets* $\bar{h}_l^k = \bar{h}_l^{k-1}$ *and signals that no data will be sent.*

Agent$_{\lambda_{l+n_l}}$ *proceeds accordingly with* $h_{l+n_l}^k$ *and* $\bar{h}_{l+n_l}^k$.

For $i \in \mathcal{N}_b$, *given* $\bar{u}_j^k$ *if* $j \in N(i) \cup \{i\}$ *and* $\bar{h}_l^k, \bar{h}_{l+n_l}^k$ *if* $l \in L(i)$, *agent*$_i$

(b) *computes*

$$\widetilde{P}_i^{g,k+1} = \underset{P_i^g \in \mathcal{P}_i}{\arg\min} \left\{ C_i(P_i^g) + \bar{u}_i^k P_i^g + \frac{c\sigma_i}{2} {P_i^g}^2 \right\} \quad \text{if } i \in \mathcal{N}_g,$$

$$\widetilde{\theta}_i^{k+1} = \underset{\theta_i \in \Theta_i}{\arg\min} \left\{ \left(\sum_{l \in L(i)} (\bar{h}_l^k - \bar{h}_{l+n_l}^k) W_{li}^{inc} - \sum_{j \in N(i) \cup \{i\}} \bar{u}_j^k B_{ji}^{bus} \right) \theta_i + \frac{c\sigma_{i+n_g}}{2} {\theta_i}^2 \right\},$$

and sends $\widetilde{\theta}_i^{k+1}$ *to agent*$_{\mu_j}$, *agent*$_{\lambda_l}$, *and agent*$_{\lambda_{l+n_l}}$ *if* $j \in N(i)$ *and* $l \in L(i)$.

For $i \in \mathcal{N}_b$ *and* $l \in \mathcal{N}_l$, *given the iterates* $\widetilde{P}_i^{g,k+1}$ *and* $\widetilde{\theta}_i^{k+1}$ *that are necessary for the computation of the partial derivatives of* $f_c((\bar{u},\bar{h})^k)$, *agent*$_{\mu_i}$, *agent*$_{\lambda_l}$, *and agent*$_{\lambda_{l+n_l}}$

(c) *compute*

$$\frac{\partial f_c(\bar{u}^k,\bar{h}^k)}{\partial u_i} = \left(I^g \widetilde{P}^{g,k+1}\right)_i - \sum_{j \in N(i) \cup \{i\}} B_{ij}^{bus} \widetilde{\theta}_j^{k+1} - \widetilde{P_i^d},$$

$$\frac{\partial f_c(\bar{u}^k,\bar{h}^k)}{\partial h_l} = \sum_{i \in G(l)} W_{li}^{inc} \widetilde{\theta}_i^{k+1} - F_l^{max},$$

$$\frac{\partial f_c(\bar{u}^k,\bar{h}^k)}{\partial h_{l+n_l}} = \sum_{i \in G(l)} -W_{li}^{inc} \widetilde{\theta}_i^{k+1} + F_l^{min},$$

and check with consensus (section 2.4)

if

$$
-\frac{L_k}{2}\sum_{l=1}^{n_b+2n_l}\left((u,h)_l^k-(\mu,\lambda)_l^k\right)^2 \leq
\sum_{l=1}^{n_b+2n_l}\left(\nabla_l f((\bar{u},\bar{h})^k)-\nabla_l f((\bar{\mu},\bar{\lambda})^k)\right)\left((u,h)_l^k-(\mu,\lambda)_l^k\right) \tag{5.11}
$$

then

continue with step 5,

else

set $L_k = L_k\gamma$ *and goe to step 3,*

5. *find*

$$
\begin{aligned}
v_i^k &= \underset{v\in Q_{\mu_i}\subset\mathbb{R}}{\arg\max}\left\{-\frac{L_k}{2}v^2+v\sum_{j=0}^{k}\frac{j+1}{2}\nabla_{\mu_i}^j\right\},\\
t_l^k &= \underset{t\in Q_{\lambda_l}\subset\mathbb{R}_+}{\arg\max}\left\{-\frac{L_k}{2}t^2+t\sum_{j=0}^{k}\frac{j+1}{2}\nabla_{\lambda_l}^j\right\},\\
t_{l+n_l}^k &= \underset{t\in Q_{\lambda_{l+n_l}}\subset\mathbb{R}_+}{\arg\max}\left\{-\frac{L_k}{2}t^2+t\sum_{j=0}^{k}\frac{j+1}{2}\nabla_{\lambda_{l+n_l}}^j\right\},
\end{aligned}
$$

6. *set*

$$
\begin{aligned}
\mu_i^{k+1} &= \frac{k+1}{k+3}u_i^k+\frac{2}{k+3}v_i^k,\\
\lambda_l^{k+1} &= \frac{k+1}{k+3}h_l^k+\frac{2}{k+3}t_l^k,\\
\lambda_{l+n_l}^{k+1} &= \frac{k+1}{k+3}h_{l+n_l}^k+\frac{2}{k+3}t_{l+n_l}^k,
\end{aligned}
$$

7. *and exchange information if necessary:*

if $\left|\bar{\mu}_i^k-\mu_i^{k+1}\right|>\Delta_{k+1}$ ***then***

agent$_{\mu_i}$ *sets* $\bar{\mu}_i^{k+1}=\mu_i^{k+1}$ *and sends* $\bar{\mu}_i^{k+1}$ *to agent*$_j$ *if* $j\in N(i)$.

else

agent$_i$ *sets* $\bar{\mu}_i^{k+1}=\bar{\mu}_i^k$ *and signals that no data will be sent.*

if $\left|\bar{\lambda}_l^k-\lambda_l^{k+1}\right|>\Delta_{k+1}$ ***then***

agent$_{\lambda_l}$ *sets* $\bar{\lambda}_l^{k+1}=\lambda_l^{k+1}$ *and sends* $\bar{\lambda}_l^{k+1}$ *to agent*$_j$ *if* $l\in L(j)$.

else

agent$_{\lambda_l}$ *sets* $\bar{\lambda}_l^{k+1}=\bar{\lambda}_l^k$ *and signals that no data will be sent.*

Agent$_{\lambda_{l+n_l}}$ *proceeds accordingly with* $\lambda_{l+n_l}^{k+1}$ *and* $\bar{\lambda}_{l+n_l}^{k+1}$.

Remark 5.1.2.

1. *For the choice $L_{-1} = L_c$, the DAPCA-EC 5.1.1 coincides with the DPCA-EC [MUA14, Algorithm 4.1] applied to solve the DC-OPF problem.*

2. *Finally, the amount of power generation P_i^g, which may be considered as sensitive information in a competitive power supplying environment, is controlled by the agent$_i$ placed at bus/generator i and not exchanged in Algorithm 5.1.1.*

Regarding the communication exchange of the agents in Algorithm 5.1.1, it follows immediately from the definition of the sets $N(i)$, $L(i)$, and $G(l)$ that the iterates have to be exchanged only locally in the above chosen multi-agent network, i.e., only agents placed at neighboring buses have to communicate with each other, which is exemplarily detailed for the first two steps of Algorithm 5.1.1 in the following (and therefore also holds for step 4 b) and 4 c)):
In step 1, *agent$_i$* only needs the iterate $\bar{\mu}_i^k$ for the computation of $P_i^{g,k+1}$ (if $i \in \mathcal{N}_\mathrm{g}$) which he controls himself. For the computation of θ_i^{k+1}, it follows from the definition of the sets $L(i)$ and $N(i)$ that only iterates $\bar{\lambda}_l^k$, $\bar{\lambda}_{l+n_\mathrm{l}}^k$, and $\bar{\mu}_j^k$ are involved that are controlled by *agent$_i$* and neighboring agents in the power system network.
In step 2, *agent$_i$* can compute the partial derivative with neighborhood information due to the definition of $N(i)$ and the fact that he controls P_i^g (if $i \in \mathcal{N}_\mathrm{g}$) as well as θ_i. Accordingly, *agent$_{\lambda_l}$* and *agent$_{\lambda_{l+n_\mathrm{l}}}$* are only compelled to communicate with neighboring agents by the definition of $G(l)$.
The parallel implementation of the consensus check in step 4 needs only local communication as well, as the multi-agent network has the same topology as the power system network according to the above placement of agents.
Finally, we derive the following bounds on the numbers η_{μ_i}, η_{λ_l}, and $\eta_{\lambda_{l+n_\mathrm{l}}}$ of the dual variables that influence the corresponding partial derivatives $\nabla_{\mu_i} f_c$, $\nabla_{\lambda_l} f_c$, and $\nabla_{\lambda_{l+n_\mathrm{l}}} f_c$:
The partial derivative of f_c with respect to μ_i in step 2 of Algorithm 5.1.1 is given by

$$\frac{\partial f_c(\bar{\mu}^k, \bar{\lambda}^k)}{\partial \mu_i} = \left(I^\mathrm{g} P^{g,k+1}\right)_i - \sum_{j \in N(i) \cup \{i\}} B_{ij}^\mathrm{bus} \theta_j^{k+1} - \widetilde{P_i^\mathrm{d}}$$

and involves the primal iterate $P_i^{g,k+1}$ if $i \in \mathcal{N}_\mathrm{g}$ and θ_j^{k+1} if $j \in N(i) \cup \{i\}$. The computation of

$$P_i^{g,k+1} = \underset{P_i^g \in \mathcal{P}_i}{\arg\min} \left\{ C_i(P_i^g) + \bar{\mu}_i^k P_i^g + \frac{c\sigma_i}{2} {P_i^g}^2 \right\} \quad \text{if } i \in \mathcal{N}_\mathrm{g}$$

involves the dual iterate $\bar{\mu}_i^k$ and is therefore controlled by $agent_i$ himself. On the other side, for $j \in N(i) \cup \{i\}$ the computation of the primal iterate

$$\theta_j^{k+1} = \underset{\theta_j \in \Theta_j}{\arg\min} \left\{ \left(\sum_{l \in L(j)} (\bar{\lambda}_l^k - \bar{\lambda}_{l+n_l}^k) W_{lj}^{\text{inc}} - \sum_{t \in N(j) \cup \{j\}} \bar{\mu}_t^k B_{tj}^{\text{bus}} \right) \theta_j + \frac{c\sigma_{j+n_g}}{2} {\theta_j}^2 \right\}$$

involves the dual iterates $\bar{\lambda}_l^k$ and $\bar{\lambda}_{l+n_l}$ if $l \in L(j)$ as well as $\bar{\mu}_t^k$ if $t \in N(j) \cup \{j\}$. As $i \in N(j) \cup \{j\}$ it follows that $agent_{\mu_i}$ controls one of the $\bar{u}_t^k$ with $t \in N(j) \cup \{j\}$ himself and it can be seen that

$$\eta_{\mu_i} < \sum_{j \in N(i) \cup \{i\}} [2|L(j)| + |N(j)|] = \sum_{j \in N(i) \cup \{i\}} 3|N(j)|.$$

The partial derivative of f_c with respect to λ_l is

$$\frac{\partial f_c(\bar{\mu}^k, \bar{\lambda}^k)}{\partial \lambda_l} = \sum_{i \in G(l)} W_{li}^{\text{inc}} \theta_i^{k+1} - F_l^{\max}$$

and involves the primal iterate θ_i^{k+1} if $i \in G(l)$. Repeating the above argumentation it follows that

$$\eta_{\lambda_l} < \sum_{i \in G(l)} [2|L(i)| + |N(i)|] = \sum_{i \in G(l)} 3|N(i)|.$$

Accordingly,

$$\eta_{\lambda_{l+n_l}} < \sum_{i \in G(l)} [2|L(i)| + |N(i)|] = \sum_{i \in G(l)} 3|N(i)|.$$

Remark 5.1.3.

In [WL10] (which as well as [LL99, WL09a] (both related to network utility maximization) inspired the decomposition of the DC-OPF problem and the resulting local communication derived above), it is shown that a reformulated version of the DC-OPF problem can be solved distributedly with event-triggered and local communication by minimizing a corresponding unconstrained augmented cost function with a gradient scheme which converges in case of strictly increasing, convex, and differentiable objective cost functions. In [BB03], the power system is divided into areas that are connected via tie-lines, resulting in an equivalent formulation of the DC-OPF problem that is separable with respect to each area up to coupling constraints that are related to the tie-lines. An iterative scheme is proposed that uses local communication with respect to the tie-line connections. In [JDR08], the DC-OPF problem is decomposed node-wisely by fixing variables (phase angles, multipliers) that have to be exchanged in the optimization process.

5.2 Application of the DAPCA-EC to LMI-constrained problems

The content of this section follows [DMUH15, sec. 4] (Deroo, Meinel, Ulbrich, and Hirche) and [DMUH14a, sec. 3.3] (Deroo, Meinel, Ulbrich, and Hirche), however, generalizes our results from there. Moreover, this section is in preparation for publication in [MU14] (Meinel and Ulbrich).

In preparation for the distributed computation of an approximate solution to the AC-OPF problem with the DAPCA-EC 3.2.2 in the following section, we combine in this section the range-space conversion method [KKMY11, sec. 5.2] with dual decomposition to solve a convex problem, that is constrained by a linear matrix inequality (LMI), in a distributed manner. We applied this combination as well in [DMUH15, sec. 4] and [DMUH14a, sec. 3.3] to solve an LMI-constrained strongly convex stability related problem with the distributed proximal center algorithm in parallel, where either local communication could be achieved if the sparsity structure of the LMI is chordal or close to local communication by considering a minimal chordal extension of the sparsity structure of the nonchrodal LMI.

To generalize these results, consider the following LMI-constrained problem similar to [KKMY11, sec. 7] and [LL12, prob. (21) - (22)]:

$$\min_{x\in X} \sum_{i=1}^{n} \Phi_i(x_i) \tag{5.12a}$$

$$\text{s.t.} \ \sum_{i=1}^{n} x_i A^i \succeq 0, \tag{5.12b}$$

where $\Phi_i \colon \mathbb{R} \to \mathbb{R}$ is a continuous and convex function, $A^i \in \mathbb{S}^n \subseteq \mathbb{R}^{n\times n}$ are given symmetric matrices, and $X \subseteq \mathbb{R}^n$ is a given compact and convex set that is component-wisely block-separable, i.e., $X = X_1 \times \ldots \times X_n$ with $X_i \subset \mathbb{R}$. By setting $X_1 = \{1\}$, we can consider LMI's of the form $A^1 + \sum_{i=2}^{n} x_i A^i \succeq 0$ too.

Unfortunately, applying dual decomposition to solve the convex problem (5.12) in parallel introduces a symmetric dual matrix-multiplier of dimension $n \times n$ that can only be updated centrally by using global information as will be more clear if we have a look at the dual problem of (5.12) that is obtained with the corresponding Lagrangian [LL12, App.]

$$\mathcal{L}(x,W) = \sum_{i=1}^{n} \Phi_i(x_i) + \mathrm{Tr}\left\{ \sum_{i=1}^{n} x_i A^i W \right\} = \sum_{i=1}^{n} \Phi_i(x_i) + \sum_{i=1}^{n} x_i A^i \bullet W,$$

where [GK02, p. 161]

$$\mathrm{Tr}(W) = \sum_{i=1}^{n} W_{ii} \quad \text{and} \quad A \bullet W = \mathrm{Tr}(AW^T) = \sum_{i,j=1}^{n} A_{ij} W_{ij}.$$

It follows that the corresponding dual problem is given by [LL12, App.]

$$\max_{W \preceq 0} f(W) = \max_{W \preceq 0} \sum_{i=1}^{n} \min_{x_i \in X_i} \left\{ \Phi_i(x_i) + x_i A^i \bullet W \right\}.$$

The application of a subgradient scheme [GK02, Algo. 6.50] to maximize the dual function $f(W)$ needs the following subgradient of the iterate W^k in iteration $k \geq 0$ [GK02, Lem. 6.23]:

$$g^k = \sum_{i=1}^{n} x_i^{k+1} A^i \in \partial f(W^k). \tag{5.13}$$

In a multi-agent network consisting of n agents, where $agent_{x_i}$ is responsible for updating the primal iterate x_i^k and $agent_W \in \{agent_{x_1}, \ldots, agent_{x_n}\}$ is responsible for updating the dual iterate W^k with the subgradient at W^k, $agent_W$ needs to communicate in each iteration $k \geq 0$ with every $agent_{x_i} \neq agent_W$ to be able to compute g^k in (5.13). This is certainly not desirable with respect to sensitivity of information and the spatial distribution of a large-scale network.

Taking these considerations into account, it is subsequently shown how to decompose LMI (5.12b) with the range-space conversion method from [KKMY11, sec. 5.2], yielding an equivalent problem to (5.12) whose augmented dual problem can be solved by the DAPCA-EC 3.2.2 with local communication if the subsequent Assumptions 5.2.1 are satisfied, stated after the following definitions from [KKMY11, sec. 5.1]:

Let $A(x) = \sum_{i=1}^{n} x_i A^i$ denote the left-hand side of LMI (5.12b). The range-space sparsity pattern (in the following simply denoted as sparsity pattern) of $A(x)$ is defined by

$$F = \left\{ (l,j) \in \mathcal{N} \times \mathcal{N} \ : \ A_{lj}(x) \neq 0 \ \text{ for some } \ x \in X,\, l \neq j \right\}, \tag{5.14}$$

where $\mathcal{N} = \{1, \ldots, n\}$. The corresponding range-space sparsity pattern graph is defined by $G = (\mathcal{N}, F)$ which is an undirected graph by identifying $(l,j) \in F$ with $(j,l) \in F$.

A graph is called chordal if every cycle with more than 3 edges has an edge connecting two nonadjacent nodes in the cycle [KKMY11, sec. 2.3].

Assumptions 5.2.1.

1. *The sparsity pattern of A^i is induced by the neighborhood of $agent_{x_i}$ in the multi-agent network, i.e., $A^i_{lj} = 0$ if $i \notin \{l,j\}$ for $l \neq j$ and $A^i_{ij} = A^i_{ji} = 0$ for $i \neq j$ if $agent_{x_j}$ and $agent_{x_i}$ are not neighbors. Moreover, $A^i_{ll} = 0$ if $agent_{x_l}$ is neither a neighbor of $agent_{x_i}$ nor $agent_{x_i}$ himself.*

2. *The graph $G = (\mathcal{N}, F)$ is chordal.*

From the first assumption it follows that the sparsity pattern of $A(x)$ is induced by the sparsity pattern of the adjacency matrix that represents the multi-agent network, i.e., if $agent_{x_l}$ and $agent_{x_j}$ are not neighbors then $(l,j) \notin F$ (and $(j,l) \notin F$). The chordality-assumption of $G = (\mathcal{N}, F)$ that we made in [DMUH15, DMUH14a] as well, ensures the local communication of the agents in the optimization process as will be discussed in detail below Algorithm 5.2.6. As mentioned above, close to local communication can be achieved if $G = (\mathcal{N}, F)$ is not chordal by finding a minimal chordal extension $\overline{G} = (\mathcal{N}, \overline{F})$ of $G = (\mathcal{N}, F)$. The communication topology of the agents that implement the DAPCA-EC in parallel will then be described by $\overline{G}$ and might not be induced by the topology of the multi-agent network anymore which is given by G.

Remark 5.2.2.
Computationally, a chordal extension of a graph $G = (\mathcal{N}, F)$ can be obtained by applying Cholesky factorization to a positive definite matrix X that has the same sparsity structure as the adjacency matrix of G as described in [FKMN01, 2.1]: Let L be the lower-triangular matrix of the Cholesky factorization of X (available in MATLAB with the function chol.m [KKMY11, Rem. 3.2]), i.e., $X = LL^T$. Then the sparsity pattern $\overline{F}$ of $L + L^T$ yields a chordal extension $\overline{G} = (\mathcal{N}, \overline{F})$ of G. Even though the problem of finding a minimal chordal extension of G is NP-complete [Yan81, Theo. 1], heuristics such as the minimum-degree ordering (available in MATLAB with the function symamd.m [KKMY11, Rem. 3.2]) exist to determine a permutation of X such that the Cholesky factorization often yields minimal fill-ins in L which correspond to the additional edges in the chordal extension of the graph.

To be able to apply the range-space conversion method, let $C_1, \ldots, C_p \subseteq \mathcal{N}$ denote the maximal cliques of the graph $G = (\mathcal{N}, F)$ (that is chordal according to Assumptions 5.2.1), i.e., the graph $G_s = (C_s, C_s \times C_s \cap F)$ is a complete subgraph of G that is not contained in a different complete subgraph of G [Gol04, p. 6]. The maximal cliques of a chordal graph G can be found in $\mathcal{O}(|\mathcal{N}| + |F|)$ time [Gol04, Theo. 4.17] and are computed in this work with the MATLAB function *maximalCliques* provided by Jeffrey Wildman. Moreover, define by [KKMY11, sec. 4]

$$\mathbb{S}^n_+(F,0) = \left\{ X \in \mathbb{S}^n_+ \; : \; X_{lj} = 0 \text{ if } l \neq j \wedge (l,j) \notin F \right\}$$

the set of positive semidefinite matrices whose sparsity pattern is induced by F and let [KKMY11, sec. 2.2]

$$\mathbb{S}^{C_s}_+ = \left\{ X \in \mathbb{S}^n_+ \; : \; X_{lj} = 0 \text{ if } (l,j) \notin C_s \times C_s \right\} \text{ for } s = 1, \ldots, p$$

be the set of positive semidefinite matrices whose sparsity pattern is induced by $C_s \times C_s$, where C_s is a maximal clique of $G = (\mathcal{N}, F)$. The following theorem builds the basis of the range-space conversion method and states that the left-hand side $A(x)$ of LMI (5.12b) is positive semidefinite if and only if it is decomposable by p positive semidefinite matrices whose sparsity patterns are induced by the maximal cliques of the chordal graph that represents the sparsity pattern of $A(x)$.

Theorem 5.2.3. [KKMY11, Theo. 4.2]
$A(x) \in \mathbb{S}^n_+(F,0)$ for $x \in X$ if and only if there exist $W^s \in \mathbb{S}^{C_s}_+$ for $s = 1, \dots, p$ which decompose $A(x)$ as

$$A(x) = \sum_{s=1}^{p} W^s.$$

Proof. The proof is given in [KKMY11]. □

Obviously, constraint (5.12b) can be decomposed with Theorem 5.2.3 to [KKMY11, sec. 5.2]

$$\sum_{s=1}^{p} W^s - A(x) = 0 \quad \text{and} \quad W^s \in \mathbb{S}^{C_s}_+ \quad \text{for} \quad s = 1, \dots, p, \tag{5.15}$$

i.e., the positive semidefinite condition to $A(x)$ in constraint (5.12b) is reduced to the positive semidefinite condition to W^s for $s = 1, \dots, p$ that can be ensured locally for every maximal clique of $G = (\mathcal{N}, F)$ as will be shown.
Finally, the symmetry and the sparsity of the matrices in (5.15) are exploited in the range-space conversion method by the following definitions from [KKMY11, sec. 2.2, sec. 5.2]: For every $(l, j) \in \mathcal{N} \times \mathcal{N}$, define

$$E_{lj} \in \mathbb{R}^{n \times n} \quad \text{with} \quad (E_{lj})_{ik} = \begin{cases} 1 & \text{if } (i,k) \in \{(l,j),(j,l)\}, \\ 0 & \text{else,} \end{cases}$$

and let

$$\begin{aligned} J(C_s) &= \{(l,j) \in C_s \times C_s \; : \; 1 \le l \le j \le n\} \text{ for } s = 1, \dots, p, \\ J &= \bigcup_{s=1}^{p} J(C_s), \\ \Gamma(l,j) &= \{s : \; l \in C_s, \; j \in C_s\} \text{ for every } (l,j) \in J. \end{aligned}$$

With the above definitions, the n^2 equality constraints in (5.15) can be reduced to $|J|$

equality constraints, yielding that problem (5.12) can be stated as [KKMY11, sec. 5.2]

$$\min_{x\in X} \sum_{i=1}^{n} \Phi_i(x_i) \tag{5.16a}$$

$$\text{s.t. } E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - E_{lj} \bullet A(x) = 0 \quad \text{for} \quad (l,j) \in J, \tag{5.16b}$$

$$W^s \in \mathbb{S}_+^{C_s} \quad \text{for} \quad s = 1,\ldots,p. \tag{5.16c}$$

Remark 5.2.4.

According to [KKMY11, sec. 5.1], the range-space conversion method can be applied even if $G = (\mathcal{N},F)$ *is not chordal, as* $A(x) \in \mathbb{S}_+^n(F,0)$ *implies* $A(x) \in \mathbb{S}_+^n(\overline{F},0)$ *for some chordal extension* $\overline{G} = (\mathcal{N},\overline{F})$ *of G, and Theorem 5.2.3 holds for the maximal cliques* $C_1,\ldots,C_p$ *of the chordal graph* $\overline{G}$*. However, the pair* $(l,j) \in J$ *might not be identifiable with a line connecting* agent_l *and* agent_j *in the multi-agent network anymore and therefore local communication may not be guaranteed as will be discussed below Algorithm 5.2.6.*

Problem (5.16) can now be dually decomposed as done in the previous sections to prepare the application of the DAPCA-EC 3.2.2. To be able to apply the derived convergence results from chapter 3, define the compact and convex set

$$\mathcal{W}^s = \left\{ W \in \mathbb{S}_+^{C_s} \,:\, W \preceq I_n R_{W^s} \right\}, \tag{5.17}$$

where $R_{W^s} > 0$ for $s = 1,\ldots,p$ and I_n is the $n \times n$ identity matrix. We assume that $\mathcal{W}^s$ contains an optimal solution $W^{s^{\text{opt}}}$ of problem (5.16) which can then be rewritten as

$$\min_{x_i\in X_i, W^s\in\mathcal{W}^s} \sum_{i=1}^{n} \Phi_i(x_i) \tag{5.18a}$$

$$\text{s.t. } E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - E_{lj} \bullet A(x) = 0 \quad \text{for} \quad (l,j) \in J. \tag{5.18b}$$

As explained in section 3.1, the Lagrangian relaxation of problem (5.18) with respect to constraint (5.18b) yields

$$\begin{aligned}
\mathcal{L}(x,W^1,\ldots,W^p,\Lambda) &= \sum_{i=1}^{n} \Phi_i(x_i) + \sum_{(l,j)\in J} \Lambda_{lj} \left(E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - E_{lj} \bullet A(x) \right) \\
&= \sum_{i=1}^{n} \Phi_i(x_i) + \sum_{(l,j)\in J} \Lambda_{lj} \left(E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - E_{lj} \bullet \sum_{i=1}^{n} A^i x_i \right) \\
&= \sum_{i=1}^{n} \Phi_i(x_i) + \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \sum_{i=1}^{n} A^i x_i \\
&= \sum_{i=1}^{n} \left[\Phi_i(x_i) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet A^i x_i \right] + \sum_{s=1}^{p} \sum_{(l,j)\in J(C_s)} \Lambda_{lj} E_{lj} \bullet W^s.
\end{aligned}$$

By smoothing the Lagrangian with scaled prox-functions $d_{x_i}(x)$ and $d_{W^s}(W)$ with convexity parameter $\sigma_{x_i} > 0$ and $\sigma_{W^s} > 0$, respectively, we obtain the smooth and concave augmented dual function

$$
\begin{aligned}
f_c(\Lambda) = \min_{x\in X,\, W^s\in\mathcal{W}^s} \Bigg\{ \sum_{i=1}^{n} \left[\Phi_i(x_i) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet A^i x_i + c d_{x_i}(x_i) \right] \\
+ \sum_{s=1}^{p} \left[\sum_{(l,j)\in J(C_s)} \Lambda_{lj} E_{lj} \bullet W^s + c d_{W^s}(W^s) \right] \Bigg\} \\
= \sum_{i=1}^{n} \min_{x_i\in X_i} \left\{ \Phi_i(x_i) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet A^i x_i + c d_{x_i}(x_i) \right\} \qquad (5.19)\\
+ \sum_{s=1}^{p} \min_{W^s\in\mathcal{W}^s} \left\{ \sum_{(l,j)\in J(C_s)} \Lambda_{lj} E_{lj} \bullet W^s + c d_{W^s}(W^s) \right\}. \qquad (5.20)
\end{aligned}
$$

With Theorem 3.1.1, it follows immediately that the partial derivatives of f_c are

$$\nabla_{\Lambda_{lj}} f_c(\Lambda) = E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s(\Lambda) - E_{lj} \bullet \sum_{i=1}^{n} A^i x_i(\Lambda) \quad \text{for} \quad (l,j) \in J,$$

where $W^s(\Lambda)$ and $x_i(\Lambda)$ solve (5.19) and (5.20). To determine the Lipschitz constant L_c of ∇f_c according to Theorem 3.1.1, let $\mathrm{v}\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ be the operator that concatenates the columns of an $n \times n$ dimensional matrix to a vector of dimension n^2. It follows that constraint (5.18b) can be stated as

$$\mathrm{v}\left(E_{lj}\right)^T \sum_{s\in\Gamma(l,j)} \mathrm{v}\left(W^s\right) - \sum_{i=1}^{n} \mathrm{v}\left(E_{lj}\right)^T \mathrm{v}\left(A^i\right) x_i = 0 \quad \text{for} \quad (l,j) \in J, \qquad (5.21)$$

i.e., the coefficient matrix of $\mathrm{v}(W^s)$ in constraint (l,j) is $\mathrm{v}(E_{lj})^T$ iff $(l,j) \in J(C_s)$ and $0^T \in \mathbb{R}^{1\times n^2}$ else, yielding the overall coefficient matrix $E_{W_s} \in \mathbb{R}^{|J|\times n^2}$ of $\mathrm{v}(W^s)$ with

$$(E_{W_s})_{ik} = \begin{cases} \left(\mathrm{v}(E_{lj})^T\right)_{ik} & \text{if } i = (l,j) \in J(C_s), \\ 0 & \text{else}. \end{cases} \qquad (5.22)$$

The determination of the coefficient matrix of x_i in (5.21) is straight forward and given by

$$E_{x_i} = \left(\mathrm{v}\left(E_{11}\right)^T \mathrm{v}\left(A^i\right), \ldots, \mathrm{v}\left(E_{nn}\right)^T \mathrm{v}\left(A^i\right) \right)^T \in \mathbb{R}^{|J|\times 1},$$

yielding the following Lipschitz constant according to Theorem 3.1.1:

$$L_c = \sum_{s=1}^{p} \frac{\|E_{W_s}\|^2}{c\sigma_{W^s}} + \sum_{i=1}^{n} \frac{\|E_{x_i}\|^2}{c\sigma_{x_i}}. \qquad (5.23)$$

Remark 5.2.5.
In [DMUH15, DMUH14a], we applied the range-space conversion method combined with dual decomposition as described above to a problem of the form (5.12), however, with strongly convex objective function, and derived similar partial derivatives for the dual function as well as a similar Lipschitz constant. The derived decomposed problem can be solved distributedly to test a sufficient condition for Lyapunov stability of an LTI system with n subsystems. More precisely, the developed test can be conducted to check if a block-diagonal matrix P exists which satisfies the Lyapunov matrix inequality

$$A^T P + PA \prec 0,$$

where the matrix A describes the interaction of the subsystems of an LTI system. For details it is referred to [DMUH15].

To state the DAPCA-EC, let $Q = Q_{11} \times \ldots \times Q_{nn} \subseteq \mathbb{R}^{|J|}$ be a compact and convex set that contains an optimal dual multiplier Λ^{opt} of the dual problem of problem(5.18) (cf. chapter 3), and consider the corresponding augmented dual problem

$$\max_{\Lambda \in Q} f_c(\Lambda). \tag{5.24}$$

To guarantee local information exchange when solving (5.24) in parallel, $agent_{\Lambda_{lj}}$ is identified either with $agent_{x_l}$ or with $agent_{x_j}$ for $l \neq j$. This setting is favorable as the index pair $(l, j) \in J$ can be identified with a line connecting $agent_{x_l}$ and $agent_{x_j}$ in the multi-agent network according to Assumptions 5.2.1. Finally, $agent_{\Lambda_{ll}}$ is identified with $agent_{x_l}$.
The initialization of the DAPCA-EC is done according to the description in section 3.2 by choosing $\gamma > 1$, $L_{-1} \in (0, L_c]$, and the starting point $\bar{\Lambda}^0 = \Lambda^0$ as the minimum of the separable prox-function $d(\Lambda) = \sum_{(l,j)\in J} d_{lj}(\Lambda_{lj})$, where $d_{lj} \colon Q_{lj} \to \mathbb{R}_+$ is an arbitrary prox-function with convexity parameter $\sigma > 0$. Moreover, let $\bar{Y}^k$ denote the outdated vector defined by (3.2.1) and set $\bar{Y}^{-1} = \bar{\Lambda}^0$. The the DAPCA-EC can be stated as (cf. [DMUH15, Algo. 3],[DMUH14a, Algo. 2])

Algorithm 5.2.6. (DAPCA-EC to solve (5.12)) *For $k \geq 0$ do in parallel:*
For $i = 1, \ldots, n$ and $s = 1, \ldots p$, given the required components of $\bar{\Lambda}^k$, $agent_{x_i}$ and $agent_{W^s}$

1. *compute*

$$x_i^{k+1} = \operatorname*{arg\,min}_{x_i \in X_i} \left\{ \Phi_i(x_i) - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet A^i x_i + c d_{x_i}(x_i) \right\}, \tag{5.25}$$

$$W^{s,k+1} = \operatorname*{arg\,min}_{W^s \in \mathcal{W}^s} \left\{ \sum_{(l,j)\in J(C_s)} \bar{\Lambda}_{lj}^k E_{lj} \bullet W^s + c d_{W^s}(W^s) \right\}, \tag{5.26}$$

and send x_i^{k+1} and $W^{s,k+1}$ to the dual agents that require it.

For $(l, j) \in J$, given the iterates x_i^{k+1} and $W^{s,k+1}$ that are necessary for the computation of $\nabla_{lj} f_c(\bar{\Lambda}^k)$, agent$_{\Lambda_{lj}}$

2. *computes*

$$\nabla_{lj} f_c(\bar{\Lambda}^k) = E_{lj} \bullet \sum_{s \in \Gamma(l,j)} W^{s,k+1} - E_{lj} \bullet \sum_{i=1}^{n} A^i x_i^{k+1},$$

and sets $L_k = L_{k-1}$,

3. *finds*

$$Y_{lj}^k = \underset{Y_{lj} \in Q_{lj}}{\arg\max} \left\{ \nabla_{lj} f_c(\bar{\Lambda}^k) Y_{lj} - L_k \Delta_k (\eta_{lj} + 1) |Y_{lj} - \Lambda_{lj}^k| - \frac{L_k}{2} \left(Y_{lj} - \Lambda_{lj}^k \right)^2 \right\}. \quad (5.27)$$

4. ***if*** $L_k < L_c$ ***then***

(a) *agent$_{\Lambda_{lj}}$ sends $\bar{Y}_{lj}^k$ to the primal agents that require it if necessary:*
if $\left| \bar{Y}_{lj}^{k-1} - Y_{lj}^k \right| > \frac{L_k}{L_c} \Delta_k$ ***then***
agent$_{\Lambda_{lj}}$ sets $\bar{Y}_{lj}^k = Y_{lj}^k$ and sends $\bar{Y}_{lj}^k$.
else
agent$_{\Lambda_{lj}}$ sets $\bar{Y}_{lj}^k = \bar{Y}_{lj}^{k-1}$ and signals that no data will be sent.

For $i = 1, \dots, n$ and $s = 1, \dots p$, given the required components of $\bar{Y}^k$, agent$_{x_i}$ and agent$_{W^s}$

(b) *compute*

$$y_i^{k+1} = \underset{x_i \in X_i}{\arg\min} \left\{ \Phi_i(x_i) - \sum_{(l,j) \in J} \bar{Y}_{lj}^k E_{lj} \bullet A^i x_i + c d_{x_i}(x_i) \right\}, \quad (5.28)$$

$$V^{s,k+1} = \underset{W^s \in \mathcal{W}^s}{\arg\min} \left\{ \sum_{(l,j) \in J(C_s)} \bar{Y}_{lj}^k E_{lj} \bullet W^s + c d_{W^s}(W^s) \right\}, \quad (5.29)$$

and send y_i^{k+1} and $V^{s,k+1}$ to the dual agents that require it.

For $(l, j) \in J$, given the iterates y_i^{k+1} and $V^{s,k+1}$ that are necessary for the computation of $\nabla_{lj} f_c(\bar{Y}^k)$, agent$_{\Lambda_{lj}}$

(c) *computes*

$$\nabla_{lj} f_c(\bar{Y}^k) = E_{lj} \bullet \sum_{s \in \Gamma(l,j)} V^{s,k+1} - E_{lj} \bullet \sum_{i=1}^{n} A^i y_i^{k+1},$$

and checks with consensus (section 2.4) *if*

$$-\frac{L_k}{2}\sum_{(l,j)\in J}\left(Y_{lj}^k-\Lambda_{lj}^k\right)^2 \le \sum_{(l,j)\in J}\left(\nabla_{lj}f(\bar{Y}^k)-\nabla_{lj}f(\bar{\Lambda}^k)\right)\left(Y_{lj}^k-\Lambda_{lj}^k\right) \quad (5.30)$$

then

continues with step 5,

else

sets $L_k = L_k\gamma$ and goes to step 3,

5. *finds*

$$Z_{lj}^k = \arg\max_{Z_{lj}\in Q_{lj}}\left\{-\frac{L_k}{\sigma}d(Z_{lj})+\sum_{t=0}^{k}\frac{t+1}{2}\nabla_{lj}f_c(\bar{\Lambda}^t)Z_{lj}\right\}, \quad (5.31)$$

6. *sets* $\Lambda_{lj}^{k+1} = \frac{2}{k+3}Z_{lj}^k + \frac{k+1}{k+3}Y_{lj}^k,$
7. *and sends $\bar{\Lambda}_{lj}^{k+1}$ to the primal agents that request it if necessary:*

   if $\left|\bar{\Lambda}_{lj}^k-\Lambda_{lj}^{k+1}\right| > \Delta_{k+1}$ ***then***

   $agent_{\Lambda_{lj}}$ sets $\bar{\Lambda}_{lj}^{k+1}=\Lambda_{lj}^{k+1}$ and sends $\bar{\Lambda}_{lj}^{k+1}$.

   else

   $agent_{\Lambda_{lj}}$ sets $\bar{\Lambda}_{lj}^{k+1}=\bar{\Lambda}_{lj}^{k}$ and signals that no data will be sent.

Even though Algorithm 5.2.6 consists of a subproblem in almost every step, analytical solutions exist for all of them if $\nabla\Phi_i(x_i)$, $\nabla d_{x_i}(x_i)$, $\nabla d_{W^s}(W^s)$, $\nabla d(\Lambda_{lj})$ can be determined analytically and if the prox-functions are chosen as

$$d_{x_i}(x_i)=\frac{\sigma_{x_i}}{2}x_i^2,\ d_{W^s}(W^s)=\frac{\sigma_{W^s}}{2}\|W^s\|_F^2,\ \text{and } d(\Lambda_{lj})=\frac{\sigma}{2}\Lambda_{lj}^2,$$

where $\|\cdot\|_F$ denotes the Frobenius norm [GK02, p. 162]. In that case, the analytical solutions to subproblems (5.25),(5.27),(5.28), and (5.31) can be obtained as in Example 2.2.6 and therefore only the analytical solution for subproblem (5.26) (and thereby for subproblem (5.29)) is derived by rewriting it similarly to Remark 2.3.1 (Michael Ulbrich, personal communication, June 17, 2013):

$$\begin{aligned}
W^{s,k+1} &= \arg\min_{W^s\in\mathcal{W}^s}\left\{\overbrace{\sum_{(l,j)\in J(C_s)}\bar{\Lambda}_{lj}^kE_{lj}}^{=X\in\mathbb{S}^{C_s}}\bullet W+\frac{c\sigma_{W^s}}{2}\|W^s\|_F^2\right\}\\
&= \arg\min_{W^s\in\mathcal{W}^s}\left\{\frac{2}{c\sigma_{W^s}}X\bullet W+\|W^s\|_F^2\right\}\\
&= \arg\min_{W^s\in\mathcal{W}^s}\left\|\frac{X}{c\sigma_{W^s}}+W^s\right\|_F^2\\
&= \arg\min_{W^s\in\mathcal{W}^s}\left\|SDS^T+W^s\right\|_F^2.
\end{aligned}$$

It follows with [Hig88, Theo. 2.1] that

$$W^{s,k+1} = -S\hat{D}S^T \text{ with } \hat{D}_{ii} = \min(\max(0, D_{ii}), R_{W^s}).$$

Finally, the matrices S and D can be efficiently obtained by the diagonalization of the $|C_s| \times |C_s|$-dimensional part of $X/(c\sigma_{W^s})$ that is nonzero.

Remark 5.2.7.
In [DMUH15, DMUH14a] we applied the DPCA (without even-triggered communication) instead of the DAPCA-EC and obtained similar analytical solutions (with the difference, that there we did not need to consider compact feasible sets for the primal variables due to the strongly convexity of the considered primal objective functions). This shows the remarkable advantage of the application of the DAPCA-EC (or different versions of it) to a problem of the form (5.12) that has been decomposed with the range-space conversion method and dual decomposition as described above.

Subsequently, it is verified that the communication exchange between the agents is local. For the computation of x_i^{k+1} in step 1 , $agent_{x_i}$ requires the dual iterate $\bar{\Lambda}_{lj}^k$ if $A_{lj}^i = A_{jl}^i \neq 0$ and therefore $\bar{\Lambda}_{lj}^k$ is either controlled by $agent_{x_i}$ himself or by a neighbor of $agent_{x_i}$ in the multi-agent network. The same holds for the computation of y_i^{k+1}.
The computation of $W^{s,k+1}$ involves $\bar{\Lambda}_{lj}^k$ for all $(l,j) \in J(C_s)$. It follows that $agent_{W^s}$ either controls $\bar{\Lambda}_{lj}^k$ by himself or $\bar{\Lambda}_{lj}^k$ is controlled by a neighbor of $agent_{W^s}$ in the clique C_s. The same holds for the computation of $V^{s,k+1}$.
Finally, $agent_{\Lambda_{lj}}$ needs to compute $\nabla_{lj} f_c(\bar{\Lambda}^k)$ in step 2 which involves $W^{s,k+1}$ for $s \in \Gamma(l,j)$ and $(A_{lj}^i + A_{jl}^i)x_i^{k+1}$ for $i = 1, \dots, n$. Firstly, for $(l,j) \in J$ we have $s \in \Gamma(l,j)$ iff $(l,j) \in C_s \times C_s$, i.e., $W^{s,k+1}$ is either controlled by $agent_{\Lambda_{lj}}$ or by a neighbor of $agent_{\Lambda_{lj}}$ in the clique C_s. Secondly, it follows immediately with Assumptions 5.2.1 that the required iterates x_i^{k+1} are either controlled by $agent_{\Lambda_{lj}}$ or by neighbors of $agent_{\Lambda_{lj}}$ in the multi-agent network. The same holds for the computation of $\nabla_{lj} f_c(\bar{Y}^k)$.
The parallel implementation of the consensus check in step 4 needs only local communication as well which follows from the definition of the consensus matrix A in section 2.4 and the setting of the multi-agent network.
Finally, in step 3 of Algorithm 5.2.6 the number η_{lj} of dual variables that influence the corresponding partial derivative $\nabla_{lj} f_c(\Lambda)$ can be estimated as follows.
In step 2 of Algorithm 5.2.6, the partial derivative is

$$\nabla_{lj} f_c(\bar{\Lambda}^k) = E_{lj} \bullet \sum_{s \in \Gamma(l,j)} W^{s,k+1} - E_{lj} \bullet \sum_{i=1}^{n} A^i x_i^{k+1}$$

which on the one hand involves the primal iterate

$$W^{s,k+1} = \underset{W^s \in \mathcal{W}^s}{\arg\min} \left\{ \sum_{(l,j) \in J(C_s)} \bar{\Lambda}_{lj}^k E_{lj} \bullet W^s + c d_{W^s}(W^s) \right\}$$

if $s \in \Gamma(l,j)$, where each iterate $W^{s,k+1}$ depends on $|J(C_s)|$ dual iterates $\bar{\Lambda}_{lj}^k$.
On the other hand, it follows for $l \neq j$ and $l \in N(j)$ with Assumptions 5.2.1 and the definition of the set J that the partial derivative $\nabla_{lj} f_c(\bar{\Lambda}^k)$ additionally depends on the primal iterates x_l^{k+1} if $A_{lj}^l \neq 0$ and x_j^{k+1} and $A_{lj}^j \neq 0$. Moreover, with Assumptions 5.2.1 it can be seen for $l = j$ that $\nabla_{lj} f_c(\bar{\Lambda}^k)$ depends on the primal iterate x_i^{k+1} if $i \in N(j) \cup \{j\}$ and $A_{jj}^i \neq 0$. Finally, each primal iterate

$$x_i^{k+1} = \underset{x_i \in X_i}{\arg\min} \left\{ \Phi_i(x_i) - \sum_{(l,j) \in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet A^i x_i + c d_{x_i}(x_i) \right\}$$

involves according to Assumptions 5.2.1 at most the dual iterate $\bar{\Lambda}_{lj}^k$ if $l = i$ and $j \in N(i)$, if $j = i$ and $l \in N(i)$, and if $j = l$ and $j \in N(i) \cup \{i\}$.
Altogether, we obtain that

$$\eta_{lj} < \sum_{s \in \Gamma(l,j)} |J(C_s)| + \sum_{j \in N(i) \cup \{i\}} (3|N(j)| + 1) \, .$$

5.3 Application of the DAPCA-EC to the AC-OPF problem

The content of this section is in preparation for publication in [MU14] (Meinel and Ulbrich).

In this section, the range-space conversion method is applied in combination with dual decomposition to a concave semidefinite dual of the nonconvex AC-OPF problem (4.15) derived in [LL10, LL12], where [LL12] extends [LL10] by considering additional constraints such as line flow limits.
The application of the DAPCA-EC 3.2.2 to the decomposed dual enables to compute a solution to the AC-OPF problem in parallel and with event-triggered communication. Furthermore, the communication exchange is local with respect to the power system topology if the considered power network is chordal which holds for distribution and subtransmission networks as they have a tree structure but not for transmission networks (such as the IEEE benchmark systems [Uni]) as they have closed loops due to stability reasons [Mom01, p. 4]. However, close to local communication can be achieved for nonchordal power systems by finding minimal chordal extensions as described in the previous section.

Due to the nonconvexity of the NP-hard AC-OPF problem

$$\min_{P^{\mathrm{g}},Q^{\mathrm{g}},V} \sum_{i\in\mathcal{N}_{\mathrm{g}}} C_i\left(P_i^{\mathrm{g}}\right) \tag{5.32a}$$

$$\text{s.t. } V_i I_i^* = \left(P_i^{\mathrm{g}} - P_i^{\mathrm{d}}\right) + \mathrm{j}\left(Q_i^{\mathrm{g}} - Q_i^{\mathrm{d}}\right) \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{5.32b}$$

$$V_i I_i^* = -P_i^{\mathrm{d}} - \mathrm{j}Q_i^{\mathrm{d}} \qquad \forall i \in \mathcal{N}_{\mathrm{b}} \setminus \mathcal{N}_{\mathrm{g}}, \tag{5.32c}$$

$$|S_{ij}| \leq S_{ij}^{\max} \qquad \forall\,(i,j) \in \mathcal{N}_{\mathrm{l}}, \tag{5.32d}$$

$$|S_{ji}| \leq S_{ij}^{\max} \qquad \forall\,(i,j) \in \mathcal{N}_{\mathrm{l}}, \tag{5.32e}$$

$$P_i^{\min} \leq P_i^{\mathrm{g}} \leq P_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{5.32f}$$

$$Q_i^{\min} \leq Q_i^{\mathrm{g}} \leq Q_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{g}}, \tag{5.32g}$$

$$V_i^{\min} \leq |V_i| \leq V_i^{\max} \qquad \forall i \in \mathcal{N}_{\mathrm{b}}, \tag{5.32h}$$

the authors of [LL12, LL12] propose to solve a concave semidefinite dual problem constrained, i.a., by a linear matrix inequality (LMI) which can be stated after the following definitions taken from [LL10, sec. 2] and [LL12, sec. 3], where the details on the derivation can be found.

Let $\mathcal{E} \subseteq \mathcal{N}_{\mathrm{b}} \times \mathcal{N}_{\mathrm{b}}$ be the symmetric relation which contains the indices of the branches and their reversals, i.e., $(t,m) \in \mathcal{N}_{\mathrm{l}} \iff (t,m) \in \mathcal{E}$ and $(m,t) \in \mathcal{E}$. Moreover, denote by $e_1,\ldots,e_{n_{\mathrm{b}}}$ the standard basis vectors in $\mathbb{R}^{n_{\mathrm{b}}}$ and define for $i \in \mathcal{N}_{\mathrm{b}}$ and $(t,m) \in \mathcal{E}$:

$$Y_i = e_i e_i^T Y^{\mathrm{bus}}, \qquad Y_{tm} = \left(y_t^{\mathrm{sh}} - Y_{tm}^{\mathrm{bus}}\right) e_t e_t^T + Y_{tm}^{\mathrm{bus}} e_t e_m^T,$$

$$\mathbf{Y}_i = \frac{1}{2}\begin{pmatrix} \mathrm{Re}\{Y_i + Y_i^T\} & \mathrm{Im}\{Y_i^T - Y_i\} \\ \mathrm{Im}\{Y_i - Y_i^T\} & \mathrm{Re}\{Y_i + Y_i^T\} \end{pmatrix}, \quad \mathbf{Y}_{tm} = \frac{1}{2}\begin{pmatrix} \mathrm{Re}\{Y_{tm} + Y_{tm}^T\} & \mathrm{Im}\{Y_{tm}^T - Y_{tm}\} \\ \mathrm{Im}\{Y_{tm} - Y_{tm}^T\} & \mathrm{Re}\{Y_{tm} + Y_{tm}^T\} \end{pmatrix},$$

$$\bar{\mathbf{Y}}_i = -\frac{1}{2}\begin{pmatrix} \mathrm{Im}\{Y_i + Y_i^T\} & \mathrm{Re}\{Y_i - Y_i^T\} \\ \mathrm{Re}\{Y_i^T - Y_i\} & \mathrm{Im}\{Y_i + Y_i^T\} \end{pmatrix}, \bar{\mathbf{Y}}_{tm} = -\frac{1}{2}\begin{pmatrix} \mathrm{Im}\{Y_{tm} + Y_{tm}^T\} & \mathrm{Re}\{Y_{tm} - Y_{tm}^T\} \\ \mathrm{Re}\{Y_{tm}^T - Y_{tm}\} & \mathrm{Im}\{Y_{tm} + Y_{tm}^T\} \end{pmatrix},$$

$$M_i = \begin{pmatrix} e_i e_i^T & 0 \\ 0 & e_i e_i^T \end{pmatrix},$$

where the bus admittance matrix Y^{bus} is given by (4.11) and y_t^{sh} is the shunt element (4.8) for $t \in \mathcal{N}_{\mathrm{b}}$.

Moreover, in [LL10, LL12] the following variables are defined for $i \in \mathcal{N}_\mathrm{b}$ and $(t,m) \in \mathcal{E}$ that allow to formulate the semidefinite dual problem of (5.32) in a compact way:

$$x_i = (\lambda_i^{\min}, \lambda_i^{\max}, \bar{\lambda}_i^{\min}, \bar{\lambda}_i^{\max}, \mu_i^{\min}, \mu_i^{\max})^T \in \mathbb{R}_+^6,$$

$$r_i = \begin{pmatrix} 1 & r_i^1 \\ r_i^1 & r_i^2 \end{pmatrix} \in \mathbb{S}_+^4 \text{ for } i \in \mathcal{N}_\mathrm{g}, \qquad r_{tm} = \begin{pmatrix} r_{tm}^1 & r_{tm}^2 & r_{tm}^3 \\ r_{tm}^2 & r_{tm}^4 & r_{tm}^5 \\ r_{tm}^3 & r_{tm}^5 & r_{tm}^6 \end{pmatrix} \in \mathbb{S}_+^6,$$

and $r_i = 0 \in \mathbb{S}_+^4$ for $i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}$. The above defined variables are the dual multipliers corresponding to the constraints of an optimization problem derived in [LL12, sec. 3] that is equivalent to the AC-OPF problem (5.32) . Furthermore, define according to [LL10, LL12] for $i \in \mathcal{N}_\mathrm{b}$

$$\boldsymbol{\lambda}_i = -\lambda_i^{\min} + \lambda_i^{\max} + a_{i1} + 2\sqrt{a_{i2}} r_i^1, \qquad \bar{\boldsymbol{\lambda}}_i = -\bar{\lambda}_i^{\min} + \bar{\lambda}_i^{\max}, \qquad \boldsymbol{\mu}_i = -\mu_i^{\min} + \mu_i^{\max},$$

where $a_{i1} = 0$ and $a_{i2} = 0$ for $i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}$. (Recall that a_{i1} and a_{i2} are the coefficients of the generator cost function C_i for $i \in \mathcal{N}_\mathrm{g}$.)
In [LL12, sec. 3], it is shown that a dual of (5.32) can be stated as

$$\max_{\substack{x_i \in \mathbb{R}_+^6,\ r_i \in \mathbb{S}_+^4 \\ r_{tm} \in \mathbb{S}_+^6}} \sum_{i \in \mathcal{N}_\mathrm{b}} \Phi_i(x_i, r_i) - \sum_{(t,m) \in \mathcal{E}} \Phi_{tm}(r_{tm}) \tag{5.33a}$$

$$\text{s.t.} \ A(x,r) \succeq 0, \tag{5.33b}$$

where

$$\begin{aligned} A(x,r) &= \sum_{i \in \mathcal{N}_\mathrm{b}} \widetilde{\mathbf{Y}}_i(x_i, r_i) + \sum_{(t,m) \in \mathcal{E}} \widetilde{\mathbf{Y}}_{tm}(r_{tm}), \\ \Phi_i(x_i, r_i) &= \lambda_i^{\min} P_i^{\min} - \lambda_i^{\max} P_i^{\max} + \boldsymbol{\lambda}_i P_i^\mathrm{d} + \bar{\lambda}_i^{\min} Q_i^{\min} - \bar{\lambda}_i^{\max} Q_i^{\max} \\ &\quad + \bar{\boldsymbol{\lambda}}_i Q_i^\mathrm{d} + \mu_i^{\min} V_i^{\min 2} - \mu_i^{\max} V_i^{\max 2} + a_{i0} - r_i^2, \\ \Phi_{tm}(r_{tm}) &= (S_{tm}^{\max})^2 r_{tm}^1 + r_{tm}^4 + r_{tm}^6, \\ \widetilde{\mathbf{Y}}_i(x_i, r_i) &= \boldsymbol{\lambda}_i \mathbf{Y}_i + \bar{\boldsymbol{\lambda}}_i \bar{\mathbf{Y}}_i + \boldsymbol{\mu}_i M_i \in \mathbb{R}^{2n_\mathrm{b} \times 2n_\mathrm{b}}, \\ \widetilde{\mathbf{Y}}_{tm}(r_{tm}) &= 2 r_{tm}^2 \mathbf{Y}_{tm} + 2 r_{tm}^3 \bar{\mathbf{Y}}_{tm} \in \mathbb{R}^{2n_\mathrm{b} \times 2n_\mathrm{b}}, \end{aligned}$$

and $a_{i0} = 0$ for $i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}$.

Finally, the following sufficient condition is provided in [LL12] which guarantees a zero duality gap for the AC-OPF problem (5.32):

Theorem 5.3.1. [LL12, part 2 of Theo. 2]
The duality gap is zero for problem (5.32) if its dual (5.33) has an optimal solution (x^{opt}, r^{opt}) such that the positive semidefinite matrix $A(x^{opt}, r^{opt})$ has a zero eigenvalue of multiplicity 2.

Proof. The proof is given detailed in [LL12] and only sketched here:
In [LL12], an equivalent problem to the AC-OPF problem (5.32) is formulated in the matrix variable $W \in \mathbb{R}^{2n_\mathrm{b} \times 2n_\mathrm{b}}$ constrained, i.a., by $W = XX^T$, where $X = (\mathrm{Re}\{V\}^T, \mathrm{Im}\{V\}^T)^T$. This constraint is denoted to be equivalent to the positive semidefiniteness of W and the rank-one constraint $\mathrm{rank}\{W\} = 1$. Relaxing this equivalent problem by removing the rank-one constraint yields a semidefinite problem that is also the dual of (5.33) and strong duality is shown for this pair. Furthermore, it is derived from the KKT conditions of the SDP and the 2-dimensionality assumption of the kernel of $A(x^{\mathrm{opt}}, r^{\mathrm{opt}})$ that an optimal solution W^{opt} with $\mathrm{rank}\{W^{\mathrm{opt}}\} = 1$ can be constructed which solves the SDP and therefore solves the equivalent problem of the AC-OPF problem, i.e., strong duality holds for the pair (5.32) and (5.33). □

Furthermore, in [LL12, sec. 4] it is discussed that this sufficient condition is satisfied generally in practice and moreover the structure of $A\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right)$ is stated as

$$A\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right) = \begin{pmatrix} T\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right) & \bar{T}\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right) \\ -\bar{T}\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right) & T\left(x^{\mathrm{opt}}, r^{\mathrm{opt}}\right) \end{pmatrix} \tag{5.34}$$

which is helpful to derive a solution for the AC-OPF problem as will be shown below. Finally, assuming problem (5.32) to be feasible, the following algorithm can be applied to find a global optimum of the AC-OPF problem (5.32) if its dual satisfies Theorem 5.3.1.

Algorithm 5.3.2. [LL12, LL10, Algorithm for Solving OPF (reduced version)]

1. *Find a solution (x^{opt}, r^{opt}) of the dual problem (5.33).*
2. *Find a vector $(v_1^T, v_2^T)^T \neq 0$ in the kernel of $A(x^{opt}, r^{opt})$.*
3. *Find scalars $\zeta_1, \zeta_2 \in \mathbb{R}$ such that $V^{opt} = (\zeta_1 + j\zeta_2)(v_1 + jv_2)$ is an optimal solution of (5.32).*

As shown in [LL12, proof of Cor. 1], step 3 in Algorithm 5.3.2 is well-defined which is concluded from the fact that subject to the sufficient condition in Theorem 5.3.1, the SPD relaxation of the AC-OPF problem (5.32), that is mentioned in the proof sketch of Theorem 5.3.1, has a rank-one optimal solution $W^{\mathrm{opt}} = X^{\mathrm{opt}} X^{\mathrm{opt}^T}$ that satisfies $A(x^{\mathrm{opt}}, r^{\mathrm{opt}}) X^{\mathrm{opt}} = 0$.

Moreover, from the structure of $A(x^{\text{opt}},r^{\text{opt}})$ given in (5.34) it is deduced in [LL12] that besides $(v_1^T,v_2^T)^T$ the orthogonal vector $(-v_2^T,v_1^T)^T$ is in the kernel of $A(x^{\text{opt}},r^{\text{opt}})$ as well, and from the 2-dimensionality of the kernel the authors infer that

$$\begin{pmatrix}\mathrm{Re}\{V^{\text{opt}}\}\\ \mathrm{Im}\{V^{\text{opt}}\}\end{pmatrix} = X^{\text{opt}} = \zeta_1\begin{pmatrix}v_1\\ v_2\end{pmatrix} + \zeta_2\begin{pmatrix}-v_2\\ v_1\end{pmatrix}$$

for some $\zeta_1, \zeta_2 \in \mathbb{R}$.

In the following we show how to implement each step of Algorithm 5.3.2 distributedly in a multi-agent network that coincides with the power system network if it is chordal. Here, the focus is on step 1, i.e., on finding a solution $(x^{\text{opt}},r^{\text{opt}})$ to the dual problem (5.33) in a distributed manner by applying the DAPCA-EC 3.2.2. Thereby, the communication can be kept local in the power system network (i.e., the branches of the power network additionally serve as communication lines) if the LMI (5.33b) in the dual (5.33) satisfies Assumptions 5.2.1, as discussed below Algorithm 5.3.4. However, according to Remark 5.2.4 and Remark 5.2.2 close to local communication can be achieved for nonchordal systems by finding minimal chordal extensions that induce the communication topology.
To determine an arbitrarily good approximation of the optimal solution $(x^{\text{opt}},r^{\text{opt}})$ distributedly with event-triggered and local (or close to local) communication in the first step of Algorithm 5.3.2, the dual is decomposed as described in section 5.2 to be able to apply the DAPCA-EC 3.2.2.
To this end, we denote in the following by $agent_{x_i,r_i}$ the agent that is placed at bus i and updates $\lambda_i^{\min},\lambda_i^{\max},\bar{\lambda}_i^{\min},\bar{\lambda}_i^{\max},\mu_i^{\min},\mu_i^{\max}$ if $i \in \mathcal{N}_{\text{b}}$ and additionally r_i if $i \in \mathcal{N}_{\text{g}}$. Moreover, denote for $(t,m) \in \mathcal{E}$ by $agent_{r_{tm}}$ the agent that is placed at bus t and updates r_{tm}. Due to the separability of the objective function (5.33a) each agent $agent_{x_i,r_i}$ and $agent_{r_{tm}}$ can update his set of variables completely in parallel.
According to (5.14) the sparsity pattern of $A(x,r)$ in (5.33b) is given by

$$F_{A(x,r)} = \left\{(l,j) \in 2\mathcal{N}_{\text{b}} \times 2\mathcal{N}_{\text{b}} : A_{lj}(x,r) \neq 0 \text{ for some } x \in \mathbb{R}_+^{6|\mathcal{N}_{\text{b}}|} \text{ and } r \in \mathbb{R}^{2|\mathcal{N}_{\text{b}}|+6|\mathcal{E}|}, l \neq j\right\},$$

where $2\mathcal{N}_{\text{b}} = \{1,\ldots,2n_{\text{b}}\}$. Obviously, $F_{A(x,r)}$ is contained in the sparsity pattern F_Y of the matrix

$$\begin{pmatrix}Y^{\text{bus}} & Y^{\text{bus}}\\ Y^{\text{bus}} & Y^{\text{bus}}\end{pmatrix}. \tag{5.35}$$

Moreover, the four cornered $n_{\text{b}} \times n_{\text{b}}$ blocks of $\widetilde{\mathbf{Y}}_i(x_i,r_i) \in \mathbb{R}^{2n_{\text{b}}\times 2n_{\text{b}}}$ in (5.33b) satisfy the first point in Assumptions 5.2.1 for $i \in \mathcal{N}_{\text{b}}$, and if $agent_{r_{tm}}$ is identified with $agent_{x_t,r_t}$ it follows that the four cornered $n_{\text{b}} \times n_{\text{b}}$ blocks of $\widetilde{\mathbf{Y}}_{tm}(r_{tm})$ in (5.33b) satisfy the first point in Assumptions 5.2.1 as well.

Remark 5.3.3.
The graph $G(2\mathcal{N}_b, F_Y)$ can be obtained by the graph $G(\mathcal{N}_b, \mathcal{E})$ that represents the power system if each bus (node) i is duplicated and the duplicate $n_b + i$ is connected to its original, the neighbors of its original, and their duplicates. In other words, identifying bus i with its duplicate $n_b + i$, the sparsity structure of $A(x,r)$ in (5.33b) is contained in the sparsity structure of the admittance matrix Y and if $G(2\mathcal{N}_b, F_Y)$ is not chordal by itself, the chordal extension of $G(2\mathcal{N}_b, F_Y)$ denoted by $G(2\mathcal{N}_b, \overline{F}_Y)$ can be obtained by the chordal extension $G(\mathcal{N}_b, \overline{\mathcal{E}})$ of $G(\mathcal{N}_b, \mathcal{E})$ as follows:
Let $C_1, \ldots, C_p \subseteq \mathcal{N}_b$ be the maximal cliques of $G(\mathcal{N}_b, \overline{\mathcal{E}})$, then $2C_1, \ldots, 2C_p \subseteq 2\mathcal{N}_b$ are maximal cliques of $G(2\mathcal{N}_b, \overline{F}_Y)$, where $2C_s = \{i,\ n_b + i \mid i \in C_s\}$.

Finally, the range-space conversion method described in section 5.2 can be applied to rewrite constraint (5.33b). To be able to apply the convergence results of the previous sections, we define for $i \in \mathcal{N}_\mathrm{g}$ the compact sets $X_i = X_i^1 \times \ldots \times X_i^6$, where $X_i^j \subset \mathbb{R}_+$. Furthermore, we denote for $i \in \mathcal{N}_\mathrm{g}$ and $(t,m) \in \mathcal{E}$ by $\mathcal{R}_i$ and $\mathcal{R}_{tm}$ the following compact sets

$$\mathcal{R}_i = \left\{ r \in \mathbb{S}_+^4 \ : \ r \preceq I_4 R_{r_i} \wedge r_{11} = 1 \right\} \text{ and } \mathcal{R}_{tm} = \left\{ r \in \mathbb{S}_+^6 \ : \ r \preceq I_6 R_{r_{tm}} \right\},$$

where $R_{r_i} > 0$ and $R_{r_{tm}} > 0$. For $i \in \mathcal{N}_\mathrm{b} \setminus \mathcal{N}_\mathrm{g}$ we set $\mathcal{R}_i = \{0\} \subset \mathbb{S}_+^4$. Finally, we denote by $\mathcal{W}^s$ the compact set that is defined according to (5.17), and assume that $\mathcal{W}^s$, X_i, $\mathcal{R}_i$, and $\mathcal{R}_{tm}$ contain an optimal solution of problem (5.33).
With the range-space conversion method, problem (5.33) can be stated as

$$\max_{\substack{x_i \in X_i,\ r_i \in \mathcal{R}_i \\ r_{tm} \in \mathcal{R}_{tm},\ W^s \in \mathcal{W}^s}} \sum_{i \in \mathcal{N}_\mathrm{b}} \Phi_i(x_i, r_i) - \sum_{(t,m) \in \mathcal{E}} \Phi_{tm}(r_{tm}) \tag{5.36a}$$

$$\text{s.t. } E_{lj} \bullet \sum_{s \in \Gamma(l,j)} W^s - E_{lj} \bullet A(x,r) = 0 \text{ for } (l,j) \in J, \tag{5.36b}$$

where $J = \bigcup_{s=1}^{p} J(2C_s)$, $\Gamma(l,j) = \{s : \ l \in 2C_s,\ j \in 2C_s\} \, \forall \ (l,j) \in J$, and E_{lj} is defined as in section 5.2 with $m = 2n_\mathrm{b}$.

The Lagrangian relaxation of problem (5.36) with respect to constraint (5.36b) yields

$$\begin{aligned}\mathcal{L}(x,r,W,\Lambda) &= \sum_{i\in\mathcal{N}_\mathrm{b}} \Phi_i(x_i,r_i) - \sum_{(t,m)\in\mathcal{E}} \Phi_{tm}(r_{tm}) \\ &\quad + \sum_{(l,j)\in J} \Lambda_{lj} \left(E_{lj} \bullet \sum_{s\in\Gamma(l,j)} W^s - E_{lj} \bullet \left(\sum_{i\in\mathcal{N}_\mathrm{b}} \widetilde{\mathbf{Y}}_i(x_i,r_i) + \sum_{(t,m)\in\mathcal{E}} \widetilde{\mathbf{Y}}_{tm}(r_{tm}) \right)\right) \\ &= \sum_{i\in\mathcal{N}_\mathrm{b}} \left[\Phi_i(x_i,r_i) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \widetilde{\mathbf{Y}}_i(x_i,r_i) \right] \\ &\quad + \sum_{(t,m)\in\mathcal{E}} \left[-\Phi_{tm}(r_{tm}) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \widetilde{\mathbf{Y}}_{tm}(r_{tm}) \right] \\ &\quad + \sum_{s=1}^{p} \sum_{(l,j)\in J(2C_s)} \Lambda_{lj} E_{lj} \bullet W^s. \end{aligned} \tag{5.37}$$

Smoothing the Lagrangian with scaled prox-functions d_{x_i,r_i}, $d_{r_{tm}}$, and d_{W^s}, where possible choices are $d_{W^s}(W^s) = (\sigma_{W^s}/2)\|W^s\|_F^2$, $d_{x_i,r_i}(x_i,r_i) = (\sigma_{x_i,r_i}/2)\,\|x_i\|^2 + (\sigma_{x_i,r_i}/2)\,\|r_i\|_F^2$, and $d_{r_{tm}}(r_{tm}) = (\sigma_{r_{tm}}/2)\,\|r_{tm}\|_F^2$, yields the smooth and concave augmented dual function

$$\begin{aligned} f_c(\Lambda) &= \sum_{i\in\mathcal{N}_\mathrm{b}} \max_{x_i\in X_i,\; r_i\in\mathcal{R}_i} \left\{ \Phi_i(x_i,r_i) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \widetilde{\mathbf{Y}}_i(x_i,r_i) - c d_{x_i,r_i}(x_i,r_i) \right\} \\ &\quad + \sum_{(t,m)\in\mathcal{E}} \max_{r_{tm}\in\mathcal{R}_{tm}} \left\{ -\Phi_{tm}(r_{tm}) - \sum_{(l,j)\in J} \Lambda_{lj} E_{lj} \bullet \widetilde{\mathbf{Y}}_{tm}(r_{tm}) - c d_{r_{tm}}(r_{tm}) \right\} \\ &\quad + \sum_{s=1}^{p} \max_{W^s\in\mathcal{W}^s} \left\{ \sum_{(l,j)\in J(2C_s)} \Lambda_{lj} E_{lj} \bullet W^s - c d_{W^s}(W^s) \right\}. \end{aligned} \tag{5.38}$$

The problems in the right-hand side of the above equation are separable with respect to each decision variable due to the separability of the objective function and $A(x,r)$ in (5.36b), however, for the sake of compact notation, we consider here the blocks (x_i,r_i), r_{tm}, and W^s.

With Theorem 3.1.1 the partial derivatives of f_c are given for $(l,j) \in J$ by

$$\nabla_{lj} f_c(\Lambda) = E_{lj} \bullet \left(\sum_{s\in\Gamma(l,j)} W^s(\Lambda) - \sum_{i\in\mathcal{N}_\mathrm{b}} \widetilde{\mathbf{Y}}_i(x_i(\Lambda),r_i(\Lambda)) - \sum_{(t,m)\in\mathcal{E}} \widetilde{\mathbf{Y}}_{tm}(r_{tm}(\Lambda)) \right), \tag{5.39}$$

where $(x_i(\Lambda),r_i(\Lambda))$, $r_{tm}(\Lambda)$, and $W^s(\Lambda)$ solve the right-hand side of (5.38). To determine the Lipschitz constant, the constraints in (5.36b) have to be sorted by the primal blocks (x_i,r_i), r_{tm}, and W^s to identify the corresponding constraint coefficient matrices for $(l,j) \in$

J as follows:

$$\begin{aligned}
&E_{lj} \bullet \left(\sum_{s \in \Gamma(l,j)} W^s - A(x,r) \right) \\
=&E_{lj} \bullet \left(\sum_{s \in \Gamma(l,j)} W^s - \sum_{i \in \mathcal{N}_\mathrm{b}} \left[\boldsymbol{\lambda}_i \mathbf{Y}_i + \bar{\boldsymbol{\lambda}}_i \bar{\mathbf{Y}}_i + \boldsymbol{\mu}_i M_i \right] - \sum_{(t,m) \in \mathcal{E}} \left[2r_{tm}^2 \mathbf{Y}_{tm} + 2r_{tm}^3 \bar{\mathbf{Y}}_{tm} \right] \right) \\
=&E_{lj} \bullet \left(\sum_{s \in \Gamma(l,j)} W^s - \sum_{i \in \mathcal{N}_\mathrm{b}} \left[\left(-\lambda_i^{\min} + \lambda_i^{\max} + a_{i1} + 2\sqrt{a_{i2}} r_i^1 \right) \mathbf{Y}_i \right.\right. \\
&+ \left.\left. \left(-\bar{\lambda}_i^{\min} + \bar{\lambda}_i^{\max} \right) \bar{\mathbf{Y}}_i + \left(-\mu_i^{\min} + \mu_i^{\max} \right) M_i \right] - \sum_{(t,m) \in \mathcal{E}} \left[2r_{tm}^2 \mathbf{Y}_{tm} + 2r_{tm}^3 \bar{\mathbf{Y}}_{tm} \right] \right) \\
=&E_{lj} \bullet \left(\sum_{s \in \Gamma(l,j)} W^s + \sum_{i \in \mathcal{N}_\mathrm{b}} \left[\mathbf{Y}_i \lambda_i^{\min} - \mathbf{Y}_i \lambda_i^{\max} + \bar{\mathbf{Y}}_i \bar{\lambda}_i^{\min} - \bar{\mathbf{Y}}_i \bar{\lambda}_i^{\max} \right.\right. \\
&+ \left.\left. M_i \mu_i^{\min} - M_i \mu_i^{\max} - \mathbf{Y}_i 2\sqrt{a_{i2}} r_i^1 - a_{i1} \mathbf{Y}_i \right] + \sum_{(t,m) \in \mathcal{E}} \left[-\mathbf{Y}_{tm} 2r_{tm}^2 - \bar{\mathbf{Y}}_{tm} 2r_{tm}^3 \right] \right) \\
=&E_{lj} \bullet \sum_{s \in \Gamma(l,j)} W^s + \sum_{i \in \mathcal{N}_\mathrm{b}} \left\langle E_{lj} \bullet \left(\mathbf{Y}_i,\, -\mathbf{Y}_i,\, \bar{\mathbf{Y}}_i,\, -\bar{\mathbf{Y}}_i,\, M_i,\, -M_i,\, -\mathbf{Y}_i 2\sqrt{a_{i2}} \right), \right. \\
&\left. \left(\lambda_i^{\min},\, \lambda_i^{\max},\, \bar{\lambda}_i^{\min},\, \bar{\lambda}_i^{\max},\, \mu_i^{\min},\, \mu_i^{\max},\, r_i^1 \right) \right\rangle - E_{lj} \bullet \sum_{i \in \mathcal{N}_\mathrm{b}} a_{i1} \mathbf{Y}_i \\
&+ \sum_{(t,m) \in \mathcal{E}} \left\langle E_{lj} \bullet \left(-\mathbf{Y}_{tm} 2,\, -\bar{\mathbf{Y}}_{tm} 2 \right), \left(r_{tm}^2,\, r_{tm}^3 \right) \right\rangle \\
=&E_{lj} \bullet \sum_{s \in \Gamma(l,j)} W^s + \sum_{i \in \mathcal{N}_\mathrm{b}} \left\langle E_{lj} \bullet \left(\mathbf{Y}_i,\, -\mathbf{Y}_i,\, \bar{\mathbf{Y}}_i,\, -\bar{\mathbf{Y}}_i\,, M_i, -M_i,\, 0,\, -\mathbf{Y}_i \sqrt{a_{i2}},\, -\mathbf{Y}_i \sqrt{a_{i2}},\, 0 \right), \right. \\
&\left. \left(\lambda_i^{\min},\, \lambda_i^{\max},\, \bar{\lambda}_i^{\min},\, \bar{\lambda}_i^{\max},\, \mu_i^{\min},\, \mu_i^{\max},\, 1,\, r_i^1,\, r_i^1,\, r_i^2 \right) \right\rangle \\
&- E_{lj} \bullet \sum_{i \in \mathcal{N}_\mathrm{b}} a_{i1} \mathbf{Y}_i + \sum_{(t,m) \in \mathcal{E}} \left\langle E_{lj} \bullet \left(0,\, -\mathbf{Y}_{tm},\, -\bar{\mathbf{Y}}_{tm},\, -\mathbf{Y}_{tm},\, 0,\, 0,\, -\bar{\mathbf{Y}}_{tm},\, 0,\, 0 \right), \right. \\
&\left. \left(r_{tm}^1,\, r_{tm}^2,\, r_{tm}^3,\, r_{tm}^2,\, r_{tm}^4,\, r_{tm}^5,\, r_{tm}^3,\, r_{tm}^5,\, r_{tm}^6 \right) \right\rangle .
\end{aligned}$$

Define by $E_{W_s} \in \mathbb{R}^{|J| \times n_\mathrm{b}{}^2}$ for $s = 1, \ldots, p$ the constraint coefficient matrix of $\mathrm{v}(W_s)$ as in (5.22), i.e.,

$$(E_{W_s})_{ik} = \begin{cases} \left(\mathrm{v}(E_{lj})^T \right)_{ik} & \text{if } i = (l,j) \in J(2C_s), \\ 0 & \text{else }, \end{cases}$$

and denote by $E_{x_i, r_i} \in \mathbb{R}^{J \times 10}$ for $i \in \mathcal{N}_\mathrm{b}$ the constraint coefficient matrix of $(x_i, \mathrm{v}(r_i))$ with

$$(E_{x_i, r_i})_{lj} = E_{lj} \bullet \left(\mathbf{Y}_i,\, -\mathbf{Y}_i,\, \bar{\mathbf{Y}}_i,\, -\bar{\mathbf{Y}}_i,\, M_i,\, -M_i,\, 0,\, -\mathbf{Y}_i \sqrt{a_{i2}},\, -\mathbf{Y}_i \sqrt{a_{i2}},\, 0 \right) \ \text{ for } \ (l,j) \in J.$$

Moreover, let $E_{r_{tm}} \in \mathbb{R}^{J\times 9}$ for $(t,m) \in \mathcal{E}$ be the constraint coefficient matrix of $\mathrm{v}(r_{tm})$ with

$$(E_{r_{tm}})_{lj} = E_{lj} \bullet \big(0,\ -\mathbf{Y}_{tm},\ -\bar{\mathbf{Y}}_{tm},\ -\mathbf{Y}_{tm},\ 0,\ 0,\ -\bar{\mathbf{Y}}_{tm},\ 0,\ 0\big) \quad \text{for} \quad (l,j) \in J.$$

Then it follows according to Theorem 3.1.1 that the Lipschitz constant of the gradient of $f_c(\Lambda)$ is given by

$$L_c = \sum_{s=1}^{p} \frac{\|E_{W^s}\|^2}{c\sigma_{W^s}} + \sum_{i\in\mathcal{N}_\mathrm{b}} \frac{\|E_{x_i,r_i}\|^2}{c\sigma_{x_i,r_i}} + \sum_{(t,m)\in\mathcal{E}} \frac{\|E_{r_{tm}}\|^2}{c\sigma_{r_{tm}}}, \tag{5.40}$$

as described detailed in section 5.2.

To state the DAPCA-EC, let $Q = Q_{11} \times \ldots \times Q_{2n_\mathrm{b}2n_\mathrm{b}} \subset \mathbb{R}^{|J|}$ be a compact and convex set that contains an optimal dual multiplier Λ^{opt} of the dual problem of problem (5.33), and consider the augmented dual problem

$$\max_{\Lambda\in Q} f_c(\Lambda).$$

Choose $\gamma > 1$, $L_{-1} \in (0, L_c]$, and the starting point $\bar{\Lambda}^0 = \Lambda^0$ as the minimum of the separable prox-function $d(\Lambda) = \sum_{(l,j)\in J} d_{lj}(\Lambda_{lj})$, where $d_{lj}\colon Q_{lj} \to \mathbb{R}_+$ is an arbitrary prox-function with convexity parameter $\sigma > 0$. Moreover, let $\bar{Y}^k$ denote the outdated vector defined by (3.2.1) and set $\bar{Y}^{-1} = \bar{\Lambda}^0$.

Algorithm 5.3.4. (DAPCA-EC to solve (5.36)) *For $k \geq 0$ do in parallel:*
For $i \in \mathcal{N}_b$, $(t,m) \in \mathcal{E}$, and $s = 1,\ldots,p$, given the required components of $\bar{\Lambda}^k$, agent$_{x_i,r_i}$, agent$_{r_{tm}}$, and agent$_{W^s}$

1. *compute*

$$\begin{aligned}
(x_i^{k+1}, r_i^{k+1}) = \operatorname*{arg\,max}_{x_i\in X_i,\ r_i\in\mathcal{R}_i} \Bigg\{ &\Phi_i(x_i,r_i) - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \widetilde{\mathbf{Y}}_i(x_i,r_i) \\
&- \frac{c\sigma_{x_i,r_i}}{2}\|x_i\|_2^2 - \frac{\sigma_{x_i,r_i}}{2}\|r_i\|_F^2 \Bigg\},
\end{aligned} \tag{5.41}$$

$$\begin{aligned}
r_{tm}^{k+1} = \operatorname*{arg\,max}_{r_{tm}\in\mathcal{R}_{tm}} \Bigg\{ &-\Phi_{tm}(r_{tm}) - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \widetilde{\mathbf{Y}}_{tm}(r_{tm}) \\
&- \frac{c\sigma_{r_{tm}}}{2}\|r_{tm}\|_F^2 \Bigg\},
\end{aligned} \tag{5.42}$$

$$W^{s,k+1} = \operatorname*{arg\,max}_{W^s\in\mathcal{W}^s} \Bigg\{ \sum_{(l,j)\in J(2C_s)} \bar{\Lambda}_{lj}^k E_{lj} \bullet W^s - \frac{c\sigma_{W^s}}{2}\|W^s\|_F^2 \Bigg\}, \tag{5.43}$$

and send (x_i^{k+1}, r_i^{k+1}), r_{tm}^{k+1}, and $W^{s,k+1}$ to the dual agents that require it.

For $(l,j) \in J$, given the blocks (x_i^{k+1}, r_i^{k+1}), r_{tm}^{k+1}, and $W^{s,k+1}$ that are necessary for the computation of $\nabla_{lj} f_c(\bar{\Lambda})^k)$, agent$_{\Lambda_{lj}}$

2. *computes*

$$\nabla_{lj} f_c(\bar{\Lambda}^k) = \sum_{s\in\Gamma(l,j)} E_{lj} \bullet W^{s,k+1} - \sum_{(l,j)\in J} E_{lj} \bullet \sum_{i\in\mathcal{N}_b} \widetilde{\mathbf{Y}}_i(x_i^{k+1}, r_i^{k+1}) \\ - \sum_{(l,j)\in J} E_{lj} \bullet \sum_{(t,m)\in\mathcal{E}} \widetilde{\mathbf{Y}}_{tm}(r_{tm}^{k+1}).$$

and sets $L_k = L_{k-1}$*,*

3. *finds*

$$Y_{lj}^k = \underset{Y_{lj}\in Q_{lj}}{\arg\min} \left\{ \nabla_{lj} f_c(\bar{\Lambda}^k) Y_{lj} + L_k \Delta_k \left(\eta_{lj} + 1\right) \left| Y_{lj} - \Lambda_{lj}^k \right| + \frac{L_k}{2} \left(Y_{lj} - \Lambda_{lj}^k \right)^2 \right\}.$$

4. ***if*** $L_k < L_c$ ***then***

   (a) *agent*$_{\Lambda_{lj}}$ *sends* $\bar{Y}_{lj}^k$ *to the primal agents that require it if necessary:*
   if $\left|\bar{Y}_{lj}^{k-1} - Y_{lj}^k\right| > \frac{L_k}{L_c}\Delta_k$ ***then***
   agent$_{\Lambda_{lj}}$ *sets* $\bar{Y}_{lj}^k = Y_{lj}^k$ *and sends* $\bar{Y}_{lj}^k$*.*
   else
   agent$_{\Lambda_{lj}}$ *sets* $\bar{Y}_{lj}^k = \bar{Y}_{lj}^{k-1}$ *and signals that no data will be sent.*

For $i \in \mathcal{N}_b$*,* $(t,m) \in \mathcal{E}$*, and* $s = 1,\dots,p$*, given the required components of* $\bar{Y}^k$*, agent*$_{x_i,r_i}$*, agent*$_{r_{tm}}$ *and agent*$_{W^s}$

   (b) *compute*

$$(y_i^{k+1}, q_i^{k+1}) = \underset{x_i\in X_i,\ r_i\in\mathcal{R}_i}{\arg\max} \left\{ \Phi_i(x_i, r_i) - \sum_{(l,j)\in J} \bar{Y}_{lj}^k E_{lj} \bullet \widetilde{\mathbf{Y}}_i(x_i, r_i) \right. \\ \left. - \frac{c\sigma_{x_i,r_i}}{2} \|x_i\|_2^2 - \frac{\sigma_{x_i,r_i}}{2} \|r_i\|_F^2 \right\}, \tag{5.44}$$

$$q_{tm}^{k+1} = \underset{r_{tm}\in\mathcal{R}_{tm}}{\arg\max} \left\{ -\Phi_{tm}(r_{tm}) - \sum_{(l,j)\in J} \bar{Y}_{lj}^k E_{lj} \bullet \widetilde{\mathbf{Y}}_{tm}(r_{tm}) \right. \\ \left. - \frac{c\sigma_{r_{tm}}}{2} \|r_{tm}\|_F^2 \right\}, \tag{5.45}$$

$$V^{s,k+1} = \underset{W^s\in\mathcal{W}^s}{\arg\max} \left\{ \sum_{(l,j)\in J(2C_s)} \bar{Y}_{lj}^k E_{lj} \bullet W^s - \frac{c\sigma_{W^s}}{2} \|W^s\|_F^2 \right\}, \tag{5.46}$$

   and send (y_i^{k+1}, q_i^{k+1})*,* q_{tm}^{k+1}*, and* $V^{s,k+1}$ *to the dual agents that require it.*

For $(l,j) \in J$, given the blocks (y_i^{k+1}, q_i^{k+1}), q_{tm}^{k+1}, and $V^{s,k+1}$ that are necessary for the computation of $\nabla_{lj} f_c(\bar{Y}^k)$, $agent_{\Lambda_{lj}}$

(c) *computes*

$$\nabla_{lj} f_c(\bar{Y}^k) = \sum_{s\in\Gamma(l,j)} E_{lj} \bullet V^{s,k+1} - \sum_{(l,j)\in J} E_{lj} \bullet \sum_{i\in\mathcal{N}_b} \widetilde{\mathbf{Y}}_i(y_i^{k+1}, q_i^{k+1}) - \sum_{(l,j)\in J} E_{lj} \bullet \sum_{(t,m)\in\mathcal{E}} \widetilde{\mathbf{Y}}_{tm}(q_{tm}^{k+1}).$$

and checks with consensus (section 2.4)

if

$$\frac{L_k}{2} \sum_{(l,j)\in J} \left(Y_{lj}^k - \Lambda_{lj}^k\right)^2 \geq \sum_{(l,j)\in J} \left(\nabla_{lj} f(\bar{Y}^k) - \nabla_{lj} f(\bar{\Lambda}^k)\right)\left(Y_{lj}^k - \Lambda_{lj}^k\right)$$

then

continues with step 5,

else

sets $L_k = L_k \gamma$ and goes to step 3,

5. *finds* $Z_{lj}^k = \operatorname*{arg\,min}_{Z_{lj}\in Q_{lj}} \left\{ \frac{L_k}{\sigma} d_{lj}(Z_{lj}) + \sum_{t=0}^{k} \frac{t+1}{2} \nabla_{lj} f_c(\bar{\Lambda}^t) Z_{lj} \right\}$,
6. *sets* $\Lambda_{lj}^{k+1} = \frac{2}{k+3} Z_{lj}^k + \frac{k+1}{k+3} Y_{lj}^k$,
7. *and sends $\bar{\Lambda}_{lj}^{k+1}$ to the primal agents that request it if necessary:*

if $\left|\bar{\Lambda}_{lj}^k - \Lambda_{lj}^{k+1}\right| > \Delta_{k+1}$ ***then***
$agent_{\Lambda_{lj}}$ sets $\bar{\Lambda}_{lj}^{k+1} = \Lambda_{lj}^{k+1}$ and sends $\bar{\Lambda}_{lj}^{k+1}$.
else
$agent_{\Lambda_{lj}}$ sets $\bar{\Lambda}_{lj}^{k+1} = \bar{\Lambda}_{lj}^k$ and signals that no data will be sent.

Regarding our desire to keep the communication local (or close to local) with respect to the branches of the considered power system, we identify $agent_{\Lambda_{lj}}$ for $(l,j) \in J$ with $agent_{x_l, r_l}$ if $l \in \mathcal{N}_\text{b}$ and with $agent_{x_{(l-n_\text{b})}, r_{(l-n_\text{b})}}$ if $l \in 2\mathcal{N}_\text{b} \setminus \mathcal{N}_\text{b}$ and choose the topology of the multi-agent network to coincide with the topology of the considered power system network. This setting is favorable as according to Remark 5.3.3 the index pair $(l,j) \in J$ can be identified with a branch in the power system if it is chordal.
Regarding the communication exchange of the consensus iterations in step 4 c) of Algorithm 5.3.4, we recall that it is determined by a symmetric Matrix A that is compatible to the multi-agent network of $agent_{\Lambda_{11}}, \ldots, agent_{\Lambda_{2n_\text{b}2n_\text{b}}}$, i.e., the information exchange is local in the consensus phase with respect to the branches of the considered power system

network. Moreover, as described above, the four cornered $n_\mathrm{b} \times n_\mathrm{b}$ blocks of $\widetilde{\mathbf{Y}}_i(x_i, r_i)$ and $\widetilde{\mathbf{Y}}_{tm}(r_{tm})$ in (5.33b) satisfy the first point in Assumptions 5.2.1 for the chosen multi-agent network, and with Remark 5.3.3 it follows that the overall information exchange in Algorithm 5.3.4 can be kept local with respect to the branches of the power system network if it is chordal as detailed in section 5.2, or close to local (cf. Remark 5.2.2 and Remark 5.2.4) if the power system is not chordal.

Finally, the analytical solutions of the subproblems in steps 1, 3, 4 *b*), and 5 of Algorithm 5.3.4 can be determined according to Example 2.2.6 and the description in section 5.2, where the analytical solution for $W^{s,k+1}$ (and accordingly for $V^{s,k+1}$) is given. The same approach can be applied to determine r_{tm}^{k+1} (and q_{tm}^{k+1}). However, the analytical solution for (x_i^{k+1}, r_i^{k+1}) (and (y_i^{k+1}, q_i^{k+1})) is more difficult to obtain and will be derived in the following. To this end, we decompose the right-hand side of (5.41) into smaller problems that can be solved in parallel and obtain

$$
\begin{aligned}
(x_i^{k+1}, r_i^{k+1}) &= \underset{x_i \in X_i,\ r_i \in \mathcal{R}_i}{\arg\max} \left\{ \Phi_i(x_i, r_i) - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \widetilde{\mathbf{Y}}_i(x_i, r_i) - \frac{\sigma_{x_i,r_i}}{2}\|x_i\|_2^2 - \frac{\sigma_{x_i,r_i}}{2}\|r_i\|_F^2 \right\} \\
&= \underset{x_i \in X_i,\ r_i \in \mathcal{R}_i}{\arg\max} \Big\{ \lambda_i^{\min} P_i^{\min} - \lambda_i^{\max} P_i^{\max} + \left(-\lambda_i^{\min} + \lambda_i^{\max} + a_{i1} + 2\sqrt{a_{i2}} r_i^1 \right) P_i^{\mathrm{d}} \\
&\quad + \bar{\lambda}_i^{\min} Q_i^{\min} - \bar{\lambda}_i^{\max} Q_i^{\max} + \left(-\bar{\lambda}_i^{\min} + \bar{\lambda}_i^{\max} \right) Q_i^{\mathrm{d}} + \mu_i^{\min} {V_i^{\min}}^2 - \mu_i^{\max} {V_i^{\max}}^2 \\
&\quad + a_{i0} - r_i^2 - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \Big[\left(-\lambda_i^{\min} + \lambda_i^{\max} + a_{i1} + 2\sqrt{a_{i2}} r_i^1 \right) \mathbf{Y}_i \\
&\quad + \left(-\bar{\lambda}_i^{\min} + \bar{\lambda}_i^{\max} \right) \bar{\mathbf{Y}}_i + \left(-\mu_i^{\min} + \mu_i^{\max} \right) M_i \Big] - \frac{\sigma_{x_i,r_i}}{2}\|x_i\|_2^2 - \frac{\sigma_{x_i,r_i}}{2}\|r_i\|_F^2 \Big\}.
\end{aligned}
$$

Further decomposition yields

$$
\begin{aligned}
{\lambda_i^{\min}}^{k+1} &= \underset{\lambda_i^{\min} \in X_i^1}{\arg\max} \left\{ \left(P_i^{\min} - P_i^{\mathrm{d}} + \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \mathbf{Y}_i \right) \lambda_i^{\min} - \frac{c\sigma_{x_i,r_i}}{2} (\lambda_i^{\min})^2 \right\}, \\
{\lambda_i^{\max}}^{k+1} &= \underset{\lambda_i^{\max} \in X_i^2}{\arg\max} \left\{ \left(-P_i^{\max} + P_i^{\mathrm{d}} - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \mathbf{Y}_i \right) \lambda_i^{\max} - \frac{c\sigma_{x_i,r_i}}{2} (\lambda_i^{\max})^2 \right\}, \\
{\bar{\lambda}_i^{\min}}^{k+1} &= \underset{\bar{\lambda}_i^{\min} \in X_i^3}{\arg\max} \left\{ \left(Q_i^{\min} - Q_i^{\mathrm{d}} + \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \bar{\mathbf{Y}}_i \right) \bar{\lambda}_i^{\min} - \frac{c\sigma_{x_i,r_i}}{2} (\bar{\lambda}_i^{\min})^2 \right\}, \\
{\bar{\lambda}_i^{\max}}^{k+1} &= \underset{\bar{\lambda}_i^{\max} \in X_i^4}{\arg\max} \left\{ \left(-Q_i^{\max} + Q_i^{\mathrm{d}} - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet \bar{\mathbf{Y}}_i \right) \bar{\lambda}_i^{\max} - \frac{c\sigma_{x_i,r_i}}{2} (\bar{\lambda}_i^{\max})^2 \right\}, \\
{\mu_i^{\min}}^{k+1} &= \underset{\mu_i^{\min} \in X_i^5}{\arg\max} \left\{ \left({V_i^{\min}}^2 + \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet M_i \right) \mu_i^{\min} - \frac{c\sigma_{x_i,r_i}}{2} (\mu_i^{\min})^2 \right\},
\end{aligned}
$$

$$\mu_i^{\max^{k+1}} = \underset{\mu_i^{\max} \in X_i^6}{\arg\max} \left\{ \left(-V_i^{\max 2} - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet M_i \right) \mu_i^{\max} - \frac{c\sigma_{x_i,r_i}}{2} (\mu_i^{\max})^2 \right\},$$

$$r_i^{k+1} = \underset{r_i \in \mathcal{R}_i}{\arg\max} \left\{ \left(2\sqrt{a_{i2}} P_i^{\mathrm{d}} - \sum_{(l,j)\in J} \bar{\Lambda}_{lj}^k E_{lj} \bullet 2\sqrt{a_{i2}} \mathbf{Y}_i \right) r_i^1 - r_i^2 - \frac{c\sigma_{x_i,r_i}}{2} \|r_i\|_F^2 \right\}, \quad (5.47)$$

where

$$\mathcal{R}_i = \left\{ r \in \mathbb{S}_+^4 \;:\; r \preceq I_4 R_{r_i} \wedge r_{11} = 1 \right\}.$$

As it is straightforward to determine analytical solutions for the iterates $\lambda_i^{\min^{k+1}}, \ldots, \mu_i^{\max^{k+1}}$, we will only provide the analytical solution for r_i^{k+1} in the following:

Lemma 5.3.5.
Subproblem (5.47) has the optimal solution

$$r_i^{1,k+1} = \min \left(\sqrt{R_{r_i} - 1}, \sqrt[3]{\frac{\tilde{a}}{4c\sigma_{x_i,r_i}} + \sqrt{D}} + \sqrt[3]{\frac{\tilde{a}}{4c\sigma_{x_i,r_i}} - \sqrt{D}} \right) \;\; \text{and} \;\; r_i^{2,k+1} = \left(r_i^{1,k+1} \right)^2,$$

where

$$\tilde{a} = \left(2\sqrt{a_{i2}} P_i^d - \sum_{(l,j)\in J} \bar{\Lambda}_{lj} E_{lj}^k \bullet 2\sqrt{a_{i2}} \mathbf{Y}_i \right) \;\; \text{and} \;\; D = \left(-\frac{\tilde{a}}{4c\sigma_{x_i,r_i}} \right)^2 + \left(\frac{c\sigma_{x_i,r_i} + 1}{3c\sigma_{x_i,r_i}} \right)^3.$$

Proof. The application of the Schur complement in combination with [Zha05, Theo. 1.12] yields that (5.47) can be rewritten with $x = r_i^1$ and $y = r_i^2$ as follows if $R_{r_i} > 1$:

$$r_i^{k+1} = \underset{x,y}{\arg\max} \left\{ \tilde{a}x - y - \frac{c\sigma_{x_i,r_i}}{2} y^2 - c\sigma_{x_i,r_i} x^2 \right\} \quad (5.48)$$

$$\text{s.t. } y - x^2 \geq 0,$$

$$R_{r_i} - y - \frac{x^2}{R_{r_i} - 1} \geq 0.$$

Let $(x^{\text{opt}}, y^{\text{opt}})$ be the optimal solution of (5.48) and assume that $y^{\text{opt}} > (x^{\text{opt}})^2$. It follows that

$$\begin{aligned} &\tilde{a}x^{\text{opt}} - c\sigma_{x_i,r_i} \left(x^{\text{opt}}\right)^2 - y^{\text{opt}} - \frac{c\sigma_{x_i,r_i}}{2} \left(y^{\text{opt}}\right)^2 \\ <&\tilde{a}x^{\text{opt}} - c\sigma_{x_i,r_i} \left(x^{\text{opt}}\right)^2 - \left(x^{\text{opt}}\right)^2 - \frac{c\sigma_{x_i,r_i}}{2} \left(x^{\text{opt}}\right)^4. \end{aligned}$$

Moreover, we have

$$\begin{aligned} \left(x^{\text{opt}}\right)^2 - \left(x^{\text{opt}}\right)^2 &\geq 0, \\ R_{r_i} - \left(x^{\text{opt}}\right)^2 - \frac{\left(x^{\text{opt}}\right)^2}{R_{r_i} - 1} \geq R_{r_i} - y^{\text{opt}} - \frac{\left(x^{\text{opt}}\right)^2}{R_{r_i} - 1} &\geq 0. \end{aligned}$$

This contradicts that $(x^{\text{opt}}, y^{\text{opt}})$ is an optimal solution of (5.48). It follows that we can consider instead of (5.48) the equivalent problem

$$\arg\max_{x} \left\{ \tilde{a}x - (c\sigma_{x_i,r_i} + 1)\, x^2 - \frac{c\sigma_{x_i,r_i}}{2} x^4 \right\} \tag{5.49}$$

$$\text{s.t. } x \leq \sqrt{R_{r_i} - 1}, \tag{5.50}$$

where (5.50) is obtained by

$$R_{r_i} - x^2 - \frac{x^2}{R_{r_i} - 1} = R_{r_i} - x^2 \left(\frac{R_{r_i}}{R_{r_i} - 1} \right). \tag{5.51}$$

As the objective function of (5.49) is strictly concave, the unique optimal solution x^{opt} can be found by determining the root of the cubic polynomial

$$f(x) = x^3 + \underbrace{\left(\frac{c\sigma_{x_i,r_i} + 1}{c\sigma_{x_i,r_i}} \right)}_{=p} x + \underbrace{\left(-\frac{\tilde{a}}{2c\sigma_{x_i,r_i}} \right)}_{=q} = 0. \tag{5.52}$$

To this end Cardano's method can be applied [Fis13, sec. 5.2]: As $f(x)$ has a negative discriminant

$$\Delta(f) = -(4p^3 + 27q^2),$$

equation (5.52) has exactly one real solution and the optimum x^{opt} is obtained by

$$x^{\text{opt}} = \min \left(\sqrt{R_{r_i} - 1}, \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \right).$$

□

Finally, we sketch how step 2 and step 3 of Algorithm 5.3.2 can be implemented in parallel by using only neighborhood information.

Starting with step 2, the following optimization problem can be considered in order to distributedly find a nonzero vector $(v_1^T, v_2^T)^T \neq 0$ in the kernel of $A(x^{\text{opt}}, r^{\text{opt}})$ (Michael Ulbrich, personal communication, July 16, 2013):

$$\min_{v \in \mathbb{R}^{2n_b}} q^T v + \frac{\sigma_v}{2} \|v\|^2 \text{ s.t. } A(x^{\text{opt}}, r^{\text{opt}})v = 0, \tag{5.53}$$

where the optimal solution v^{opt} of (5.53) satisfies $v^{\text{opt}} \neq 0$ if $q \notin range(A(x^{\text{opt}}, r^{\text{opt}}))$.

Moreover, problem (5.53) is convex with a separable objective function and coupled linear constraints, i.e., it can be solved distributedly with event-triggered and local communication by the application of the DAPCA-EC 3.2.2. Here, the strongly convexity of the

objective function of (5.53) makes a smoothing of the dual function unnecessary. Therefore, it is possible to allow $v \in \mathbb{R}^{2n_\mathrm{b}}$ in (5.53) as a straightforward revision of the corresponding convergence proofs shows. Moreover, the application of the DAPCA-EC 3.2.2 ensures that the variables v_{1i} and v_{2i} are only known at bus $i \in \mathcal{N}_\mathrm{b}$.
According to [LL10, LL12], the scalars $\zeta_1, \zeta_2 \in \mathbb{R}$ in step 3 of Algorithm 5.3.2, that satisfy $V^\mathrm{opt} = (\zeta_1 + \mathrm{j}\zeta_2)(v_1 + \mathrm{j}v_2)$, can be found by solving two equations of the abovementioned system

$$\begin{pmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\{V^\mathrm{opt}\} \\ \mathrm{Im}\{V^\mathrm{opt}\} \end{pmatrix}.$$

Following [LL10, LL12] (and Javad Lavaei, personal communication, August 1, 2013), these equations can be obtained by letting the swing bus $i \in \mathcal{N}_\mathrm{b}$ coincide with a bus, where the voltage constraint is active in the optimum, and an agent placed at the corresponding bus i can determine ζ_1 and ζ_2 by solving

$$\begin{pmatrix} v_{1i} & -v_{2i} \\ v_{2i} & v_{1i} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} = \begin{pmatrix} V_i^\mathrm{min}(\text{ or } V_i^\mathrm{max}) \\ 0 \end{pmatrix}.$$

After that ζ_1 and ζ_2 are spread through the network. In contrast to a centralized approach, where V^opt is determined by a single entity and then spread through the network, the value of v_{1i} and v_{2i} is only known at the corresponding bus $i \in \mathcal{N}_\mathrm{b}$ in a distributed implementation of Algorithm 5.3.2.
Finally, $P_i^{\mathrm{g}^\mathrm{opt}}$ and $Q_i^{\mathrm{g}^\mathrm{opt}}$ for $i \in \mathcal{N}_\mathrm{g}$ can be computed locally with each agent using neighborhood information according to the power balance equation (4.15b) as
$V_i I_i^* = V_i \sum_{j \in N(i)} Y_{ij}^{\mathrm{bus}^*} V_j^*$, where $N(i)$ is the set of buses that are connected to bus i by a branch.

Remark 5.3.6.
In [LZT12, DZG13], OPF problems are solved distributedly by the application of semidefinite matrix completion as well. In this works, however, a matrix completion technique (which differs from the range-space conversion method for the decomposition of an LMI) is used to decompose a matrix variable of an SDP that relaxes the OPF problem by neglecting a rank-1 constraint and for which equivalence of the optimal solutions can be shown if the considered network has a special structure (e.g., tree or lossless cycle [LZT12]). In [LZT12], two algorithms (primal and dual) are proposed that use a (sub)gradient scheme, whereas in [DZG13] the alternating direction method of multipliers is used. Similarly, the domain-space conversion method [KKMY11] is applied in [Jab12] for the decomposition of a matrix variable to reduce the computation time of a primal-dual interior-point solver that is applied to solve an OPF relaxation in semidefinite form.

6 Numerical results

This chapter follows our numerical investigation of the DPCA-EC applied to the DC-OPF problem in [MUA14, sec. 4] (Meinel, Ulbrich, and Albrecht) and significantly extends it. Moreover, the results that are related to the AC-OPF problem are in preparation for publication in [MU14] (Meinel and Ulbrich).

In this chapter, the numerical results of the DAPCA-EC applied to the DC-OPF problem (5.1) and the dual of the AC-OPF problem (5.32) are presented. More precisely, the IEEE benchmark test cases are considered for the analysis of the DPCA-EC (DAPCA-EC 3.2.2 with $L_{-1} = L_c$) applied to the DC-OPF problem. The IEEE test cases are archived at [Uni] and represent portions of the American Electric Power System in the Midwestern US, where here the portions with 14, 30, and 57 buses are considered. The data for these test cases were obtained in this work with MATPOWER [ZMST11] which is a free scientific tool for power flow analysis in MATLAB. (Moreover, Javad Lavaei was so friendly to send us his model data for the IEEE 57 bus test case (related to [LL12]) for comparison, personal communication, August 26, 2013). Partly, the data needed to be converted to per unit, see for instance [ST06, sec. 3.1] and [BV00, sec. 5.5].

The implementation of the different versions of the DAPCA-EC was done in this work with MATLAB R2014a [Mat14]. Additionally, the modeling toolbox YALMIP [LÖ4] for optimization problems together with the solver $SDPT3$ [TTT99] were used to compute the optimal solutions to the DC-OPF problems and the duals of the AC-OPF problems, as well as the corresponding optimal dual multipliers as reference values for the determination of the primal gap and the constraint violation at the approximate solutions obtained with the DAPCA-EC. $SDPT3$ is a semidefinite programming solver that is recommended in the YALMIP Wiki. Finally, concerning the decomposition of the dual of the AC-OPF problem as described in section 5.3 , the chordal extension of the graph that represents the sparsity structure of LMI (5.33b) was determined according to Remark 5.2.2 with the MATLAB functions *chol.m* and *amd.m* to obtain a chordal extension with minimal additional edges. Moreover, the maximal cliques of the chordalized graph were computed with the MATLAB function *maximalCliques* provided by Jeffrey Wildman.

In the following sections, the subsequent steps are carried out for each test case to investigate the DAPCA-EC.
In the first step, the PCA 3.1.2 is compared with the DPCA-EC to find out to what extent the communication exchange can be reduced by the usage of event-triggered communication for a pre-given number of iterations according to Theorem 3.1.5. Moreover, the comparison is repeated with a stopping-criterion for the primal gap and the constraint violation, to firstly investigate the tightness of the pre-given number of iterations in Theorem 3.1.5, and to secondly find out if there is a trade-off between the communication savings due to the usage of event-triggered communication and the necessary number of iterations to obtain a desired accuracy.
In the second step, the same stopping criterion is used to investigate how the adaptive step-size strategy in the DAPCA-EC helps to reduce the number of iterations compared to the DPCA-EC. Furthermore, the impact of event-triggered communication in combination with the adaptive step-size strategy is studied. Finally, as the application of the consensus technique in step 4 $c)$ of the DAPCA-EC 3.2.2 can be a bottleneck regarding the computation time, it is examined if the algorithm still converges if the L_k-updates in step 4 $c)$ are not allowed in every iteration.
The numerical results are presented in the following sections with numerous tables to provide a compact overview. To this end, the following abbreviations are used:

PG : primal gap at approximate solution,

CV : constraint violation at approximate solution,

NoI : number of iterations of the DAPCA-EC,

NoCI : number of consensus iterations of the DAPCA-EC,

TC : total communication (number of exchanged primal and dual iterates plus exchanged iterates in the consensus phase if the DAPCA-EC is considered),

DC : dual communication (number of exchanged dual iterates),

CC : consensus communication (number of exchanged iterates in the consensus phase),

L_k**-Up** : number of L_k-updates in the DAPCA-EC,

$L_k^{\mathbf{max}}$: maximal L_k in the DAPCA-EC,

MCTpA : maximal computation time per agent in seconds.

6.1 Choice of parameters for the DAPCA-EC

For the implementation of the different versions of the DAPCA-EC, the convexity parameters of the prox-functions, that are used to smoothen the dual functions, were chosen optimally as described in section 3.1.1, resulting in a minimal Lipschitz constant and thus reducing the necessary number of iterations according to the convergence results derived in this work. As the convergence result of the PCA in Theorem 3.1.5 builds the basis of comparison in the first step of the numerical investigation, the scaling technique described in section 3.1.1 was used to balance the bounds in Theorem 3.1.5. The scaling factor s and the accuracy ϵ were chosen in a way that the absolute values of the upper and lower bound on the primal gap

$$-\frac{1}{s}\left\|(\mu,\lambda)^{\text{opt}}\right\|\left(\frac{1}{s}\left\|(\mu,\lambda)^{\text{opt}}\right\|+\sqrt{\frac{1}{s^2}\left\|(\mu,\lambda)^{\text{opt}}\right\|^2+2}\right)\epsilon \leq \sum_{i=1}^{n}\Phi_i(\hat{x}_i)-f^{\text{opt}} \leq \epsilon \tag{6.1}$$

are approximately $1/100$ of the absolute value of the primal gap at the starting point which is zero (with appropriate dimension) according to the choice of the prox-function in the initialization of Algorithm 5.1.1 and Algorithm 5.3.4. To give an example, if the primal gap at the starting point is -0.5 the choice $\epsilon = 0.005$ yields an upper bound in (6.1) that is $1/100$ of the absolute value of the primal gap at the starting point. Moreover, s has to be chosen such that the lower bound in (6.1) is approximately -0.005 for $\epsilon = 0.005$.
Regarding the choice of the threshold $\Delta_k = \beta\delta^k$, which describes the extend of event-triggered communication, numerical tests showed that the results are comparable if either both parameters β and δ are varied or only one of them. For the ease of presentation only the parameter β was varied in this work for all test cases and the parameter δ was chosen as

$$\delta^{k_{\text{fin}}/2} = 0.025, \tag{6.2}$$

where k_{fin} is the necessary number of iterations that is given by Theorem 3.1.5 to achieve the required accuracy for given ϵ and s. This choice is done in order to prevent δ^k from getting to small (> 0.025) before half of the necessary iterations are executed. Numerical tests showed that with this choice of δ, the potential of event-triggered communication can be fully unlocked by solely varying the parameter β.

Finally, for the step-size initialization of the DAPCA-EC 3.2.2, the update parameter $\gamma > 1$ and the starting value $L_{-1} \in (0, L_c)$ were chosen as follows, where L_c is the Lipschitz constant of the gradient of the augmented dual function, given by (5.8) concerning the DC-OPF problem, and by (5.40) regarding the dual of the AC-OPF problem:
For a given test case and

$$(\gamma, L_{-1}) \in \{1.01, 1.1, 1.2, 1.5, 2, 3\} \times \left\{10^{-1}, \ldots, 10^{-6}\right\} L_c \tag{6.3}$$

the results of the DAPCA-EC for $\Delta_k = 0$, i.e., without event-triggered communication, were compared with respect to the number of iterations, the number of consensus iterations and the computation time. The range of γ in (6.3) gives the amount by which L_k is raised in step 4 c) of the DAPCA-EC if an update is necessary, i.e., by 1 %,…,300 %. The pair (γ, L_{-1}) with the best results were chosen for further investigation of the DAPCA-EC.

6.2 IEEE 57 bus test case (DC-OPF)

In this section, the numerical results of the application of the DAPCA-EC to the DC-OPF problem (5.1) are exemplarily discussed for the IEEE 57 bus test case [ZMS11, Uni] with 7 generators and 80 branches. The results for the IEEE test cases with 14 and 30 buses from [ZMS11, Uni] can be found in the appendix 7.
Firstly, we compare the results of the PCA 3.1.2 with the results of the DPCA-EC (DAPCA-EC 3.2.2 with $L_{-1} = L_c$) that are given in Table 6.1, to find out to what extent the communication exchange can be reduced by the usage of event-triggered communication.
Regarding the optimal dual multipliers $(\mu, \lambda)^{\text{opt}} \in \mathbb{R}^{57+160}$, we have $\|(\mu, \lambda)^{\text{opt}}\| = 31.4364$ and according to Theorem 3.1.5 combined with the scaling technique described in section 3.1.1, the choice $\epsilon = 0.4$ and $s = 62$ yields the following bounds on the primal gap:

$$\begin{aligned}
-0.4075 = -\frac{1}{s}\left\|(\mu,\lambda)^{\text{opt}}\right\| \left(\frac{1}{s}\left\|(\mu,\lambda)^{\text{opt}}\right\| + \sqrt{\frac{1}{s^2}\left\|(\mu,\lambda)^{\text{opt}}\right\|^2 + 2} \right) \epsilon \\
\leq \sum_{i \in \mathcal{N}_{\text{g}}} C_i\left(\hat{P}_i^{\text{g}}\right) - f_c^{\text{opt}} \leq \\
\epsilon = 0.4,
\end{aligned} \tag{6.4}$$

and the following bound on the constraint violation:

$$\left\| \begin{matrix} B^{\text{bus}}\hat{\theta} - I^{\text{g}}\hat{P}^{\text{g}} + \widetilde{P^{\text{d}}} \\ \begin{bmatrix} W^{\text{inc}}\hat{\theta} - F^{\max} \\ F^{\min} - W^{\text{inc}}\hat{\theta} \end{bmatrix}^+ \end{matrix} \right\| \leq \frac{\epsilon}{s}\left(\frac{1}{s}\left\|(\mu,\lambda)^{\text{opt}}\right\| + \sqrt{\frac{1}{s^2}\left\|(\mu,\lambda)^{\text{opt}}\right\|^2 + 2} \right) = 0.013, \tag{6.5}$$

where

$$\hat{P^{\rm g}} = \sum_{j=0}^{k_{\rm fin}} \frac{2(j+1)}{(k_{\rm fin}+1)(k_{\rm fin}+2)} P^{{\rm g},j+1} \in \mathbb{R}^{n_{\rm g}} \text{ and } \hat{\theta} = \sum_{j=0}^{k_{\rm fin}} \frac{2(j+1)}{(k_{\rm fin}+1)(k_{\rm fin}+2)} \theta^{j+1} \in \mathbb{R}^{n_{\rm b}} \quad (6.6)$$

are the convex sums of the primal iterates and $k_{\rm fin} = 130957$ is the number of necessary iterations given by Theorem 3.1.5. It follows for the choice of δ in the threshold $\Delta_k = \beta\delta^k$ that

$$\delta^{k_{\rm fin}/2} = 0.025$$

is satisfied by $\delta \approx 0.9999$. As described in section 6.1, ϵ and s were chosen such that the absolute values of the bounds on the primal gap in (6.4) are approximately $1/100$ of the primal gap at the starting point $(\mu,\lambda)^0 = 0 \in \mathbb{R}^{57+160}$ which is -41.0067. This value coincides with the negative of the optimal function value $f_c^{\rm opt}$ as $P^{{\rm g},1} = 0 \in \mathbb{R}^{n_{\rm g}}$ and for each $i \in \mathcal{N}_{\rm g}$ one has $a_{i0} = 0$ in the quadratic power generation cost function $C_i(P_i^{\rm g})$ (4.14). Finally, the constraint violation at the starting point is 4.4571.
For a better overview, the above figures are resumed in the following list:

IEEE 57 bus test case (DC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 57 + 7 = 64, \qquad \text{dual: } 57 + 160 = 217,$$

- Accuracy $\epsilon = 0.4$ and scaling factor $s = 62$,
- Norm of optimal dual multipliers: $\|(\mu,\lambda)^{\rm opt}\| = 31.4364$,
- Necessary number of iterations (Theorem 3.1.5): $k_{\rm fin} = 130957$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -0.4075, \qquad \text{upper bound: } 0.4, \quad (6.7)$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$0.013, \quad (6.8)$$

- Primal gap at starting point $(\mu,\lambda)^0 = 0 \in \mathbb{R}^{57+160}$: $-41.0067 = -f_c^{\rm opt}$,
- Constraint violation at starting point: 4.4571,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9999^k$.

The abbreviations used in Table 6.1 and the following tables are given in the introduction of this chapter. The results in Table 6.1 are similar to the results in [MUA14, Table 1], where we, however, considered a different objective function and a simpler model for the DC-OPF problem. However, the impact of even-triggered communication is the same:
In row 1 of Table 6.1, we see the result of the PCA 3.1.2 implemented in a distributed manner without event-triggered communication which means that all iterates are exchanged in every iteration, however, the exchange is local according to the discussion on the communication topology of the multi-agent network in section 5.1. As expected, the primal gap (column 3) and the constraint violation at the approximate solution (column 4) of the PCA satisfy the bounds (6.7) and (6.8) according to Theorem 3.1.5.
In row 2 of Table 6.1, the result of the DPCA-EC is given for the threshold $\Delta_k = 0 \cdot 0.9999^k = 0$ which means that an iterate is sent by a controlling agent only if it differs from the previous iterate. This version of the DPCA-EC coincides with a not naive implementation of the PCA, where an iterate is send only if it provides new information. Obviously, the primal gap and the constraint violation in row 2 are the same as in row 1 as identical information is used in the optimization process, however, without explicitly using event-triggered communication, the saving regarding the total communication is 33 % (column 5) and regarding the dual communication 67 % (column 6). This is due to the fact that we have $\lambda^{\text{opt}} = 0 \in \mathbb{R}^{160}$ and according to the choice of the starting point $(\mu, \lambda)^0$, the corresponding iterates stay zero in the optimization process and do not have to be exchanged.
If event-triggered communication is introduced by choosing $\beta > 0$, the results in row 3 - 7 of Table 6.1 show that the total communication can be reduced by up to 42 % and the dual communication by up to 76 %, still satisfying the bounds on the primal gap (6.7) and the constraint violation (6.8). If β is chosen larger as in row 8 and 9, the communication savings are bigger too, however, the bound on the constraint violation is not satisfied anymore by the approximate solutions of the DPCA-EC.
To see the pure impact of event-triggered communication, the same results are given in Table 6.2 with the difference that the communication savings are considered in relation to the result of the DPCA-EC with $\beta = 0$.
It can be seen by the results in row 2 - 6 of Table 6.2 that the total communication can still be reduced by up to 13 % and the dual communication by up to 27 %, satisfying the bounds on the primal gap (6.7) and the constraint violation (6.8) which is quite a great saving, considering the fact that 160 of 217 dual iterates do not need to be exchanged in every iteration, even if no event-triggered communication is used. Moreover, the maximal computation time that an agent needs is given in the last column of Table 6.2.

	β	PG	CV	TC	DC
1	-	-0.1094	0.0035	1.3e8 (100 %)	6.3e7 (100 %)
2	0	-0.1094	0.0035	8.4e7 (67 %)	2.1e7 (33 %)
3	1e-6	-0.1135	0.0036	8.3e7 (66 %)	2.0e7 (32 %)
4	5e-6	-0.0985	0.0032	8.1e7 (65 %)	1.9e7 (30 %)
5	1e-5	-0.1186	0.0039	8.0e7 (64 %)	1.8e7 (29 %)
6	5e-5	-0.0567	0.0051	7.6e7 (60 %)	1.6e7 (26 %)
7	1e-4	-0.1273	0.0093	7.3e7 (58 %)	1.5e7 (24 %)
8	5e-4	-0.0747	0.0385	6.5e7 (52 %)	1.3e7 (20 %)
9	1e-3	-0.0662	0.0520	6.0e7 (48 %)	1.2e7 (18 %)

Table 6.1: Results of the DPCA-EC – IEEE 57 bus (DC-OPF)

	β	PG	CV	TC	DC	MCTpA
1	0	-0.1094	0.0035	8.4e7 (100 %)	2.1e7 (100 %)	6.9
2	1e-6	-0.1135	0.0036	8.3e7 (99 %)	2.0e7 (96 %)	7.1
3	5e-6	-0.0985	0.0032	8.1e7 (97 %)	1.9e7 (89 %)	7.1
4	1e-5	-0.1186	0.0039	8.0e7 (96 %)	1.8e7 (87 %)	7.1
5	5e-5	-0.0567	0.0051	7.6e7 (90 %)	1.6e7 (77 %)	6.8
6	1e-4	-0.1273	0.0093	7.3e7 (87 %)	1.5e7 (73 %)	6.8
7	5e-4	-0.0747	0.0385	6.5e7 (77 %)	1.3e7 (61 %)	6.7
8	1e-3	-0.0662	0.0520	6.0e7 (72 %)	1.2e7 (55 %)	6.5

Table 6.2: Results of the DPCA-EC – IEEE 57 bus (DC-OPF)

In the next step, we investigate the impact of event-triggered communication if the following stopping criterion is used in the DPCA-EC instead of the pre-given number $k_{\text{fin}} = 130957$ of iterations which may not be necessary to obtain the primal gap of -0.1094 and the constraint violation of 0.0035 in row 1 of Table 6.2.
To this end, let

$$\hat{p^{\text{g}}} = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} p^{\text{g},j+1} \in \mathbb{R}^{n_{\text{g}}} \text{ and } \hat{\theta} = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} \theta^{j+1} \in \mathbb{R}^{n_{\text{b}}}, \tag{6.9}$$

be the approximate solutions after k iterations.

DAPCA-EC stopping criterion for the IEEE 57 bus test case (DC-OPF):

For $k \geq 0$

if

$$|\text{primal gap at (6.9)}| \leq 0.1094 \text{ and constraint violation at (6.9) } \leq 0.0035 \tag{6.10}$$

then

stop.

else

continue.

In Table 6.3, the results of the DPCA-EC with stopping criterion (6.10) are presented. The result in row 1 shows that only 1.2467e5 iterations (column 3) are needed to compute an approximate solution without even-triggered communication which satisfies the same primal gap and constraint violation as the solution in row 1 of Table 6.2. This is not surprising as the application of the (adaptive) Nesterov-Algorithm in the DAPCA-EC does not guarantee monotonically increasing dual function values.

As can be seen in row 2 and 3 of Table 6.3, event-triggered communication does not necessarily result in a higher number of iterations, compared to the result in row 1, if β in the threshold $\Delta_k = \beta\delta^k$ is chosen small enough. Additionally, row 4 - 6 of Table 6.3 show that the information exchange can still be reduced even if the number of iterations is higher compared to the result obtained without event-triggered communication in row 1. However, if β is chosen to large, as in row 7 and 8, the error in the event-triggered communication becomes so big that even more communication is required compared to row 1.

	β	NoI	TC	DC	MCTpA
1	0	1.2467e5	8.0e7 (100 %)	2.0e7 (100 %)	6.7
2	1e-6	1.1907e5	7.5e7 (94 %)	1.8e7 (91 %)	6.5
3	5e-6	1.2299e5	7.6e7 (95 %)	1.7e7 (87 %)	6.6
4	1e-5	1.2640e5	7.7e7 (97 %)	1.7e7 (87 %)	6.8
5	5e-5	1.2965e5	7.5e7 (94 %)	1.6e7 (80 %)	6.8
6	1e-4	1.4384e5	8.1e7 (102 %)	1.7e7 (87 %)	7.6
7	5e-4	1.8622e5	1.0e8 (126 %)	2.2e7 (108 %)	9.6
8	1e-3	2.0254e5	1.1e8 (133 %)	2.3e7 (115 %)	10.4

Table 6.3: Results of DPCA-EC with stopping criterion (6.10) – IEEE 57 bus (DC-OPF)

Finally, the results of the DAPCA-EC 5.1.1 are given in Table 6.4 for the step-size parameters $\gamma = 1.5$ and $L_{-1} = 10^{-4}L_c$ which were selected among the candidates in (6.3) as described in section 6.1. For the IEEE 57 bus DC-OPF problem, the corresponding Lipschitz constant is

$$L_c = 1.7149e9 \tag{6.11}$$

according to (5.9). Moreover, for the computation of the results that are given in Table 6.4, the stopping criterion (6.10) was used to find out to what extent the number of iterations can be reduced by the adaptive step-size strategy in the DAPCA-EC compared to the results of the DPCA-EC in Table 6.3 which are obtained with the same stopping criterion.
Indeed, as can be seen in row 1 of Table 6.4 that only $3.1e4$ iterations are needed to compute a solution without event-triggered communication which is a decrease compared to the number of iterations of the DPCA-EC in row 1 of Table 6.3 by approximately 75 %. In column 5 of the first row of Table 6.4, the maximal L_k is given by $1.1e8$ obtained after 16 L_k-updates (step 4 c) of the DAPCA-EC 5.1.1). In other words, the smallest step-size in the DAPCA-EC, which is the inverse of $L_k^{\max}$, is approximately 15 times bigger than the inverse of the Lipschitz constant L_c (6.11) which is the step-size in the DPCA-EC as described in Remark 2.3.1. However, due to the large number of consensus iterations (column 6), to control the step-size distributedly in the DAPCA-EC 5.1.1, the consensus information exchange (column 9) is large too which results in a total information exchange (column 7) that is bigger compared to the results in Table 6.3. Accordingly, the maximal computation time of an agent (last column) is with 120.8 seconds bigger as well. So, if no event-triggered communication is used, the consensus technique (section 2.4) is clearly a bottle-neck regarding the computation time as well as the amount of information exchange. However, concerning the number of iterations of the DAPCA-EC, the adaptive step-size strategy yields a reduction of up to 75 % which would save 75 % of the computation time as well if the algorithm would be implemented centrally.
Remarkably, the results in row 2 - 8 of Table 6.4 show that the application of event-triggered communication can reduce the consensus communication drastically by up to 99% (row 2). This can be explained by the fact that the usage of outdated dual multipliers in (5.11) makes consensus iterations unnecessary if the right-hand side in (5.11) becomes zero due to $(\bar{u},\bar{h})^k = (\bar{\mu},\bar{\lambda})^k$. As the outdated dual iterates $(\bar{u},\bar{h})^k$ and $(\bar{\mu},\bar{\lambda})^k$ correspond both to the dual multipliers (μ,λ), it follows that equality $(\bar{u},\bar{h})^k = (\bar{\mu},\bar{\lambda})^k$ is satisfied much more often compared to the case, where by $\beta = 0$ no event-triggered communication is used.
However, regarding the maximal computation time, only for the choice $\beta = 1e-6$ a result was obtained in 6.3 seconds which is slightly faster compared to the results obtained with

the DPCA-EC in Table 6.3. If β is chosen larger, as can be seen in row 3 - 8, the number of iterations and the computation times are larger compared to the results in Table 6.3 as well.

To sum it up, even if event-triggered communication in combination with the consensus technique can reduce the total information exchange as well as the maximal computation time by up to 95 % compared to the result computed without event-triggered communication in row 1, the consensus technique is clearly a bottle-neck in the DAPCA-EC. As a result, the application of the DAPCA-EC 5.1.1 yields only for very tight thresholds $\Delta_k = \beta\delta^k$ approximate solutions that are slightly better than the solutions obtained by the DPCA-EC with respect to the amount of information exchange and the computation time. However, the DAPCA-EC needs much less iterations compared to the DPCA-EC which was the main goal of the application of the adaptive step-size strategy.

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	3.1e4	16	1.1e8	1.1e7	1.7e9 (100 %)	1.0e7 (100 %)	1.7e9 (100 %)	120.8
2	1e-6	4.7e4	18	2.5e8	1.5e5	8.2e7 (5 %)	1.4e7 (139 %)	2.3e7 (1 %)	6.3
3	5e-6	1.2e5	23	1.7e9	5.7e5	2.2e8 (13 %)	3.1e7 (305 %)	8.8e7 (5 %)	16.2
4	1e-5	1.2e5	23	1.7e9	1.8e5	1.5e8 (9 %)	2.9e7 (286 %)	2.8e7 (2 %)	11.5
5	5e-5	1.4e5	23	1.7e9	1.5e6	3.8e8 (22 %)	3.5e7 (346 %)	2.4e8 (14 %)	28.0
6	1e-4	1.4e5	23	1.7e9	8.6e5	2.9e8 (17 %)	3.7e7 (372 %)	1.3e8 (8 %)	21.3
7	5e-4	1.8e5	23	1.7e9	4.8e6	9.5e8 (55 %)	4.8e7 (479 %)	7.5e8 (44 %)	70.0
8	1e-3	1.9e5	23	1.7e9	5.2e6	1.0e9 (58 %)	4.9e7 (487 %)	8.1e8 (47 %)	73.1

Table 6.4: Results of DAPCA-EC – IEEE 57 bus (DC-OPF)

To remedy the drawback that the consensus algorithm has to be executed at least once in each iteration of the DAPCA-EC (more than once if an L_k-update is necessary), we implemented a simple heuristic that allows L_k-updates only in the first iteration of the DAPCA-EC 5.1.1 as we observed for all AC/DC-OPF test cases that most of the L_k-updates were done in iteration $k = 1$.

The results of the DAPCA-EC combined with this heuristic, denoted by H1 in the following, can be seen in Table 6.5 which shows that 14 L_k-updates are done in the first iteration independent of the choice of β. Compared to the results of the DAPCA-EC in Table 6.4, the application of heuristic H1 reduces the number of consensus iterations by up to 97 %, the total information exchange by up to 81 % and the maximal computation time by up to 78 %. Moreover, row 2 - 4 of Table 6.4 show that the application of event-triggered communication reduces the total and dual communication by up to 23 % and 35 %. Accordingly, compared to the results of the DPCA-EC in Table 6.3, the application of the DAPCA-EC combined with heuristic H1 results in a reduction of the number of iterations by up to 80 %, the amount of total information exchange by up to 79 %, and the maximal computation time by up to 79 % as well.

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	3.0e4	14	5.0e7	5.3e3	2.0e7 (100 %)	4.9e6 (100 %)	8.2e5 (100 %)	1.7
2	1e-6	3.0e4	14	5.0e7	5.3e3	1.9e7 (95 %)	4.1e6 (85 %)	8.2e5 (100 %)	1.7
3	5e-6	2.4e4	14	5.0e7	5.3e3	1.6e7 (77 %)	3.1e6 (65 %)	8.2e5 (100 %)	1.4
4	1e-5	2.6e4	14	5.0e7	5.3e3	1.7e7 (83 %)	3.5e6 (71 %)	8.2e5 (100 %)	1.5
5	5e-5	5.0e4	14	5.0e7	5.3e3	3.2e7 (156 %)	6.9e6 (142 %)	8.2e5 (100 %)	2.7
6	1e-4	6.1e4	14	5.0e7	5.3e3	3.8e7 (189 %)	8.2e6 (169 %)	8.2e5 (100 %)	3.3
7	5e-4	1.2e5	14	5.0e7	5.3e3	7.2e7 (357 %)	1.6e7 (321 %)	8.2e5 (100 %)	6.5
8	1e-3	1.4e5	14	5.0e7	5.3e3	8.5e7 (421 %)	1.8e7 (379 %)	8.2e5 (100 %)	7.9

Table 6.5: Results of DAPCA-EC with heuristic H1 – IEEE 57 bus (DC-OPF)

Even though the DAPCA-EC combined with heuristic H1 converges for all AC/DC-OPF test cases that are considered in this work, we implemented the DAPCA-EC with two other simple heuristics, H3 and H6, that may in general be more robust regarding the convergence. In heuristic H3, the L_k-updates are allowed in the first iteration of the DAPCA-EC as well as in iteration k if k is a whole multiple of a rounded third of the number of iterations of the DAPCA-EC for $\beta = 0$ from Table 6.4, i.e., 3.1e4. Accordingly, in heuristic H6, the L_k-updates are allowed in iteration k if $k = 1$ or if k is a whole multiple of a rounded sixth of 3.1e4.

The results of the DAPCA-EC combined with H3 can be seen in Table 6.6 and with H6 in Table 6.7, showing that this more robust heuristics on the one side yields much better results compared the ones in Table 6.3 and Table 6.4. On the other side, slightly more iterations as well as a higher number of consensus iterations as compared to the application of H1 have to be accepted. Finally, the results in row 2-4 of Table 6.6 and Table 6.7 show that event-triggered communication helps to reduce the total communication by up to 29 % and 32 %, respectively, as well as the dual communication by up to 39 % and 41 %, respectively, and the consensus communication by up to 23 % and 34 %, respectively, if β is chosen small enough.

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	3.5e4	17	1.7e8	6.8e3	2.4e7 (100 %)	5.6e6 (100 %)	1.1e6 (100 %)	2.0
2	1e-6	3.5e4	17	1.7e8	6.3e3	2.3e7 (97 %)	5.1e6 (91 %)	9.9e5 (93 %)	2.0
3	5e-6	4.1e4	17	1.7e8	6.5e3	2.6e7 (110 %)	5.4e6 (96 %)	1.0e6 (96 %)	2.3
4	1e-5	2.6e4	14	5.0e7	5.3e3	1.7e7 (71 %)	3.5e6 (61 %)	8.2e5 (77 %)	1.5
5	5e-5	5.0e4	14	5.0e7	5.6e3	3.2e7 (134 %)	6.9e6 (122 %)	8.8e5 (82 %)	2.8
6	1e-4	6.9e4	15	7.5e7	5.9e3	4.3e7 (183 %)	9.4e6 (167 %)	9.2e5 (87 %)	3.8
7	5e-4	1.2e5	14	5.0e7	6.0e3	7.3e7 (307 %)	1.6e7 (278 %)	9.3e5 (88 %)	6.4
8	1e-3	1.4e5	16	1.1e8	6.2e3	8.1e7 (342 %)	1.7e7 (306 %)	9.7e5 (91 %)	7.3

Table 6.6: Results of DAPCA-EC with heuristic H3 – IEEE 57 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	3.7e4	17	1.7e8	8.5e3	2.5e7 (100 %)	5.9e6 (100 %)	1.3e6 (100 %)	2.1
2	1e-6	3.7e4	17	1.7e8	6.4e3	2.4e7 (98 %)	5.5e6 (93 %)	9.9e5 (75 %)	2.1
3	5e-6	3.5e4	16	1.1e8	6.2e3	2.2e7 (89 %)	4.6e6 (78 %)	9.7e5 (73 %)	2.0
4	1e-5	2.6e4	14	5.0e7	5.6e3	1.7e7 (68 %)	3.5e6 (59 %)	8.8e5 (66 %)	1.5
5	5e-5	5.0e4	14	5.0e7	5.9e3	3.2e7 (127 %)	6.9e6 (117 %)	9.3e5 (70 %)	2.8
6	1e-4	6.9e4	15	7.5e7	6.6e3	4.3e7 (174 %)	9.4e6 (160 %)	1.0e6 (77 %)	3.8
7	5e-4	1.2e5	14	5.0e7	7.0e3	7.3e7 (292 %)	1.6e7 (266 %)	1.1e6 (82 %)	6.5
8	1e-3	1.4e5	16	1.1e8	6.9e3	8.1e7 (325 %)	1.7e7 (292 %)	1.1e6 (81 %)	7.4

Table 6.7: Results of DAPCA-EC with heuristic H6 – IEEE 57 bus (DC-OPF)

Similar results of the DAPCA-EC applied to the DC-OPF problem (5.1) for the IEEE 14 and 30 bus test cases [ZMS11, Uni] can be found in the appendix 7.

6.3 Graphical representation of the IEEE power system test cases

Before we discuss the results of the DAPCA-EC applied to solve the AC-OPF problems, we have a look at the graphical representations of the IEEE test cases with 14, 30, and 57 buses from [ZMS11, Uni] as well as their chordal extensions. As detailed in section 5.3, the chordal extension of a power system network describes the communication topology of the agents that implement the DAPCA-EC 5.3.4 to solve problem (5.36) in parallel.
In Figure 6.1, the IEEE 14 bus test case with 20 branches (solid lines) is depicted as well as its chordal extension with 24 lines (solid and dotted lines). It follows that the communication topology of the agents does not coincide with the topology of the power system network, however, only 4 additional edges are needed for the communication which are approximately 5 % of the number of edges that would fill the 14 bus network to a complete graph.

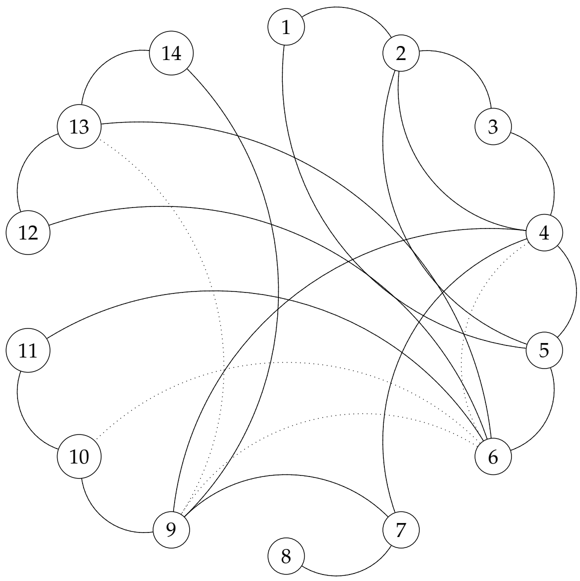

Figure 6.1: IEEE 14 bus system and its chordal extension.

In Figure 6.2, we see the IEEE 30 bus test case with 41 branches and its chordal extension with 55 lines, i.e., only 14 additional edges are needed for the communication which are approximately 3 % of the number of edges that would fill the 30 bus network to a complete graph.

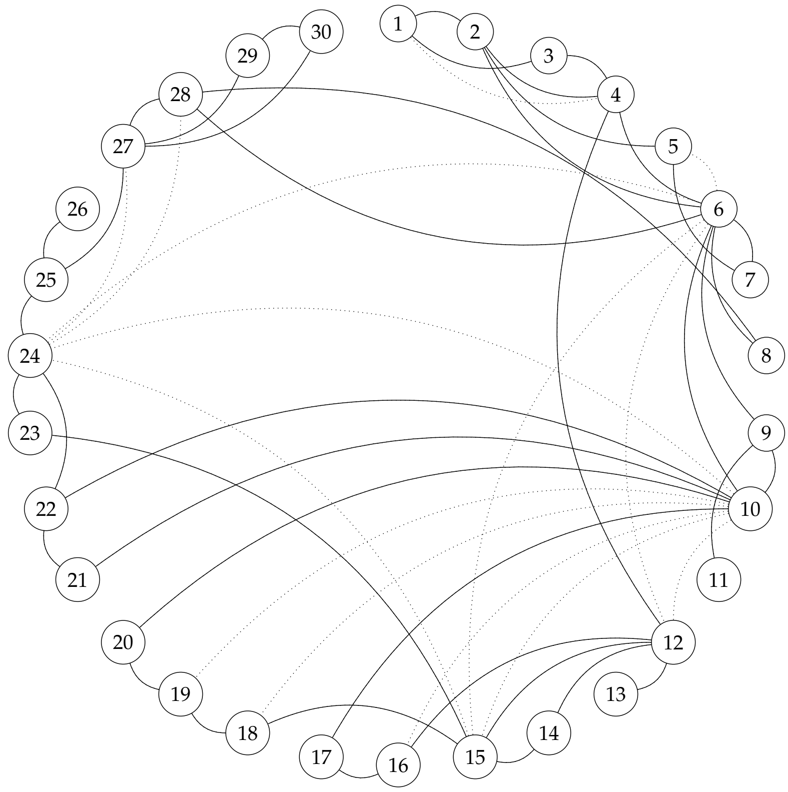

Figure 6.2: IEEE 30 bus system and its chordal extension (follows [DMUH14b, Fig. 3]).

In Figure 6.3, the IEEE 57 bus test case with 80 branches is depicted as well as its chordal extension with 137 lines, i.e., only 57 additional edges are needed for the communication which are approximately 3 % of the number of edges that would fill the 57 bus network to a complete graph.

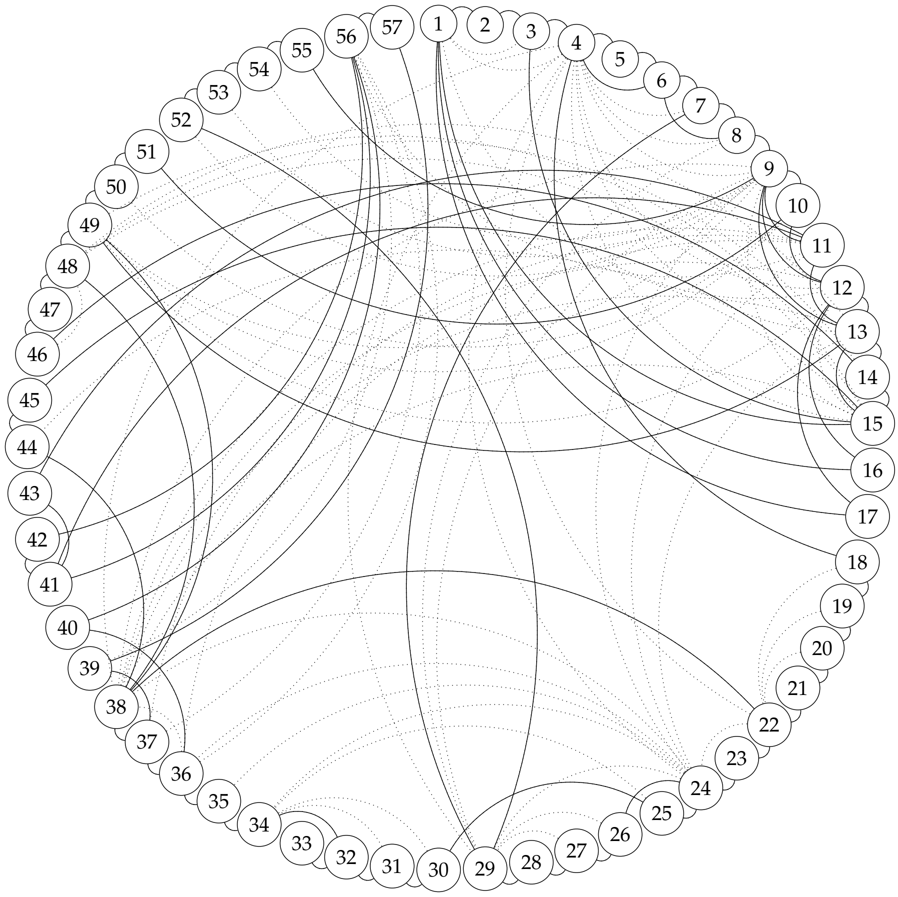

Figure 6.3: IEEE 57 bus system and its chordal extension.

6.4 IEEE 57 bus test case (AC-OPF)

Building on section 6.2, where the results of the DAPCA-EC to the DC-OPF problem are presented, we discuss in this section the results of the DAPCA-EC applied to the dual of the AC-OPF problem (5.36) for the IEEE 57 bus test case [ZMS11, Uni] with 7 generators and 80 branches as well. The results for the IEEE test cases with 14 and 30 buses from [ZMS11, Uni] can be found in the appendix 7.
As done in section 6.2, we firstly compare the results of the DPCA-EC (DAPCA-EC 3.2.2 with $L_{-1} = L_c$) for $\beta = 0$ with the results for $\beta > 0$ that are given in Table 6.8, to find out to what extent the communication exchange can be reduced by the usage of event-triggered communication.
Regarding the optimal dual multipliers $\Lambda^{\text{opt}} \in \mathbb{R}^{719}$ (see Remark 5.3.3 to understand the size of the dimension), we have $\|\Lambda^{\text{opt}}\| = 9.9463$. According to Theorem 3.1.5 combined with the scaling technique described in section 6.1, the choice $\epsilon = 5.52$ and $s = 19$ yields the following bounds on the primal gap:

$$
\begin{aligned}
-5.8703 = -\frac{1}{s}\left\|\Lambda^{\text{opt}}\right\|\left(\frac{1}{s}\left\|\Lambda^{\text{opt}}\right\| + \sqrt{\frac{1}{s^2}\left\|\Lambda^{\text{opt}}\right\|^2 + 2}\right)\epsilon \\
\leq \sum_{i \in \mathcal{N}_{\text{b}}} \Phi_i(\hat{x}_i, \hat{r}_i) - \sum_{(t,m)\in\mathcal{E}} \Phi_{tm}(\hat{r}_{tm}) - f_c^{\text{opt}} \leq \\
\epsilon = 5.52,
\end{aligned}
\tag{6.12}
$$

and the following bound on the constraint violation for $|J|$ equality constraints in (5.36b):

$$
\begin{aligned}
&\left\| \begin{matrix} E_{11} \bullet \sum_{s\in\Gamma(1,1)} \hat{W}^s - E_{11} \bullet A(\hat{x}, \hat{r}) \\ \vdots \\ E_{2n_{\text{b}}2n_{\text{b}}} \bullet \sum_{s\in\Gamma(2n_{\text{b}},2n_{\text{b}})} \hat{W}^s - E_{2n_{\text{b}}2n_{\text{b}}} \bullet A(\hat{x}, \hat{r}) \end{matrix} \right\| \\
&\leq \frac{\epsilon}{s}\left(\frac{1}{s}\left\|\Lambda^{\text{opt}}\right\| + \sqrt{\frac{1}{s^2}\left\|\Lambda^{\text{opt}}\right\|^2 + 2}\right) = 0.5902,
\end{aligned}
\tag{6.13}
$$

where

$$
\hat{x}_i = \sum_{j=0}^{k_{\text{fin}}} \frac{2(j+1)}{(k_{\text{fin}}+1)(k_{\text{fin}}+2)} x_i^{j+1} \in \mathbb{R}^6, \tag{6.14a}
$$

$$
\hat{r}_i = \sum_{j=0}^{k_{\text{fin}}} \frac{2(j+1)}{(k_{\text{fin}}+1)(k_{\text{fin}}+2)} r_i^{j+1} \in \mathbb{S}^4, \tag{6.14b}
$$

$$
\hat{r}_{tm} = \sum_{j=0}^{k_{\text{fin}}} \frac{2(j+1)}{(k_{\text{fin}}+1)(k_{\text{fin}}+2)} r_{tm}^{j+1} \in \mathbb{S}^6, \tag{6.14c}
$$

$$
\hat{W}^s = \sum_{j=0}^{k_{\text{fin}}} \frac{2(j+1)}{(k_{\text{fin}}+1)(k_{\text{fin}}+2)} W^{s,j+1} \in \mathbb{S}^{2C_s} \tag{6.14d}
$$

are the convex sums of the primal iterates for $i \in \mathcal{N}_\text{b}$, $(t,m) \in \mathcal{E}$, and $s = 1,\dots,p = 52$. Here, the maximal cliques $2C_s$ are defined as in Remark 5.3.3 and one has $6 \leq 2C_s \leq 12$ for $s = 1,\dots,52$. The dimension of the primal variable space is given by 2793. Moreover, $k_\text{fin} = 783884$ is the number of necessary iterations given by Theorem 3.1.5. We notice that the bounds in (6.12) and (6.13) follow immediately by rewriting the Frobenius product in constraint (5.36b) of problem (5.36) in vectorized form, applying Theorem 3.1.5, and retrieve the notation with the Frobenius product.
It follows for the choice of δ in the threshold $\Delta_k = \beta\delta^k$, that

$$\delta^{k_\text{fin}/2} = 0.025$$

is satisfied by $\delta \approx 0.9999$. As described in section 6.1, ϵ and s were chosen such that the absolute values of the bounds on the primal gap in (6.12) are approximately 1/100 of the primal gap at the starting point $\Lambda^0 = 0 \in \mathbb{R}^{719}$ which is -552.1739. The constraint violation at the starting point is $2.0104e3$.
For a better overview the above figures are resumed in the following list:

IEEE 57 bus test case (AC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 2793, \qquad \text{dual: } 719,$$

- Accuracy $\epsilon = 5.52$ and scaling factor $s = 19$,
- Norm of optimal dual multipliers: $\|\Lambda^\text{opt}\| = 9.9463$,
- Necessary number of iterations (Theorem 3.1.5): $k_\text{fin} = 783884$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -5.8703, \qquad \text{upper bound: } 5.52, \tag{6.15}$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$0.5902, \tag{6.16}$$

- Primal gap at starting point $\Lambda^0 = 0 \in \mathbb{R}^{719}$: -552.1739,
- Constraint violation at starting point: $2.0104e3$,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9999^k$.

In row 1 of Table 6.8, the result of the DPCA-EC is given for the threshold $\Delta_k = 0$, i.e., no event-triggered communication was used. As expected, the primal gap (column 3) and the constraint violation at the approximate solution (column 4) satisfy the bounds (6.15) and (6.16) according to Theorem 3.1.5.
If event-triggered communication is introduced by choosing $\beta > 0$, the results in row 2 - 7 of Table 6.8 show that the total communication can be reduced by up to 11 % and the dual communication by up to 21 %, still satisfying the bounds on the primal gap (6.15) and the constraint violation (6.16). If β is chosen larger as in row 8, the bound on the primal gap is not satisfied anymore.
With respect to the results of the corresponding IEEE 57 bus DC-OPF problem in Table 6.2, where the total communication could be reduced by up to 13 % and the dual communication by up to 27 %, the communication savings here are comparable. However, the maximal computation times of an agent in the last column of Table 6.8 are not surprisingly bigger compared to approximately 7 seconds for the DC-OPF problem in Table 6.2.

	β	PG	CV	TC	DC	MCTpA
1	0	-0.9049	0.1527	4.8e9 (100 %)	2.4e9 (100 %)	307.8
2	5e-9	-0.8601	0.1536	4.8e9 (100 %)	2.4e9 (99 %)	307.6
3	1e-8	-0.8730	0.1553	4.8e9 (99 %)	2.4e9 (98 %)	311.3
4	5e-8	-0.8799	0.1563	4.7e9 (97 %)	2.3e9 (94 %)	304.8
5	1e-7	-0.8633	0.1564	4.6e9 (96 %)	2.2e9 (91 %)	312.7
6	5e-7	-0.2521	0.1721	4.4e9 (92 %)	2.0e9 (83 %)	309.7
7	1e-6	1.0127	0.2076	4.3e9 (89 %)	1.9e9 (79 %)	313.7
8	5e-6	12.4716	0.4644	4.0e9 (83 %)	1.7e9 (69 %)	306.2

Table 6.8: Results of the DPCA-EC – IEEE 57 bus (AC-OPF)

Just like in section 6.2, we investigate the impact of event-triggered communication if the following stopping criterion is used in the DPCA-EC instead of the pre-given number $k_{\text{fin}} = 783884$ of iterations which may be not necessary to obtain the primal gap of -0.9049 and the constraint violation of 0.1527 in row 1 of Table 6.8.

To this end, let

$$\hat{x}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} x_i^{j+1} \in \mathbb{R}^6, \tag{6.17a}$$

$$\hat{r}_i = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} r_i^{j+1} \in \mathbb{S}^4, \tag{6.17b}$$

$$\hat{r}_{tm} = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} r_{tm}^{j+1} \in \mathbb{S}^6, \tag{6.17c}$$

$$\hat{W}^s = \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} W^{s,j+1} \in \mathbb{S}^{2C_s} \tag{6.17d}$$

be the approximate solutions after k iterations for $i \in \mathcal{N}_\mathrm{b}$, $(t,m) \in \mathcal{E}$, and $s = 1,\ldots,p$.

DAPCA-EC stopping criterion for the IEEE 57 bus test case (AC-OPF):

For $k \geq 0$

if

$$|\text{primal gap at (6.17)}| \leq 0.9049 \text{ and constraint violation at (6.17) } \leq 0.1527 \tag{6.18}$$

then

stop.

else

continue.

In Table 6.9 the results of the DPCA-EC with stopping criterion (6.18) are given.

As we observed for the DC-OPF problem in Table 6.3 as well, the result in row 1 shows that a similar number of iterations (column 3) is needed to compute an approximate solution without event-triggered communication that satisfies the same primal gap and constraint violation as the solution in row 1 of Table 6.8. Actually, the number of iterations in row 1 of Table 6.9 is slightly higher compared to row 1 of Table 6.8 as the values of the primal gap and the constraint violation in (6.18) are rounded off values given by MATLAB.

Consistent with the observation made for the DC-OPF problem, row 4 - 6 of Table 6.9 show that the information exchange can still be reduced even if the number of iterations is higher compared to the result obtained without event-triggered communication in row 1. However, if β is chosen to large, as in row 7 and 8, even more communication is required compared to row 1.

	β	NoI	TC	DC	MCTpA
1	0	7.8391e5	4.8e9 (100 %)	2.4e9 (100 %)	314.3
2	5e-9	7.8467e5	4.8e9 (100 %)	2.4e9 (99 %)	305.5
3	1e-8	7.8739e5	4.8e9 (100 %)	2.4e9 (99 %)	312.1
4	5e-8	7.9017e5	4.7e9 (98 %)	2.3e9 (95 %)	315.6
5	1e-7	7.8940e5	4.7e9 (96 %)	2.2e9 (92 %)	320.2
6	5e-7	8.4125e5	4.8e9 (99 %)	2.2e9 (91 %)	338.7
7	1e-6	9.0544e5	5.1e9 (105 %)	2.3e9 (95 %)	360.5
8	5e-6	3.2169e6	1.9e10 (393 %)	9.2e9 (379 %)	1186.8

Table 6.9: Results of DPCA-EC with stopping criterion (6.18) – IEEE 57 bus (AC-OPF)

In Table 6.10 the results of the DAPCA-EC 5.3.4 are given for the step-size parameters $\gamma = 1.2$ and $L_{-1} = 10^{-2} L_c$. The corresponding Lipschitz constant (5.40) is

$$L_c = 8.4797e11, \tag{6.19}$$

compared to $L_c = 1.7149e9$ regarding the IEEE 57 bus DC-OPF problem. Stopping criterion (6.18) was used for the computation of the results given in Table 6.10, to be able to compare with the results of the DPCA-EC in Table 6.9.

As can be seen in row 1 of Table 6.10 only $1.5e5$ iterations are needed to compute a solution without event-triggered communication which is a decrease, compared to the number of iterations of the DPCA-EC in row 1 of Table 6.9, by approximately 81 %. As a reminder, the corresponding decrease concerning the IEEE 57 bus DC-OPF problem amounts to 75 %, i.e., no significant difference can be observed here.

Similar to the results for the DC-OPF problem in Table 6.4, the large number of consensus iterations (column 6) in row 1 of Table 6.10 results in a total information exchange (column 7) that is bigger compared to the results in row 1 - 7 of Table 6.9. However, the results in row 2 - 7 of Table 6.10 show here as well that the application of event-triggered communication reduces the consensus communication by up to over 99 % which yields a reduction of the total communication by up to 82 %.

Unlike the corresponding results for the DC-OPF problem in Table 6.4, the maximal computation time could be reduced by approximately 66 %, compared to the results of the DPCA-EC in Table 6.9, which makes the application of the DAPCA-EC 5.3.4 attractive even if no heuristic is used. This can be explained by the fact that only 6 - 10 L_k-updates needed to be executed due to the high starting value $L_{-1} = 10^{-2} L_c$.

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.5e5	7	3.0e10	4.9e7	9.5e9 (100 %)	9.5e8 (100 %)	7.6e9 (100 %)	701.1
2	5e-9	1.4e5	6	2.5e10	1.4e5	1.7e9 (18 %)	8.3e8 (87 %)	2.2e7 (0 %)	102.9
3	1e-8	1.4e5	6	2.5e10	1.4e5	1.7e9 (18 %)	8.3e8 (87 %)	2.2e7 (0 %)	104.5
4	5e-8	1.4e5	6	2.5e10	1.4e5	1.7e9 (18 %)	8.2e8 (86 %)	2.2e7 (0 %)	106.2
5	1e-7	1.5e5	6	2.5e10	1.5e5	1.8e9 (19 %)	8.4e8 (88 %)	2.3e7 (0 %)	117.2
6	5e-7	1.8e5	6	2.5e10	1.9e5	2.2e9 (23 %)	1.0e9 (106 %)	2.9e7 (0 %)	143.6
7	1e-6	3.8e5	6	2.5e10	3.9e5	4.6e9 (48 %)	2.1e9 (225 %)	6.0e7 (1 %)	297.6
8	5e-6	1.4e6	10	5.3e10	1.1e7	1.8e10 (188 %)	7.9e9 (829 %)	1.7e9 (22 %)	1133.2

Table 6.10: Results of DAPCA-EC – IEEE 57 bus (AC-OPF)

Nevertheless, we investigate in the following if the heuristics H1, H3, and H6, described in section 6.2, can reduce the number of iterations and the maximal computation time even more.

The results of the DAPCA-EC combined with heuristic H1 can be seen in Table 6.11, showing that 6 L_k-updates are done in the first iteration independent of the choice of β, i.e., compared to the results in Table 6.10, the number of iterations can not be further reduced by heuristic H1, however, the consensus iterations are reduced by up to 99 % as well as the consensus communication which results in a further computation time reduction by up to 49 %.

Moreover, row 2 - 5 of Table 6.11 show that the application of event-triggered communication does not have a remarkable impact in this case. However, compared to the results of the DPCA-EC in Table 6.9, the application of the DAPCA-EC combined with heuristic H1 results in a reduction of the number of iterations by up to 83 %, the amount of total information exchange by up to 83 %, and the maximal computation time by up to 83 % as well which coincides with the observations made for the IEEE 57 bus DC-OPF problem in section 6.2.

Finally, the results of the DAPCA-EC implemented with heuristics H3 and H6 can be seen in Table 6.12 and 6.13, showing that this heuristics yield the same slightly worse results compared to heuristic H1, since no more than 6 - 7 L_k-updates are needed which are all done in the first iteration. Thus allowing additional updates in heuristics H3 and H6 does not yield an improvement regarding the number of iterations and the maximal computation time for this case.

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.4e5	6	2.5e10	2.2e3	8.4e8 (100 %)	4.2e8 (100 %)	3.5e5 (100 %)	54.9
2	5e-9	1.4e5	6	2.5e10	2.2e3	8.4e8 (99 %)	4.1e8 (98 %)	3.5e5 (100 %)	53.5
3	1e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (99 %)	4.1e8 (98 %)	3.5e5 (100 %)	55.6
4	5e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (98 %)	3.9e8 (93 %)	3.5e5 (100 %)	56.1
5	1e-7	1.5e5	6	2.5e10	2.2e3	8.5e8 (101 %)	3.9e8 (93 %)	3.5e5 (100 %)	58.6
6	5e-7	1.8e5	6	2.5e10	2.2e3	1.0e9 (121 %)	4.5e8 (106 %)	3.5e5 (100 %)	73.2
7	1e-6	3.8e5	6	2.5e10	2.2e3	2.2e9 (256 %)	9.6e8 (229 %)	3.5e5 (100 %)	148.7
8	5e-6	1.4e6	6	2.5e10	2.2e3	8.0e9 (950 %)	3.8e9 (899 %)	3.5e5 (100 %)	487.8

Table 6.11: Results of DAPCA-EC with heuristic H1 – IEEE 57 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.4e5	6	2.5e10	2.7e3	8.4e8 (100 %)	4.2e8 (100 %)	4.3e5 (100 %)	52.3
2	5e-9	1.4e5	6	2.5e10	2.2e3	8.4e8 (99 %)	4.1e8 (98 %)	3.5e5 (81 %)	54.4
3	1e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (99 %)	4.1e8 (98 %)	3.5e5 (81 %)	54.1
4	5e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (98 %)	3.9e8 (93 %)	3.5e5 (81 %)	56.3
5	1e-7	1.5e5	6	2.5e10	2.2e3	8.5e8 (101 %)	3.9e8 (93 %)	3.5e5 (81 %)	56.3
6	5e-7	1.8e5	6	2.5e10	2.2e3	1.0e9 (121 %)	4.5e8 (106 %)	3.5e5 (81 %)	70.4
7	1e-6	3.8e5	6	2.5e10	2.2e3	2.2e9 (256 %)	9.6e8 (229 %)	3.5e5 (81 %)	139.7
8	5e-6	1.4e6	7	3.0e10	2.6e3	8.0e9 (947 %)	3.8e9 (897 %)	4.0e5 (95 %)	477.1

Table 6.12: Results of DAPCA-EC with heuristic H3 – IEEE 57 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.4e5	6	2.5e10	3.8e3	8.4e8 (100 %)	4.2e8 (100 %)	6.0e5 (100 %)	58.3
2	5e-9	1.4e5	6	2.5e10	2.2e3	8.4e8 (99 %)	4.1e8 (98 %)	3.5e5 (58 %)	54.0
3	1e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (99 %)	4.1e8 (98 %)	3.5e5 (58 %)	53.9
4	5e-8	1.4e5	6	2.5e10	2.2e3	8.3e8 (98 %)	3.9e8 (93 %)	3.5e5 (58 %)	54.7
5	1e-7	1.5e5	6	2.5e10	2.2e3	8.5e8 (101 %)	3.9e8 (93 %)	3.5e5 (58 %)	62.0
6	5e-7	1.8e5	6	2.5e10	2.2e3	1.0e9 (121 %)	4.5e8 (106 %)	3.5e5 (58 %)	75.9
7	1e-6	3.8e5	6	2.5e10	2.2e3	2.2e9 (256 %)	9.6e8 (229 %)	3.5e5 (58 %)	154.2
8	5e-6	1.4e6	7	3.0e10	2.9e3	8.0e9 (947 %)	3.8e9 (897 %)	4.5e5 (75 %)	517.4

Table 6.13: Results of DAPCA-EC with heuristic H6 – IEEE 57 bus (AC-OPF)

Similar results of the DAPCA-EC applied to the dual of the AC-OPF problem (5.36) for the IEEE 14 and 30 bus test cases [ZMS11, Uni] can be found in the appendix 7.

7 Appendix - numerical results

7.1 Numerical results for the DC-OPF problem

7.1.1 IEEE 14 bus test case (DC-OPF)

In this section, the results of the DAPCA-EC applied to the DC-OPF problem (5.1) for the IEEE 14 bus test case [ZMS11, Uni] with 5 generators and 20 branches are presented without further comment as they are similar to the results discussed detailed in section 6.2 for the IEEE 57 bus test case.

IEEE 14 bus test case (DC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 14+5=19, \qquad \text{dual: } 14+40=54,$$

- Accuracy $\epsilon = 0.07$ and scaling factor $s = 29$,
- Norm of optimal dual multipliers: $\|(\mu,\lambda)^{\text{opt}}\| = 14.5985$,
- Lipschitz constant $L_c = 3.3556e7$ (5.9),
- Necessary number of iterations (Theorem 3.1.5): $k_{\text{fin}} = 43790$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -0.0706, \qquad \text{upper bound: } 0.07, \tag{7.1}$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$0.0048, \tag{7.2}$$

- Primal gap at starting point $(\mu,\lambda)^0 = 0 \in \mathbb{R}^{14+40}$: $-7.6426 = -f_c^{\text{opt}}$,
- Constraint violation at starting point: 0.7473,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9998^k$,
- Step-size parameters for DAPCA-EC: $\gamma = 1.1$ and $L_{-1} = 10^{-2} L_c$.

	β	PG	CV	TC	DC	MCTpA
1	0	-0.0167	0.0012	7.0e6 (100 %)	1.8e6 (100 %)	2.2
2	1e-6	-0.0170	0.0012	7.0e6 (100 %)	1.7e6 (99 %)	2.2
3	5e-6	-0.0171	0.0012	6.9e6 (99 %)	1.7e6 (96 %)	2.2
4	1e-5	-0.0170	0.0012	6.9e6 (98 %)	1.7e6 (94 %)	2.2
5	5e-5	-0.0172	0.0013	6.7e6 (96 %)	1.5e6 (86 %)	2.2
6	1e-4	-0.0143	0.0014	6.6e6 (94 %)	1.4e6 (82 %)	2.1
7	5e-4	0.0406	0.0013	6.3e6 (89 %)	1.3e6 (74 %)	2.1
8	1e-3	0.0735	0.0045	6.0e6 (86 %)	1.2e6 (70 %)	2.1

Table 7.1: Results of the DPCA-EC – IEEE 14 bus (DC-OPF)

DAPCA-EC stopping criterion for the IEEE 14 bus test case (DC-OPF):

For $k \geq 0$

if

$$|\text{primal gap at (6.9)}| \leq 0.0167 \text{ and constraint violation at (6.9) } \leq 0.0012 \tag{7.3}$$

then

stop.

else

continue.

	β	NoI	TC	DC	MCTpA
1	0	4.2473e4	6.8e6 (100 %)	1.7e6 (100 %)	2.1
2	1e-6	4.2333e4	6.8e6 (99 %)	1.7e6 (99 %)	2.1
3	5e-6	4.2303e4	6.7e6 (99 %)	1.6e6 (96 %)	2.1
4	1e-5	4.2317e4	6.6e6 (98 %)	1.6e6 (94 %)	2.1
5	5e-5	4.4400e4	6.8e6 (100 %)	1.5e6 (90 %)	2.2
6	1e-4	4.2105e4	6.3e6 (93 %)	1.4e6 (81 %)	2.0
7	5e-4	5.3557e4	7.8e6 (115 %)	1.7e6 (99 %)	2.6
8	1e-3	7.6406e4	1.1e7 (165 %)	2.5e6 (149 %)	3.7

Table 7.2: Results of DPCA-EC with stopping criterion (7.3) – IEEE 14 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.8e4	30	5.9e6	1.6e6	7.1e7 (100 %)	1.4e6 (100 %)	6.5e7 (100 %)	18.3
2	1e-6	2.0e4	32	7.1e6	5.2e4	8.6e6 (12 %)	1.6e6 (114 %)	2.1e6 (3 %)	2.4
3	5e-6	4.4e4	49	3.4e7	7.2e4	1.5e7 (21 %)	2.9e6 (206 %)	2.9e6 (4 %)	4.3
4	1e-5	4.5e4	49	3.4e7	3.7e4	1.2e7 (16 %)	2.4e6 (168 %)	1.5e6 (2 %)	3.3
5	5e-5	4.3e4	49	3.4e7	2.1e5	2.1e7 (29 %)	2.9e6 (206 %)	8.4e6 (13 %)	5.8
6	1e-4	3.7e4	48	3.3e7	4.0e5	2.8e7 (39 %)	2.8e6 (201 %)	1.6e7 (24 %)	7.8
7	5e-4	8.2e4	46	2.7e7	2.8e6	1.4e8 (193 %)	6.4e6 (455 %)	1.1e8 (169 %)	38.0
8	1e-3	8.9e4	46	2.7e7	2.7e6	1.4e8 (194 %)	6.9e6 (485 %)	1.1e8 (168 %)	36.3

Table 7.3: Results of DAPCA-EC – IEEE 14 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	9.4e3	12	1.1e6	1.2e3	1.6e6 (100 %)	3.8e5 (100 %)	4.9e4 (100 %)	0.5
2	1e-6	1.1e4	12	1.1e6	1.2e3	1.8e6 (116 %)	4.4e5 (116 %)	4.9e4 (100 %)	0.6
3	5e-6	9.2e3	12	1.1e6	1.2e3	1.5e6 (97 %)	3.6e5 (96 %)	4.9e4 (100 %)	0.5
4	1e-5	9.8e3	12	1.1e6	1.2e3	1.6e6 (104 %)	3.8e5 (102 %)	4.9e4 (100 %)	0.5
5	5e-5	9.3e3	12	1.1e6	1.2e3	1.5e6 (97 %)	3.3e5 (88 %)	4.9e4 (100 %)	0.5
6	1e-4	2.0e4	12	1.1e6	1.2e3	3.2e6 (205 %)	7.4e5 (196 %)	4.9e4 (100 %)	1.0
7	5e-4	6.0e4	12	1.1e6	1.2e3	9.6e6 (616 %)	2.3e6 (608 %)	4.9e4 (100 %)	3.0
8	1e-3	8.2e4	12	1.1e6	1.2e3	1.3e7 (830 %)	3.1e6 (821 %)	4.9e4 (100 %)	4.0

Table 7.4: Results of DAPCA-EC with heuristic H1 – IEEE 14 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.8e4	38	1.3e7	3.8e3	3.1e6 (100 %)	7.3e5 (100 %)	1.5e5 (100 %)	0.9
2	1e-6	1.1e4	12	1.1e6	1.3e3	1.8e6 (59 %)	4.4e5 (60 %)	5.2e4 (34 %)	0.6
3	5e-6	9.2e3	12	1.1e6	1.3e3	1.5e6 (49 %)	3.6e5 (50 %)	5.2e4 (34 %)	0.5
4	1e-5	9.1e3	21	2.5e6	2.1e3	1.5e6 (50 %)	3.5e5 (48 %)	8.4e4 (55 %)	0.5
5	5e-5	9.3e3	12	1.1e6	1.2e3	1.5e6 (49 %)	3.3e5 (46 %)	4.9e4 (32 %)	0.5
6	1e-4	2.0e4	33	7.8e6	3.4e3	3.2e6 (106 %)	7.2e5 (98 %)	1.4e5 (89 %)	1.0
7	5e-4	5.9e4	21	2.5e6	2.3e3	9.4e6 (306 %)	2.2e6 (304 %)	9.2e4 (60 %)	2.9
8	1e-3	7.7e4	31	6.4e6	3.4e3	1.2e7 (398 %)	2.9e6 (395 %)	1.4e5 (88 %)	3.9

Table 7.5: Results of DAPCA-EC with heuristic H3 – IEEE 14 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	1.9e4	36	1.0e7	3.9e3	3.1e6 (100 %)	7.5e5 (100 %)	1.6e5 (100 %)	1.0
2	1e-6	1.2e4	24	3.3e6	2.4e3	2.1e6 (65 %)	4.9e5 (65 %)	9.5e4 (60 %)	0.6
3	5e-6	1.2e4	23	3.0e6	2.3e3	2.0e6 (63 %)	4.6e5 (61 %)	9.3e4 (59 %)	0.6
4	1e-5	10.0e3	20	2.3e6	2.2e3	1.7e6 (53 %)	3.8e5 (51 %)	8.8e4 (56 %)	0.5
5	5e-5	1.1e4	18	1.9e6	2.0e3	1.8e6 (56 %)	3.7e5 (49 %)	8.1e4 (51 %)	0.7
6	1e-4	1.8e4	30	5.9e6	3.3e3	3.0e6 (94 %)	6.5e5 (87 %)	1.3e5 (84 %)	1.0
7	5e-4	5.9e4	21	2.5e6	2.3e3	9.4e6 (298 %)	2.2e6 (296 %)	9.4e4 (59 %)	3.0
8	1e-3	7.8e4	28	4.8e6	3.4e3	1.2e7 (396 %)	2.9e6 (394 %)	1.3e5 (85 %)	4.0

Table 7.6: Results of DAPCA-EC with heuristic H6 – IEEE 14 bus (DC-OPF)

7.1.2 IEEE 30 bus test case (DC-OPF)

In this section, the results of the DAPCA-EC applied to the DC-OPF problem (5.1) for the IEEE 30 bus test case [ZMS11, Uni] with 6 generators and 41 branches are presented without further comment as they are similar to the results discussed detailed in section 6.2 for the IEEE 57 bus test case.

IEEE 30 bus test case (DC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 30 + 6 = 36, \qquad \text{dual: } 30 + 82 = 112,$$

- Accuracy $\epsilon = 0.005$ and scaling factor $s = 4$,
- Norm of optimal dual multipliers: $\|(\mu, \lambda)^{\text{opt}}\| = 2.0754$,
- Lipschitz constant $L_c = 6.6003e6$ (5.9),
- Necessary number of iterations (Theorem 3.1.5): $k_{\text{fin}} = 72666$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -0.0053, \qquad \text{upper bound: } 0.005, \tag{7.4}$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$7.0284e - 4, \tag{7.5}$$

- Primal gap at starting point $(\mu, \lambda)^0 = 0 \in \mathbb{R}^{30+82}$: $-0.5652 = -f_c^{\text{opt}}$,
- Constraint violation at starting point: 0.3617,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9998^k$,
- Step-size parameters for DAPCA-EC: $\gamma = 1.5$ and $L_{-1} = 10^{-4} L_c$.

	β	PG	CV	TC	DC	MCTpA
1	0	-0.0014	0.0007	2.4e7 (100 %)	6.0e6 (100 %)	3.8
2	1e-6	-0.0014	0.0007	2.4e7 (99 %)	5.8e6 (97 %)	3.8
3	5e-6	-0.0014	0.0007	2.3e7 (98 %)	5.5e6 (91 %)	3.7
4	1e-5	-0.0014	0.0007	2.3e7 (97 %)	5.3e6 (89 %)	3.7
5	5e-5	-0.0008	0.0006	2.2e7 (93 %)	4.8e6 (81 %)	3.7
6	1e-4	-0.0010	0.0010	2.2e7 (91 %)	4.6e6 (77 %)	3.6
7	5e-4	0.0030	0.0032	2.0e7 (83 %)	4.0e6 (67 %)	3.5
8	1e-3	0.0103	0.0066	1.9e7 (79 %)	3.7e6 (62 %)	3.4

Table 7.7: Results of the DPCA-EC – IEEE 30 bus (DC-OPF)

DAPCA-EC stopping criterion for the IEEE 30 bus test case (DC-OPF):

For $k \geq 0$

if

$$|\text{primal gap at (6.9)}| \leq 0.0014 \text{ and constraint violation at (6.9) } \leq 7.0284e-4 \qquad (7.6)$$

then

stop.

else

continue.

	β	NoI	TC	DC	MCTpA
1	0	6.8775e4	2.3e7 (100 %)	5.7e6 (100 %)	3.6
2	1e-6	6.8875e4	2.2e7 (99 %)	5.5e6 (97 %)	3.6
3	5e-6	6.8940e4	2.2e7 (98 %)	5.2e6 (91 %)	3.5
4	1e-5	6.9260e4	2.2e7 (97 %)	5.0e6 (89 %)	3.5
5	5e-5	6.7019e4	2.0e7 (90 %)	4.4e6 (77 %)	3.4
6	1e-4	6.9818e4	2.1e7 (92 %)	4.4e6 (78 %)	3.5
7	5e-4	9.3364e4	2.7e7 (118 %)	5.7e6 (100 %)	4.6
8	1e-3	1.2311e5	3.5e7 (156 %)	7.8e6 (138 %)	6.0

Table 7.8: Results of DPCA-EC with stopping criterion (7.6) – IEEE 30 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	2.2e4	17	6.5e5	4.3e6	3.6e8 (100 %)	3.5e6 (100 %)	3.5e8 (100 %)	46.8
2	1e-6	3.3e4	19	1.5e6	1.1e5	3.0e7 (8 %)	5.2e6 (147 %)	8.8e6 (3 %)	4.2
3	5e-6	5.0e4	22	4.9e6	5.6e4	3.7e7 (10 %)	7.8e6 (220 %)	4.6e6 (1 %)	5.2
4	1e-5	6.1e4	23	6.6e6	3.9e4	3.2e7 (9 %)	6.7e6 (189 %)	3.2e6 (1 %)	4.7
5	5e-5	6.1e4	23	6.6e6	7.6e4	3.9e7 (11 %)	7.7e6 (216 %)	6.3e6 (2 %)	5.6
6	1e-4	7.0e4	23	6.6e6	3.0e5	6.1e7 (17 %)	8.7e6 (246 %)	2.4e7 (7 %)	8.6
7	5e-4	1.0e5	23	6.6e6	1.6e5	6.5e7 (18 %)	1.2e7 (347 %)	1.3e7 (4 %)	9.5
8	1e-3	1.3e5	23	6.6e6	2.9e5	8.6e7 (24 %)	1.5e7 (414 %)	2.4e7 (7 %)	12.4

Table 7.9: Results of DAPCA-EC – IEEE 30 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	2.3e4	16	4.3e5	2.8e3	8.6e6 (100 %)	2.7e6 (100 %)	2.3e5 (100 %)	1.2
2	1e-6	1.8e4	16	4.3e5	2.8e3	5.9e6 (68 %)	1.3e6 (50 %)	2.3e5 (100 %)	0.9
3	5e-6	1.8e4	16	4.3e5	2.8e3	5.8e6 (67 %)	1.2e6 (45 %)	2.3e5 (100 %)	1.0
4	1e-5	1.8e4	16	4.3e5	2.8e3	5.9e6 (68 %)	1.2e6 (45 %)	2.3e5 (100 %)	1.0
5	5e-5	2.9e4	16	4.3e5	2.8e3	9.2e6 (107 %)	2.0e6 (76 %)	2.3e5 (100 %)	1.5
6	1e-4	4.2e4	16	4.3e5	2.8e3	1.3e7 (156 %)	2.9e6 (110 %)	2.3e5 (100 %)	2.2
7	5e-4	8.8e4	16	4.3e5	2.8e3	2.8e7 (321 %)	6.2e6 (229 %)	2.3e5 (100 %)	4.5
8	1e-3	1.0e5	16	4.3e5	2.8e3	3.2e7 (372 %)	7.2e6 (269 %)	2.3e5 (100 %)	5.2

Table 7.10: Results of DAPCA-EC with heuristic H1 – IEEE 30 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	2.4e4	18	9.8e5	3.7e3	8.4e6 (100 %)	2.2e6 (100 %)	3.0e5 (100 %)	1.3
2	1e-6	1.8e4	16	4.3e5	2.8e3	5.9e6 (70 %)	1.3e6 (61 %)	2.3e5 (75 %)	1.0
3	5e-6	2.1e4	17	6.5e5	2.9e3	6.8e6 (81 %)	1.4e6 (65 %)	2.4e5 (79 %)	1.1
4	1e-5	2.1e4	17	6.5e5	3.1e3	6.8e6 (80 %)	1.4e6 (64 %)	2.6e5 (85 %)	1.1
5	5e-5	2.9e4	16	4.3e5	2.8e3	9.2e6 (110 %)	2.0e6 (92 %)	2.3e5 (75 %)	1.5
6	1e-4	4.2e4	16	4.3e5	3.3e3	1.3e7 (160 %)	2.9e6 (134 %)	2.7e5 (89 %)	2.2
7	5e-4	8.8e4	16	4.3e5	2.8e3	2.8e7 (329 %)	6.2e6 (279 %)	2.3e5 (75 %)	4.5
8	1e-3	1.0e5	16	4.3e5	3.2e3	3.2e7 (382 %)	7.2e6 (327 %)	2.6e5 (87 %)	5.2

Table 7.11: Results of DAPCA-EC with heuristic H3 – IEEE 30 bus (DC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	3.0e4	19	1.5e6	4.5e3	1.0e7 (100 %)	2.6e6 (100 %)	3.7e5 (100 %)	1.6
2	1e-6	1.8e4	16	4.3e5	2.8e3	5.9e6 (57 %)	1.3e6 (51 %)	2.3e5 (61 %)	0.9
3	5e-6	2.1e4	17	6.5e5	3.1e3	6.9e6 (67 %)	1.5e6 (56 %)	2.6e5 (69 %)	1.1
4	1e-5	1.8e4	16	4.3e5	3.2e3	5.9e6 (57 %)	1.2e6 (46 %)	2.6e5 (70 %)	1.0
5	5e-5	2.9e4	16	4.3e5	3.0e3	9.3e6 (89 %)	2.0e6 (78 %)	2.4e5 (65 %)	1.5
6	1e-4	4.0e4	17	6.5e5	3.2e3	1.3e7 (123 %)	2.8e6 (107 %)	2.6e5 (70 %)	2.1
7	5e-4	8.8e4	16	4.3e5	3.2e3	2.8e7 (266 %)	6.2e6 (236 %)	2.6e5 (70 %)	4.5
8	1e-3	1.0e5	16	4.3e5	3.8e3	3.2e7 (309 %)	7.2e6 (276 %)	3.1e5 (84 %)	5.3

Table 7.12: Results of DAPCA-EC with heuristic H6 – IEEE 30 bus (DC-OPF)

7.2 Numerical results for the AC-OPF problem

7.2.1 IEEE 14 bus test case (AC-OPF)

In this section, the results of the DAPCA-EC applied to the dual of the AC-OPF problem (5.36) for the IEEE 14 bus test case [ZMS11, Uni] with 5 generators and 20 branches are presented without further comment as they are similar to the results discussed detailed in section 6.4 for the IEEE 57 bus test case.

IEEE 14 bus test case (AC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 575, \quad \text{dual: } 138,$$

- Accuracy $\epsilon = 1.37$ and scaling factor $s = 9$,
- Norm of optimal dual multipliers: $\|\Lambda^{\text{opt}}\| = 4.7064$,
- Lipschitz constant $L_c = 8.9064e9$ (5.40),
- Necessary number of iterations (Theorem 3.1.5): $k_{\text{fin}} = 161259$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -1.4549, \quad \text{upper bound: } 1.37, \tag{7.7}$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$0.3091, \tag{7.8}$$

- Primal gap at starting point $\Lambda^0 = 0 \in \mathbb{R}^{138}$: -137.8802,
- Constraint violation at starting point: $1.1549e3$,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9999^k$,
- Step-size parameters for DAPCA-EC: $\gamma = 1.1$ and $L_{-1} = 10^{-3} L_c$.

	β	PG	CV	TC	DC	MCTpA
1	0	-0.2079	0.0797	2.1e8 (100 %)	1.1e8 (100 %)	16.7
2	5e-9	-0.2095	0.0804	2.1e8 (100 %)	1.1e8 (100 %)	16.6
3	1e-8	-0.2006	0.0788	2.1e8 (100 %)	1.1e8 (100 %)	16.9
4	5e-8	-0.2069	0.0801	2.1e8 (99 %)	1.1e8 (99 %)	14.9
5	1e-7	-0.2170	0.0824	2.1e8 (99 %)	1.0e8 (98 %)	15.0
6	5e-7	-0.2013	0.0798	2.1e8 (97 %)	1.0e8 (94 %)	15.0
7	1e-6	-0.2008	0.0801	2.0e8 (95 %)	9.6e7 (90 %)	15.1
8	5e-6	-0.1198	0.0867	1.9e8 (91 %)	8.8e7 (82 %)	15.0

Table 7.13: Results of the DPCA-EC – IEEE 14 bus (AC-OPF)

DAPCA-EC stopping criterion for the IEEE 14 bus test case (AC-OPF):
For $k \geq 0$

if

$$|\text{primal gap at (6.17)}| \leq 0.2079 \text{ and constraint violation at (6.17) } \leq 0.0797 \qquad (7.9)$$

then

stop.

else

continue.

	β	NoI	TC	DC	MCTpA
1	0	1.6130e5	2.1e8 (100 %)	1.1e8 (100 %)	14.9
2	5e-9	1.6194e5	2.1e8 (100 %)	1.1e8 (100 %)	14.9
3	1e-8	1.5999e5	2.1e8 (99 %)	1.1e8 (99 %)	15.0
4	5e-8	1.6233e5	2.1e8 (100 %)	1.1e8 (100 %)	15.1
5	1e-7	1.6332e5	2.1e8 (100 %)	1.1e8 (100 %)	15.1
6	5e-7	1.6066e5	2.1e8 (96 %)	10.0e7 (93 %)	14.9
7	1e-6	1.6107e5	2.0e8 (95 %)	9.6e7 (90 %)	14.9
8	5e-6	1.6393e5	2.0e8 (93 %)	9.0e7 (84 %)	15.3

Table 7.14: Results of DPCA-EC with stopping criterion (7.9) – IEEE 14 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	4.0e4	43	5.4e8	3.5e6	2.5e8 (100 %)	5.3e7 (100 %)	1.4e8 (100 %)	49.6
2	5e-9	4.7e4	46	7.1e8	8.1e5	1.6e8 (64 %)	6.2e7 (117 %)	3.2e7 (23 %)	18.6
3	1e-8	4.6e4	45	6.5e8	3.4e5	1.3e8 (54 %)	6.0e7 (114 %)	1.4e7 (10 %)	12.7
4	5e-8	3.6e4	40	4.0e8	4.2e4	9.7e7 (40 %)	4.8e7 (90 %)	1.7e6 (1 %)	6.7
5	1e-7	3.0e4	36	2.8e8	3.3e4	7.9e7 (32 %)	3.8e7 (73 %)	1.3e6 (1 %)	5.4
6	5e-7	2.8e4	35	2.5e8	3.1e4	7.2e7 (29 %)	3.5e7 (66 %)	1.2e6 (1 %)	5.1
7	1e-6	5.0e5	47	7.9e8	3.4e7	2.7e9 (1090 %)	6.6e8 (1253 %)	1.4e9 (966 %)	441.1
8	5e-6	1.1e6	48	8.6e8	8.8e7	6.5e9 (2643 %)	1.5e9 (2818 %)	3.5e9 (2509 %)	1117.6

Table 7.15: Results of DAPCA-EC – IEEE 14 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	2.9e4	35	2.5e8	3.0e3	3.8e7 (100 %)	1.9e7 (100 %)	1.2e5 (100 %)	2.7
2	5e-9	2.9e4	35	2.5e8	3.0e3	3.8e7 (99 %)	1.9e7 (99 %)	1.2e5 (100 %)	2.7
3	1e-8	2.7e4	35	2.5e8	3.0e3	3.6e7 (93 %)	1.8e7 (93 %)	1.2e5 (100 %)	2.5
4	5e-8	2.7e4	35	2.5e8	3.0e3	3.5e7 (92 %)	1.7e7 (91 %)	1.2e5 (100 %)	2.5
5	1e-7	2.7e4	35	2.5e8	3.0e3	3.5e7 (92 %)	1.7e7 (90 %)	1.2e5 (100 %)	2.5
6	5e-7	2.8e4	35	2.5e8	3.0e3	3.5e7 (91 %)	1.7e7 (87 %)	1.2e5 (100 %)	2.7
7	1e-6	4.6e5	35	2.5e8	3.0e3	6.1e8 (1580 %)	3.0e8 (1582 %)	1.2e5 (100 %)	42.5
8	5e-6	1.0e6	35	2.5e8	3.0e3	1.4e9 (3603 %)	6.9e8 (3608 %)	1.2e5 (100 %)	95.8

Table 7.16: Results of DAPCA-EC with heuristic H1 – IEEE 14 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	4.9e4	45	6.5e8	4.2e3	6.5e7 (100 %)	3.2e7 (100 %)	1.7e5 (100 %)	4.5
2	5e-9	4.9e4	45	6.5e8	4.0e3	6.5e7 (100 %)	3.2e7 (100 %)	1.6e5 (96 %)	4.6
3	1e-8	2.7e4	35	2.5e8	3.0e3	3.6e7 (55 %)	1.8e7 (55 %)	1.2e5 (72 %)	2.6
4	5e-8	2.7e4	35	2.5e8	3.0e3	3.5e7 (55 %)	1.7e7 (54 %)	1.2e5 (72 %)	2.5
5	1e-7	2.7e4	35	2.5e8	3.0e3	3.5e7 (55 %)	1.7e7 (54 %)	1.2e5 (72 %)	2.5
6	5e-7	2.8e4	35	2.5e8	3.0e3	3.5e7 (54 %)	1.7e7 (51 %)	1.2e5 (72 %)	2.6
7	1e-6	4.6e5	42	4.9e8	6.2e3	6.1e8 (945 %)	3.0e8 (945 %)	2.5e5 (147 %)	42.7
8	5e-6	1.1e6	49	9.5e8	1.1e4	1.4e9 (2161 %)	7.0e8 (2162 %)	4.3e5 (260 %)	97.7

Table 7.17: Results of DAPCA-EC with heuristic H3 – IEEE 14 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	4.7e4	44	5.9e8	4.4e3	6.2e7 (100 %)	3.1e7 (100 %)	1.7e5 (100 %)	4.4
2	5e-9	4.4e4	43	5.4e8	3.9e3	5.8e7 (94 %)	2.9e7 (94 %)	1.6e5 (90 %)	4.1
3	1e-8	2.7e4	35	2.5e8	3.1e3	3.6e7 (58 %)	1.8e7 (57 %)	1.2e5 (71 %)	2.5
4	5e-8	2.7e4	35	2.5e8	3.0e3	3.5e7 (57 %)	1.7e7 (56 %)	1.2e5 (69 %)	2.5
5	1e-7	2.7e4	35	2.5e8	3.0e3	3.5e7 (57 %)	1.7e7 (56 %)	1.2e5 (69 %)	2.5
6	5e-7	2.8e4	35	2.5e8	3.0e3	3.5e7 (56 %)	1.7e7 (54 %)	1.2e5 (69 %)	2.6
7	1e-6	4.6e5	42	4.9e8	7.5e3	6.0e8 (976 %)	3.0e8 (976 %)	3.0e5 (171 %)	42.4
8	5e-6	1.0e6	45	6.5e8	1.6e4	1.4e9 (2230 %)	6.9e8 (2232 %)	6.3e5 (359 %)	95.5

Table 7.18: Results of DAPCA-EC with heuristic H6 – IEEE 14 bus (AC-OPF)

7.2.2 IEEE 30 bus test case (AC-OPF)

In this section, the results of the DAPCA-EC applied to the dual of the AC-OPF problem (5.36) for the IEEE 30 bus test case [ZMS11, Uni] with 6 generators and 41 branches are presented without further comment as they are similar to the results discussed detailed in section 6.4 for the IEEE 57 bus test case.

IEEE 30 bus test case (AC-OPF):

- Dimension of primal and dual variable space:

$$\text{primal: } 1227, \qquad \text{dual: } 310,$$

- Accuracy $\epsilon = 1.77$ and scaling factor $s = 13$,
- Norm of optimal dual multipliers: $\|\Lambda^{\text{opt}}\| = 6.5599$,
- Lipschitz constant $L_c = 3.3676e10$ (5.40),
- Necessary number of iterations (Theorem 3.1.5): $k_{\text{fin}} = 275870$,
- Bounds on primal gap at approximate solution (Theorem 3.1.5):

$$\text{lower bound: } -1.7918, \qquad \text{upper bound: } 1.77, \tag{7.10}$$

- Bound on constraint violation at approximate solution (Theorem 3.1.5):

$$0.2731, \tag{7.11}$$

- Primal gap at starting point $\Lambda^0 = 0 \in \mathbb{R}^{138}$: -177.3612,
- Constraint violation at starting point: $1.6085e3$,
- Threshold for event-triggered communication (6.2): $\Delta_k \approx \beta \cdot 0.9999^k$,
- Step-size parameters for DAPCA-EC: $\gamma = 1.2$ and $L_{-1} = 10^{-5} L_c$.

	β	PG	CV	TC	DC	MCTpA
1	0	-0.2736	0.0699	7.7e8 (100 %)	3.8e8 (100 %)	96.1
2	5e-9	-0.2560	0.0692	7.7e8 (100 %)	3.8e8 (100 %)	93.4
3	1e-8	-0.2548	0.0695	7.6e8 (100 %)	3.8e8 (99 %)	93.5
4	5e-8	-0.2554	0.0700	7.6e8 (99 %)	3.8e8 (98 %)	92.9
5	1e-7	-0.2581	0.0706	7.5e8 (98 %)	3.7e8 (96 %)	92.1
6	5e-7	-0.2573	0.0706	7.3e8 (95 %)	3.4e8 (89 %)	95.7
7	1e-6	-0.2293	0.0721	7.1e8 (93 %)	3.3e8 (86 %)	92.8
8	5e-6	0.1377	0.1002	6.7e8 (87 %)	2.9e8 (76 %)	92.8

Table 7.19: Results of the DPCA-EC – IEEE 30 bus (AC-OPF)

DAPCA-EC stopping criterion for the IEEE 30 bus test case (AC-OPF):

For $k \geq 0$

if

$$|\text{primal gap at (6.17)}| \leq 0.2736 \text{ and constraint violation at (6.17) } \leq 0.0699 \tag{7.12}$$

then

stop.

else

continue.

	β	NoI	TC	DC	MCTpA
1	0	2.7596e5	7.7e8 (100 %)	3.8e8 (100 %)	95.0
2	5e-9	2.7459e5	7.6e8 (99 %)	3.8e8 (99 %)	99.4
3	1e-8	2.7481e5	7.6e8 (99 %)	3.8e8 (99 %)	96.2
4	5e-8	2.7603e5	7.6e8 (99 %)	3.8e8 (98 %)	93.3
5	1e-7	2.7689e5	7.6e8 (99 %)	3.7e8 (97 %)	93.5
6	5e-7	2.7764e5	7.3e8 (95 %)	3.5e8 (90 %)	94.3
7	1e-6	2.7962e5	7.2e8 (94 %)	3.3e8 (87 %)	96.6
8	5e-6	3.2525e5	8.1e8 (105 %)	3.6e8 (94 %)	110.6

Table 7.20: Results of DPCA-EC with stopping criterion (7.12) – IEEE 30 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	5.4e4	45	1.2e9	9.8e6	1.1e9 (100 %)	1.5e8 (100 %)	8.0e8 (100 %)	138.5
2	5e-9	7.2e4	48	2.1e9	8.7e4	4.0e8 (37 %)	2.0e8 (132 %)	7.1e6 (1 %)	47.1
3	1e-8	5.3e4	45	1.2e9	6.3e4	3.0e8 (27 %)	1.5e8 (97 %)	5.2e6 (1 %)	34.6
4	5e-8	3.7e4	41	5.9e8	4.4e4	2.0e8 (19 %)	10.0e7 (66 %)	3.6e6 (0 %)	24.7
5	1e-7	3.7e4	41	5.9e8	4.5e4	2.1e8 (19 %)	9.9e7 (66 %)	3.7e6 (0 %)	24.1
6	5e-7	1.5e6	48	2.1e9	2.4e8	2.8e10 (2516 %)	4.1e9 (2708 %)	2.0e10 (2443 %)	3468.0
7	1e-6	1.9e6	48	2.1e9	3.3e8	3.8e10 (3420 %)	5.4e9 (3596 %)	2.7e10 (3353 %)	4734.9
8	5e-6	3.1e6	48	2.1e9	5.5e8	6.3e10 (5699 %)	8.7e9 (5815 %)	4.5e10 (5654 %)	7727.9

Table 7.21: Results of DAPCA-EC – IEEE 30 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	4.0e4	41	5.9e8	7.7e3	1.1e8 (100 %)	5.6e7 (100 %)	6.4e5 (100 %)	13.9
2	5e-9	3.7e4	41	5.9e8	7.7e3	1.0e8 (91 %)	5.1e7 (91 %)	6.4e5 (100 %)	12.7
3	1e-8	3.7e4	41	5.9e8	7.7e3	1.0e8 (91 %)	5.1e7 (90 %)	6.4e5 (100 %)	13.3
4	5e-8	3.7e4	41	5.9e8	7.7e3	1.0e8 (89 %)	4.9e7 (87 %)	6.4e5 (100 %)	12.4
5	1e-7	3.7e4	41	5.9e8	7.7e3	1.0e8 (89 %)	4.8e7 (86 %)	6.4e5 (100 %)	12.2
6	5e-7	1.5e6	41	5.9e8	7.7e3	4.3e9 (3795 %)	2.1e9 (3814 %)	6.4e5 (100 %)	518.1
7	1e-6	1.9e6	41	5.9e8	7.7e3	5.4e9 (4787 %)	2.7e9 (4811 %)	6.4e5 (100 %)	619.8
8	5e-6	3.0e6	41	5.9e8	7.7e3	8.4e9 (7488 %)	4.2e9 (7519 %)	6.4e5 (100 %)	986.9

Table 7.22: Results of DAPCA-EC with heuristic H1 – IEEE 30 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	5.8e4	45	1.2e9	9.0e3	1.6e8 (100 %)	8.1e7 (100 %)	7.4e5 (100 %)	20.8
2	5e-9	3.7e4	41	5.9e8	7.8e3	1.0e8 (63 %)	5.1e7 (63 %)	6.4e5 (86 %)	13.0
3	1e-8	3.7e4	41	5.9e8	7.8e3	1.0e8 (63 %)	5.1e7 (63 %)	6.4e5 (86 %)	12.8
4	5e-8	3.7e4	41	5.9e8	7.8e3	1.0e8 (62 %)	4.9e7 (61 %)	6.4e5 (86 %)	12.6
5	1e-7	3.7e4	41	5.9e8	7.8e3	1.0e8 (61 %)	4.8e7 (59 %)	6.4e5 (86 %)	12.4
6	5e-7	1.5e6	46	1.5e9	2.3e4	4.3e9 (2639 %)	2.1e9 (2648 %)	1.9e6 (253 %)	520.7
7	1e-6	1.9e6	45	1.2e9	2.7e4	5.4e9 (3321 %)	2.7e9 (3332 %)	2.2e6 (304 %)	628.2
8	5e-6	3.0e6	46	1.5e9	3.9e4	8.5e9 (5195 %)	4.2e9 (5210 %)	3.2e6 (430 %)	1023.4

Table 7.23: Results of DAPCA-EC with heuristic H3 – IEEE 30 bus (AC-OPF)

	β	NoI	L_k-Up	$L_k^{\max}$	NoCI	TC	DC	CC	MCTpA
1	0	7.7e4	48	2.1e9	1.0e4	2.2e8 (100 %)	1.1e8 (100 %)	8.5e5 (100 %)	28.3
2	5e-9	3.7e4	41	5.9e8	7.8e3	1.0e8 (48 %)	5.1e7 (47 %)	6.4e5 (75 %)	13.5
3	1e-8	3.7e4	41	5.9e8	7.8e3	1.0e8 (48 %)	5.1e7 (47 %)	6.4e5 (75 %)	12.8
4	5e-8	3.7e4	41	5.9e8	7.8e3	1.0e8 (47 %)	4.9e7 (46 %)	6.4e5 (75 %)	12.3
5	1e-7	3.7e4	41	5.9e8	7.8e3	1.0e8 (46 %)	4.8e7 (45 %)	6.4e5 (75 %)	13.2
6	5e-7	1.5e6	48	2.1e9	3.7e4	4.2e9 (1956 %)	2.1e9 (1961 %)	3.0e6 (358 %)	504.0
7	1e-6	1.9e6	49	2.6e9	4.5e4	5.4e9 (2503 %)	2.7e9 (2510 %)	3.7e6 (432 %)	654.2
8	5e-6	3.0e6	46	1.5e9	6.9e4	8.5e9 (3929 %)	4.2e9 (3936 %)	5.6e6 (665 %)	989.7

Table 7.24: Results of DAPCA-EC with heuristic H6 – IEEE 30 bus (AC-OPF)

List of Figures

Figures were created with the LaTeX package PGF/TikZ.

List of Tables

Bibliography

[BB03] A.G. Bakirtzis and P.N. Biskas. A Decentralized Solution to the DC-OPF of Interconnected Power Systems. *IEEE Transactions on Power Systems*, 18(3):1007–1013, Aug 2003.

[BCG11] S.R. Becker, E.J. Candes, and M.C. Grant. Templates for Convex Cone Problems with Applications to Sparse Signal Recovery. *Mathematical Programming Computation*, 3(3):165–218, 2011.

[Beu14] A. Beutelspacher. *Lineare Algebra: Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen*. Springer Spektrum. Springer Fachmedien Wiesbaden, 2014.

[Blu08] S.W. Blume. *Electric Power System Basics for the Nonelectrical Professional*. IEEE Press Series on Power Engineering. Wiley, 2008.

[BT89] D.P. Bertsekas and J.N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, 1989.

[BV00] A.R. Bergen and V. Vittal. *Power Systems Analysis*. Prentice Hall, 2000.

[CCW10] G.C. Calafiore, L. Carlone, and M. Wei. A Distributed Gradient Method for Localization of Formations Using Relative Range Measurements. In *IEEE International Symposium on Computer-Aided Control System Design (CACSD)*, pages 1146–1151, 2010.

[Cra12] V. Crastan. *Elektrische Energieversorgung 1 Netzelemente, Modellierung, Stationäres Verhalten, Bemessung, Schalt- und Schutztechnik*. Elektrische Energieversorgung. Springer Berlin Heidelberg, 2012.

[DMUH14a] F. Deroo, M. Meinel, M. Ulbrich, and S. Hirche. Distributed control design with local model information and guaranteed stability. In *19th IFAC World Congress*, volume 19, pages 4010–4017, 2014.

[DMUH14b] F. Deroo, M. Meinel, M. Ulbrich, and S. Hirche. Distributed stability tests for large-scale systems with limited model information. Extended version of [DMUH15], 2014.

[DMUH15] F. Deroo, M. Meinel, M. Ulbrich, and S. Hirche. Distributed stability tests for large-scale systems with limited model information. *IEEE Transactions on Control of Network Systems*, 2015.

[DUAH12] F. Deroo, M. Ulbrich, B.D.O. Anderson, and S. Hirche. Accelerated Iterative Distributed Controller Synthesis with a Barzilai-Borwein Step Size. In *IEEE 51st Annual Conference on Decision and Control (CDC)*, pages 4864–4870, 2012.

[DZG13] E. Dall'Anese, Hao Zhu, and G.B. Giannakis. Distributed Optimal Power Flow for Smart Microgrids. *IEEE Transactions on Smart Grid*, 4(3):1464–1475, 2013.

[Fis13] G. Fischer. *Lehrbuch der Algebra: Mit lebendigen Beispielen, ausführlichen Erläuterungen und zahlreichen Bildern*. Springer Fachmedien Wiesbaden, 2013.

[FKMN01] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting Sparsity in Semidefinite Programming via Matrix Completion I: General Framework. *SIAM Journal on Optimization*, 11(3):647–674, 2001.

[GK02] C. Geiger and C. Kanzow. *Theorie und Numerik Restringiererter Optimierungsaugaben*. Springer-Lehrbuch Masterclass. Springer Berlin Heidelberg, 2002.

[Gol04] M.C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs: Second Edition*. Annals of Discrete Mathematics. Elsevier Science, 2004.

[GS94] J.J. Grainger and W.D. Stevenson. *Power System Analysis*. McGraw-Hill Series in Electrical and Computer Engineering: Power and Energy. McGraw-Hill, 1994.

[Hig88] N.J. Higham. Computing a Nearest Symmetric Positive Semidefinite Matrix. *Linear Algebra and its Applications*, 103(0):103–118, 1988.

[Jab12] R.A. Jabr. Exploiting Sparsity in SDP Relaxations of the OPF Problem. *IEEE Transactions on Power Systems*, 27(2):1138–1139, 2012.

[JDR08] J.C. Jacobo and D. De Roure. A Decentralised DC Optimal Power Flow Model. In *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, pages 484–490, April 2008.

[KB00] B.H. Kim and R. Baldick. A Comparison of Distributed Optimal Power Flow Algorithms. *IEEE Transactions on Power Systems*, 15(2):599–604, 2000.

[KCD15] A. Kozma, C. Conte, and M. Diehl. Benchmarking Large-Scale Distributed Convex Quadratic Programming Algorithms. *Optimization Methods and Software*, 30(1):191–214, 2015.

[KKMY11] S. Kim, M. Kojima, M. Mevissen, and M. Yamashita. Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. *Mathematical Programming*, 129(1):33–68, 2011.

[LÖ4] J. Löfberg. YALMIP : a toolbox for modeling and optimization in MATLAB . In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[Lem10] M. Lemmon. Event-Triggered Feedback in Control, Estimation, and Optimization. In Alberto Bemporad, Maurice Heemels, and Mikael Johansson, editors, *Networked Control Systems*, volume 406 of *Lecture Notes in Control and Information Sciences*, pages 293–358. Springer London, 2010.

[LL99] S.H. Low and D.E. Lapsley. Optimization Flow Control - I. Basic Algorithm and Convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, Dec 1999.

[LL10] J. Lavaei and S.H. Low. Convexification of Optimal Power Flow Problem. In *48th Annual Allerton Conference on Communication, Control, and Computing*, pages 223–232, 2010.

[LL12] J. Lavaei and S.H. Low. Zero Duality Gap in Optimal Power Flow Problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

[LZT12] A.Y.S. Lam, Baosen Zhang, and D.N. Tse. Distributed Algorithms for Optimal Power Flow Problem. In *IEEE 51st Annual Conference on Decision and Control (CDC)*, pages 430–437, 2012.

[Mat14] *MATLAB and Statistics Toolbox Release 2014a*. The MathWorks Inc., Natick, Massachusetts, 2014.

[Mom01] J.A. Momoh. *Electric Power System Applications of Optimization*. Markel Dekker, 2001.

[MU14] M. Meinel and M. Ulbrich. A Distributed Algorithm with Event-Triggered Communication for Convex LMI-Constrained Optimization Problems. Preprint in preparation, Department of Mathematics, TU München, 2014.

[MUA14] M. Meinel, M. Ulbrich, and S. Albrecht. A class of distributed optimization methods with event-triggered communication. *Computational Optimization and Applications*, 57(3):517–553, 2014. Springer US.

[Nes05] Yu. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programmming*, 103(1):127–152, 2005.

[Nes13] Yu. Nesterov. Gradient methods for minimizing composite functions. *Mathematical Programming*, 140(1):125–161, 2013.

[NO09] A. Nedic and A. Ozdaglar. Cooperative Distributed Multi-Agent Optimization. In D.P. Palomar and Y.C. Eldar, editors, *Convex Optimization in Signal Processing and Communications*, pages 340–386. Cambridge University Press, 2009.

[NS08] I. Necoara and J. Suykens. Application of a Smoothing Technique to Decomposition in Convex Optimization. *IEEE Transactions on Automatic Control*, 53(11):2674–2679, 2008.

[PC06] D.P. Palomar and Mung Chiang. A Tutorial on Decomposition Methods for Network Utility Maximization. *IEEE Journal on Selected Areas in Communications*, 24(8):1439–1451, 2006.

[PMVDB05] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans. Usefulness of DC power Flow for Active Power Flow Analysis. In *Proceedings of the IEEE Power Engineering Society General Meeting*, volume 1, pages 454–459, 2005.

[SBG07] S. Samar, S. Boyd, and D. Gorinevsky. Distributed Estimation via Dual Decomposition. In *Proceedings of the European Control Conference*, pages 1511–1516, 2007.

[SFJ06] A. Speranzon, C. Fischione, and K.H. Johansson. Distributed and Collaborative Estimation over Wireless Sensor Networks. In *45th IEEE Conference on Decision and Control*, pages 1025–1030, 2006.

[ST06] J. Sun and L. Tesfatsion. DC Optimal Power Flow Formulation and Solution Using QuadProgJ. *Economics Department, Iowa State University*, 6014:1–36, 2006.

[TTT99] K.C. Toh, M.J. Todd, and R.H. Tütüncü. SDPT3 - a MATLAB software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.

[Uni] University of Washington - Department of Electrical Engineering. Power systems test case archive (1999). http://www.ee.washington.edu/research/pstca/.

[UU12] M. Ulbrich and S. Ulbrich. *Nichtlineare Optimierung*. Mathematik Kompakt. Springer Basel, 2012.

[WL09a] P. Wan and M.D. Lemmon. Distributed Network Utility Maximization using Event-triggered augmented Lagrangian methods. In *American Control Conference*, pages 3298–3303, 2009.

[WL09b] P. Wan and M.D. Lemmon. Distributed Network Utility Maximization using Event-triggered Barrier Methods. In *European Control Conference (ECC)*, pages 3785–3790, 2009.

[WL10] P. Wan and M.D. Lemmon. Optimal power flow in microgrids using event-triggered optimization. In *American Control Conference (ACC)*, pages 2521–2526, 2010.

[XBL06] L. Xiao, S. Boyd, and S. Lall. Distributed Average Consensus with Time-Varying Metropolis Weights, 2006.

[Yan81] M. Yannakakis. Computing the Minimum Fill-In is NP-Complete. *SIAM Journal on Algebraic Discrete Methods*, 2(1):77–79, 1981.

[ZC10] M. Zhong and C.G. Cassandras. Asynchronous Distributed Optimization with Event-Driven Communication. *IEEE Transactions on Automatic Control*, 55(12):2735–2750, 2010.

[Zha05] F. Zhang. *The Schur Complement and Its Applications*. Numerical Methods and Algorithms. Springer, 2005.

[ZMS11] R.D. Zimmerman and C.E. Murillo-Sanchez. MATPOWER 4.1 User's Manual, 2011.

[ZMST11] R.D. Zimmerman, C.E. Murillo-Sanchez, and R.J. Thomas. MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education. *Power Systems, IEEE Transactions on*, 26(1):12–19, 2011.