SPARSITY AND STRUCTURE EXPLOITING DIAGONALLY DOMINANT RELAXATION OF THE OPF PROBLEM

Laith Mubaslat

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

August 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science.

2020/08/29

Abstract

The Optimal Power Flow (OPF) is an optimization problem which tackles both the economy and physics of power systems operation. Due to high non-linearity in the power flow equations, the OPF problem is non-convex. Consequently, optimally solving for the OPF problem at a reasonable computational time presents a serious challenge.

Several approaches were presented to solve the OPF problem. These include local solvers, heuristic methods and the approximation of non-linear equations. However, these approaches either do not bound the true value of the objective function or are lacking in the trade-off they provide between solution time and quality.

As an alternative, convex relaxation techniques could be used to address this challenge. A convex relaxation is obtained by means of finding a convex representation of the problem's feasible space. As a natural byproduct of the convexity of the resulting problem, a wide array of convex optimization techniques could be utilized. Furthermore, the solution obtained presents a lower bound on the global solution of the original non-convex problem.

Several factors influence the tightness and scalability of convex relaxations. Those include the number and type of constraints used in the relaxation of the original non-convex problem. Most relaxations of the optimal power flow problem are based on second order conic or positive semidefinite type of constraints. Alternatively, in this dissertation we address the utilization of the linearly representable diagonally dominant cone in relaxing the optimal power flow problem.

First, we investigate the diagonally-dominant-sum-of-squares relaxation of the problem. We evaluate the reasons behind its poor optimality gaps and scalability issue. We demonstrate that diagonal dominance could be utilized in creating a similar, yet tighter relaxation. The relaxation we propose is based on the semidefinite relaxation of the problem.

This dissertation then follows to improve the tractability of the aforementioned relaxation. We achieve that by an investigation into the optimal exploitation of the sparsity and structure of the OPF problem.

Several methods exist for the exploitation of sparsity in semidefininte programming. Specifically, chordal decomposition has been applied with great success to improve the tractability of the semidefinite relaxation of the optimal power flow problem. Accordingly, we investigate the utilization of chordal decomposition in improving the diagonal dominance based relaxation proposed in this thesis.

We find that the direct exploitation of sparsity requires a number of linear inequalities that scales linearly with the size of the problem. Alternatively, chordal decomposition introduces equality and inequality constraints into the problem which needlessly increases its computational demand. We prove the direct exploitation of sparsity to be more beneficial in the case of a relaxation similar to that of this dissertation. Additionally, we exploit the structure of the problem in further reducing the number of linear inequalities by half. We further suggest two more relaxations based on the empirical results of the improved relaxation proposed.

Résumé

Le flux de puissance optimal est un problème d'optimisation qui concerne à la fois l'économie et la physique du fonctionnement des systèmes électriques. En raison de la forte non-linéarité des équations de flux de puissance, le problème de flux de puissance optimal est non convexe. Par conséquent, la solution optimale du problème de flux de puissance optimale à un moment de calcul raisonnable représente un grand défi.

Plusieurs approches ont été présentées pour résoudre le problème du flux de puissance optimale. Elles comprennent des solveurs locaux, des méthodes heuristiques et l'approximation d'équations non linéaires. Toutefois, ces approches ne limitent pas la valeur réelle de la fonction objective ou ne permettent pas de trouver un compromis entre le temps de solution et la qualité des résultats.

Comme alternative, des techniques de relaxation convexe pourraient être utilisées pour relever ce défi. Une relaxation convexe est obtenue en trouvant une représentation convexe de l'espace réalisable du problème. Comme sous-produit naturel de la convexité du problème résultant, un large éventail de techniques d'optimisation convexes pourrait être utilisé. En outre, la solution obtenue présente une limite inférieure à la solution globale du problème non convexe originale.

Plusieurs facteurs influencent l'étroitesse et l'extensibilité des relaxations convexes. Parmi ceux-ci figurent le nombre et le type de contraintes utilisées dans la relaxation du problème originale non convexe. La plupart des relaxations du problème de flux de puissance optimal sont basées sur des contraintes de type conique du second ordre ou semi-définies positives. Alternativement, dans cette thèse, nous abordons l'utilisation du cône diagonalement dominant linéairement représentable dans la relaxation du problème de flux de puissance optimal.

Tout d'abord, nous étudions la relaxation du problème par la somme des carrés diagonalement dominante. Nous évaluons les raisons de ses faibles écarts d'optimalité et de son problème de scalabilité. Nous démontrons que la dominance diagonale pourrait être utilisée pour créer une relaxation similaire, mais plus étroite. La relaxation que nous proposons est

basée sur la relaxation semi-définie du problème.

Cette thèse suit ensuite pour améliorer la tractabilité de la relaxation susmentionnée. Nous y parvenons par une étude de l'exploitation optimale de la structure et de sparsité du problème de flux de puissance optimale.

Plusieurs méthodes existent pour l'exploitation de sparsité dans la programmation semi-définie. En particulier, la décomposition cordale a été appliquée avec beaucoup de succès pour améliorer la traçabilité de la relaxation semi-définie du problème de flux de puissance optimal. En conséquence, nous étudions l'utilisation de la décomposition cordale dans l'amélioration de la relaxation basée sur la dominance diagonale proposée dans cette thèse.

Nous constatons que l'exploitation directe de la sparsité nécessite un certain nombre d'inégalités linéaires qui s'échelonnent linéairement en fonction de la taille du problème. Alternativement, la décomposition cordale introduit des contraintes d'égalité et d'inégalité dans le problème, ce qui augmente inutilement sa demande de calcul. Nous prouvons que l'exploitation directe de la sparsité est plus bénéfique dans le cas d'une relaxation similaire à celle de cette thèse. En outre, nous exploitons la structure du problème en réduisant encore de moitié le nombre d'inégalités linéaires. Nous suggérons en outre deux autres relaxations basées sur les résultats empiriques de la relaxation améliorée proposée.

Acknowledgments

Throughout my father's severe illness and all the hardships we have been through, my family had given me every support and pushed me towards taking every opportunity. As such, I would like to begin by acknowledging the unwavering support of my mother, my late father and brother for which I am forever grateful.

I would like to express my sincere appreciation to Prof. Francois Bouffard for his trust, guidance, support and understanding throughout my work. I gratefully acknowledge the funding and constant support I had received from Alghurair Foundation for Education (AGFE), Professor Bouffard, the Natural Sciences and Engineering Research Council of Canada (NSERC) and McGill University towards the completion of my Master's degree.

I would like to give my heartfelt thanks to my friends and colleagues Jneid Jneid, Elio Abi-Younes and Hamzeh Salameh for their friendship and constant support. I would also like to recognize the invaluable insights towards research I had received during my internship at Airbus by Mr. Franck Flourens, Mr. Michael Bowery and Mr. Frederic Therond.

Preface

I was the major contributor for all the chapters in this dissertation where I was responsible for the totality of dissertation composition.

All the text and equations that are taken from previously published articles were cited in the dissertation. The corresponding sources of the tools that were adopted in this dissertation were cited where appropriate.

Professor Bouffard contributed by supervising me and providing feedback on the technical contents and results, as well as giving best practice guidelines in academic writing.

Contents

List of Figures

List of Tables

List of Acronyms

AC	Alternating Current
BFM	Branch Flow Model
BIM	Bus Injection Model
DC	Direct Current
DD	Diagonally Dominant
DSOS	Diagonally-Dominant-Sum-Of-Squares
ED	Economic Dispatch
GA	Genetic Algorithms
LP	Linear Programming
MILP	Mixed Integer Linear Program
OPF	Optimal Power Flow
PP	Polynomial Program
PSO	Particle Swarm Optimization
QC	Quadratic Convex
QCQP	Quadratically Constrained Quadratic Program
SOC	Second Order Conic
SOCP	Second Order Cone Programming
SOS	Sum-Of-Squares
SDD	Scaled Diagonally Dominant
SDP	Semidefinite Programming

SDSOS	Scaled-Diagonally-Dominant-Sum-Of-Squares
UC	Unit Commitment

Nomenclature

$\mathbf{Y} = \mathbf{G} + j\mathbf{B}$	Power system admittance matrix
$\mathcal{N}$	Set of nodes in the power system
$\mathcal{N}_g$	Set of nodes with generation in the power system
L	Set of Edges in the power system
$S_{Di} = P_{Di} + jQ_{Di}$	AC power demand at bus i
$V_i = V_{di} + jV_{qi}$	AC voltage at bus i
$S_{Gg} = P_{Gg} + jQ_{Gg}$	AC power generation of generating bus g
$P_{inj,i}$	Active power injections at bus i
$Q_{inj,i}$	Reactive power injections at bus i
$f_g^q(.)$	Quadratic active power cost function of generator g
c_{g0}	Constant cost coefficient for active power generation of generator g
c_{g1}	Linear cost coefficient for active power generation of generator g
c_{g2}	Quadratic cost coefficient for active power generation of generator g
$S_{lm} = P_{lm} + jQ_{lm}$	AC power flow between busses l and m

S_n	Real symmetric matrices of dimension $n \times n$
P_n	Positive semidefinite symmetric matrices of dimension $n \times n$
$\mathbb{R}$	Real numbers
$\mathbb{R}^+$	Non-negative real numbers
$\mathbb{R}^n$	Real vectors of dimension n
$\mathbb{R}^{n\times n}$	Real matrices of dimension $n \times n$
A_z, C	Data matrices of a real positive semidefinite program
b	Data vector of the real positive semidefinite program
$\mathcal{X}$	Decision variable of the primal real semidefinite program
$\mathcal{A}$	Decision variable of the dual real semidefinite program
DD_n	Diagonally dominant matrices of dimension $n \times n$
SDD_n	Scaled diagonally dominant matrices of dimension $n \times n$
$\mathcal{U}_{n,z}$	Vectors in $\mathbb{R}^n$ with at most z non-zeros which are either $+1$ or -1
DD_n^*	The dual cone of DD_n
$\mathbb{C}^n$	Complex vectors of dimension n
$\mathbf{V}$	Complex voltage vector in $\mathbb{C}^n$
$\mathbf{X} = [Re(\mathbf{V}), Im(\mathbf{V})]$	Real voltage vector in $\mathbb{R}^{2n}$
e_i	ith standard basis vector in $\mathbb{R}^n$
b_{lm}	Shunt susceptance of the edge connecting busses l and m
y_{lm}	Admittance of the edge connecting busses l and m

M_i	Matrix employed in the reformulation of the voltage magnitude constraints
$Y_i, \mathbf{Y}_i, \bar{\mathbf{Y}}_i$	Matrices employed in the reformulation of the bus active and reactive power injection constraints
$Y_{lm}, \mathbf{Y}_{lm}, \bar{\mathbf{Y}}_{lm}$	Flow matrices of the edge connecting busses l and m
$f(x), g_z(x), p(x)$	Polynomials in x
$\mathcal{G}$	A set of polynomials $\{g_z(x) : z = 1, ..., w\}$
$S_{\mathcal{G}}$	A basic closed semi-algebriec set defined by $\mathcal{G}$
$\mathcal{P}_d(S_{\mathcal{G}})$	Cone of non-negative polynomials of degree at most d over the set $S_{\mathcal{G}}$
σ_0, σ_z	SOS polynomials
$z(x, d)$	Polynomial in x of degree d
$\mathbf{Q}$	Gram matrix
$SOS_{n,2d}$	Cone of n variable SOS polynomials of degree at most $2d$
SOS_r	Cone of SOS polynomials of degree at most r
SOS_2	First order in the SOS hierarchy
$\bar{\lambda}_i, \bar{\gamma}_i, \bar{\mu}_i$	Upper bound inequality constraints Lagrange multipliers
$\underline{\lambda}_i, \underline{\gamma}_i, \underline{\mu}_i$	Lower bound inequality constraints Lagrange multipliers
$\hat{a}_{lm}, \hat{b}_i, \hat{c}_{lm}, \hat{d}_{lm}$	Equality constraints Lagrange multipliers
$\Lambda(x)$	Polynomial in x of order at most $2d$

$\mathbf{Q}_{\mathbf{X}}, \mathbf{Q}_{P_{Gg}}, \mathbf{Q}_{(P_{lm},Q_{lm})}$	Gram matrices
$m(x)$	Monomial in x
$DSOS_{n,2d}$	Cone of n variable Diagonally Dominant Sum-Of-Squares polynomials of degree at most $2d$
$SDSOS_{n,2d}$	Cone of n variable Scaled Diagonally Dominant Sum-Of-Squares polynomials of degree at most $2d$
$PSD_{n,2d}$	Cone of n variable non-negative polynomials of degree at most $2d$
$\mathbf{W} = \mathbf{X}\mathbf{X}^T$	Voltage product matrix in $\mathbb{R}^{2n \times 2n}$
$\mathcal{L}$	Set of transmission Lines in the power system
$\mathbb{G}_i^{pw}$	Generators with a piece-wise linear cost function at bus i
$\mathbb{G}_i^{q}$	Generators with a quadratic cost function at bus i
$\mathbb{G}$	All generators in the power system
$f_g^{pw}(.)$	Piece-wise linear active power cost function of generator g
m_{gt}, a_{gt}, b_{gt}	Piece-wise linear generation cost coefficients (for $t \in [1, r]$ segments)
$S_k = P_k + jQ_k$	AC power flow at line k
τ_k	Ideal transformer tap ratio magnitude at line k
θ_k	Ideal transformer phase shift angle at line k
f_i	ith standard basis vector in $\mathbb{R}^{2n}$
b_k	Shunt susceptance at k
$Z_{k_u}, \bar{Z}_{k_u}$	Flow matrices of the of line k at bus u

λ_i	Locational marginal price at bus i
ζ_{gt}	Lagrange multiplier for line segment t of generator g with a piece-wise linear cost function
$\mathcal{C}_q$	Maximal clique q
$\mathcal{A}_{\mathcal{C}_q}$	Sub-matrix of the $\mathcal{A}$ corresponding to maximal clique q
$E_{\mathcal{C}_q}$	Index matrix
$G(v, \epsilon)$	Un-directed graph of vertices v and edges ϵ
$\mathbb{C}^{n\times n}$	Complex matrices of dimension $n \times n$
$\hat{A}_z, C$	Data matrices of the complex positive semidefinite program
$\hat{b}$	Data vector of the complex positive semidefinite program
Ψ	Decision variable of the primal complex Semidefinite program

Chapter 1

Introduction

1.1 Background and Motivation

The physics of power systems and their stability require that generation-demand balance be met at all time. For that purpose, fast intervention mechanisms are introduced into the operation of power systems. Due to the serious ramifications of failure in power system operation, these mechanisms often prioritize system integrity above all else, including economy. Accordingly, the improper introduction of such mechanisms could significantly strain the economics of power systems. Cyclical changes in demand, the introduction of renewable energy, protection against contingencies and asset congestion are all some of the factors crucial to power systems operation. Therefore, and considering the high cost of operating the system in a non-economic fashion, it is paramount to properly navigate the trade-off between the different operational aspects and the economy of power systems. Accordingly, the balance of energy generation and demand is one of the most challenging issues in power systems from the perspectives of both operation and economy [2].

Power system optimization aims at the proper integration of the different conflicting factors in power systems. The field includes a variety of problems which differ in both complexity and scope. The most common problem formulations aim at reducing the overall cost in the power system with respect to certain operational considerations. Economic Dispatch (ED), Unit Commitment (UC) and the Alternating Current Optimal Power Flow (AC-OPF) problems are some of the most common in this field. Below, we will provide an abstraction of these problem formulations, gradually progressing into the topic of this dissertation.

ED is one of the simplest variants in power system optimization. In summary, it relates to the problem of matching supply with demand for a given number of generators during real-time operation. These generators operate between an individual lower and upper generation bound. Accordingly, this problem assumes the availability of all generators and that they all operate under the same marginal cost. It also neglects all other operational constraints. Despite not being that representative of the operation of power systems, this problem reduces to solving a linear program and thus is significantly scalable and amenable to real-time solution. On the other hand, UC could be considered as a more complicated variant of ED. In UC, generator selection (on/off status) is included in the problem definition. Accordingly, the simplest definition of this problem boils down to solving a non-convex Mixed Integer Linear Program (MILP) [3].

There are many variants of the aforementioned problems. For the sake of consistency we will confine our discussion to the definitions introduced earlier. Introduced in the work of [4], the AC-OPF problem relates to solving an ED problem with the inclusion of operational and network constraints. The problem introduces a much needed link between the economical and operational aspects of power systems. In view of the inherent non-convexity of the power flow equations, that link results in the non-convexity of the problem's feasible space. Due to the highly non-convex nature of the AC-OPF problem, obtaining a solution is a computationally expensive process. Accordingly, the non-convexity of the problem presents a serious challenge for its implementation in real time operation. Local and heuristic solution techniques are utilized in solving this non-convex problem albeit with some limitations to scalability and solution optimality [5, 6].

Power systems need to react to changes in a timely manner, as well as operate at the most economically advantageous settings for a given scenario. The Direct Current Optimal Power Flow (DC-OPF) is an approximation of the AC-OPF problem. The DC-OPF approximation is obtained by making a number of simplifying assumptions. These assumptions are derived from the usual operating conditions of power systems. For one, the power system is assumed to work close to unity per unit voltage magnitudes throughout the entire system. Additionally, the difference in voltage angles is assumed to be so small that the problem structure could be further simplified. As a consequence of these assumptions, this formulation could only solve for real powers in a given power system [1]. DC-OPF reduces to solving a linear program and is an extremely scalable variant of the OPF problem. However, the DC-OPF leaves much to be desired. While it tackles the scalability issue of its parent formulation, it lacks guarantees and bounds on the obtained objective value. Additionally, the solution of this approach is not one of the original problem and, as a result, could be infeasible [2].

A relaxation is an technique by which a bounded solution could be obtained for an otherwise hard to solve optimization problem. The earlier solution is obtained by solving for a simplified optimization problem (i.e., a relaxation). The relaxation itself is obtained by either eliminating or replacing complicating constraints in the original problem. Convex relaxations in particular relate to when this process results in a convex optimization problem.

To circumvent the earlier limitations, convex relaxations have been applied to the OPF problem. A convex hull of a non-convex set, is a convex set which includes the non-convex set [7]. Convex relaxations are in essence formulations of the OPF problem which replace the non-convex set representing the AC-OPF by a convex hull [8]. The hull is defined based on the manner by which a given relaxation is obtained. Therefore, the problem is transformed

[1] This assumption allows for the replacement of trigonometric functions by the first term of their respective Taylor series expansion.

into one that closely resembles the original. The resulting convex problem could be solved to global optimality, which is not the case for the original non-convex formulation. As such, this approach provides a lower bound on the objective value of the original problem. It follows that the end goal of convex relaxations is to globally solve the non-convex problem. Optimally, we would like a relaxation to also provide a solution (in terms of the decision variables) that is feasible with regards to the operation of the power system.

The research on convex relaxations has been targeting three major aspects. The first is the tightness of the relaxation, where tightness represents the fidelity of the bound obtained, relative to the true objective value. The second aspect is scalability, while the third relates to the possibility of extracting feasible operating variables from a given relaxation [9]. This thesis addresses both the tightness and scalability aspects of an OPF convex relaxation.

In summary, The problem that this thesis tackles is the non-convexity of the OPF problem. We approach this problem by obtaining a convex representation of the original non-convex feasible space. Our work is based on a pre-existing linear relaxation which fails at producing acceptable bounds [1]. Using the underlying concept of diagonal dominance, addressed in Chapter 3, we attempt to create a tighter and a more scalable variant.

1.2 Thesis Statement and Contribution

Thesis statement

In this dissertation, we explore the first order Diagonally Dominant Sum-Of-Squares (DSOS) [10] relaxation of the AC-OPF problem. This relaxation suffers from two important limitations. First, it demonstrates poor bounds. The second limitation lies in the scalability of the relaxation, which is demonstrated by how it is incapable of solving cases as large as a 300 bus system [1].

We begin by an investigation into the reasons behind the poor optimality gaps demonstrated by the aforementioned relaxation. We then move to remedy the cause. In our work, we utilize the same underlying concept of the DSOS relaxation (i.e. diagonal dominance). We then follow by an exploration into approaches that could help improve our relaxation's scalability. For that purpose we investigate how to best utilize the inherent sparsity of power systems in the making of a more tractable relaxation.

Contribution

Our contribution is threefold. We first manage to significantly improve the tightness of the first order DSOS relaxation. Second, we investigate the utilization of chordal decomposition in our relaxation. We then demonstrate how this approach is, contrary to belief, counter-productive in the case of problems defined in this manner. Following that, we exploit the structure of the problem in further reducing the size of our relaxation, and propose a slightly faster relaxation based on empirical findings.

1.3 Organization

This section describes the organization of this dissertation. In Chapter 2, we present the classical formulation of the AC-OPF problem. This chapter further discusses the literature surrounding the non-convexity issue of the problem and its convex relaxations. Chapter 3 presents the mathematical preliminaries, as well as a detailed review of relevant OPF convex relaxations. In Chapter 4, we present a way for tightening the first order DSOS relaxation of the OPF problem. An investigation into how to best exploit sparsity in improving the tractability of our relaxation follows in Chapter 5. In Chapter 5, we discuss the merit of chordal decomposition and structure in improving our relaxation. Furthermore, this chapter provides a new slightly faster relaxation based on empirical findings derived from the relaxation proposed in Chapter 4. We conclude our work in Chapter 6 which summarizes the final findings, as well as possible future directions of our work.

Chapter 2

Literature Review

This chapter reviews the literature relevant to the AC-OPF problem. The first section provides a background on the problem and introduces its classical formulation. The second reviews the non-convexity issue of the problem. It further draws a canvas of the available literature in the area of OPF convex relaxations. An outline of the work done in this thesis in the context of the literature is also provided.

2.1 Optimal Power Flow (OPF)

2.1.1 Introduction

Optimal Power Flow (OPF), Economic Dispatch (ED) and Unit Commitment (UC) are all example realizations of resource allocation in power system optimization. These realizations relate to the proper allocation (i.e. dispatch) of generation resources so that a predefined cost is minimized for a given power network [3].

Load flow is the set of non-linear equations that relate the power flow in a power system to network parameters (e.g., line reactances), control and state variables (e.g., voltages) and nodal complex power injections [2]. An optimal operating point is determined in AC-OPF with the incorporation of load flow equations and other operating constraints. Therefore, the identifying (relative to its peers) aspect of the OPF problem is the inclusion of the complicating network constraints. Non-convexity is introduced into the OPF problem as a natural artifact of that inclusion [11]. Having said that, the OPF problem is considered as one of the most challenging optimization problems in the field of power systems.

Numerous aspects influence the operation of power systems. Accordingly, many formulations of the AC-OPF problem have been adapted to account for security (security constrained OPF)[12], uncertainty (e.g. robust or stochastic OPF) [13] and many of the other conditions detrimental to power system operation.

The AC-OPF, which is the focus of this thesis, is the paradigm that deals with the exact description of a power system's physical quantities. Accordingly, this paradigm makes up a basis on top of which all of the aforementioned considerations are adapted. Although the AC-OPF is the most true realization to the steady-state analysis of power systems , this paradigm is seldom utilized. Due to its computational intractability [14, 15] and lack of guarantee for optimality [16], cheaper and less accurate variations are leveraged in the work of both academia and industry (such as the linear DC-OPF approximation [17]).

2.1.2 Classical OPF

This subsection serves to introduce the classical AC-OPF for an n bus arbitrary power network with an admittance matrix $\mathbf{Y} = \mathbf{G} + j\mathbf{B}$ [18]. This formulation is relaxed into an SDP in the work of [19]. The SDP relaxation of [19] was further expanded in [20, 21] to incorporate the more practical considerations of parallel lines, multiple generation units at a node and piece-wise linear generator cost functions.

Notation

The following notation will be used throughout the second and third chapters. The fourth chapter will utilize the more generalized representation introduced in [20]. An n bus power system is modeled as an un-directed graph with sets of vertices $\mathcal{N}$ and edges L. It follows that $\mathcal{N}$ has the following definintion

$$\mathcal{N} := \{1, 2, 3, ..., n\} \tag{2.1}$$

Each vertex corresponds to a bus $i \in \mathcal{N}$ and each edge a line $(l, m) \in L$. Since not all busses contribute to the generation of power in the network, we define the subset $\mathcal{N}_g \subseteq \mathcal{N}$ as the set of generator busses. Consider an arbitrary bus $i \in \mathcal{N}$. The apparent power demand (S_{Di}) and complex voltage (V_i) at bus i are defined as

$$S_{Di} = P_{Di} + jQ_{Di} \tag{2.2}$$

$$V_i = V_{di} + jV_{qi} \tag{2.3}$$

Similarly, at a generator bus $g \in \mathcal{N}_g$, the apparent power generation is defined as

$$S_{Gg} = P_{Gg} + jQ_{Gg} \tag{2.4}$$

Thermal limits are of great importance to the operation of power systems [22]. For a line $(l, m) \in L$, these limits are imposed through the line's apparent power flow which is denoted as S_{lm}. Power flow equations capture the relationship between the power injections and voltages in the network. Therefore, these equations are utilized in solving for the power mismatch as a function of the voltages. For the utilization of the power flow equations in rectangular coordinates [1], the rectangular voltage magnitudes should satisfy the following constraint

$$|V_i|^2 = V_{di}^2 + V_{qi}^2 \tag{2.5}$$

Accordingly, the power flow equations in rectangular coordinates can be written as

[1]The resulting power flow equations are polynomials in the real and imaginary parts of the voltage. As such, OPF in rectangular coordinates could be treated as a polynomial programming problem [9].

$$P_{inj,i} = V_{di}\sum_{h=1}^{n}(\mathbf{G}_{ih}V_{dh} - \mathbf{B}_{ih}V_{qh}) + V_{qi}\sum_{h=1}^{n}(\mathbf{B}_{ih}V_{dh} + \mathbf{G}_{ih}V_{qh}) \tag{2.6}$$

$$Q_{inj,i} = V_{di}\sum_{h=1}^{n}(-\mathbf{G}_{ih}V_{dh} - \mathbf{B}_{ih}V_{qh}) + V_{qi}\sum_{h=1}^{n}(\mathbf{B}_{ih}V_{dh} - \mathbf{G}_{ih}V_{qh}) \tag{2.7}$$

As shown in the two equations above, power flow provides the required link between the operating parameters of the network. However, the resulting problem is non-convex. By virtue of the power flows being defined in terms of bi-linear and quadratic voltage products, the power flow equations are non-linear. Consequently, the non-convexity of the problem comes as a natural result of the inherent non-linearity of these equations [9].

The OPF solves for the minimum value of an objective function. In our work, we solve for the commonly used objective; the total cost of generation. This formulation considers generators with quadratically defined cost functions. Accordingly, the total cost of generation is the sum of the individual cost for each generator bus. For a generator bus g, let the quadratic cost function f_g^q be defined as follows

$$f_g^q(P_{Gg}) = c_{g2}(P_{Gg})^2 + c_{g1}(P_{Gg}) + c_{g0} \tag{2.8}$$

The classical OPF formulation includes thermal limits, generator capabilities, line thermal limits and limits on the voltages of each bus. It additionally accounts for the mismatch in both active and reactive power. The upper and lower bounds of a variable are denoted by their respective superscripts. Accordingly, the classical OPF formulation (in rectangular coordinates) can be cast as the following optimization problem

$$\min \sum_{g \in \mathcal{N}_g} f_g^q(P_{Gg}) \tag{2.9a}$$

subject to

$$P_{Gg}^{min} \leq P_{Gg} \leq P_{Gg}^{max} \qquad \forall g \in \mathcal{G} \tag{2.9b}$$

$$Q_{Gg}^{min} \leq Q_{Gg} \leq Q_{Gg}^{max} \qquad \forall g \in \mathcal{G} \tag{2.9c}$$

$$|V_i^{min}|^2 \leq |V_i|^2 \leq |V_i^{max}|^2 \qquad \forall i \in \mathcal{N} \tag{2.9d}$$

$$-S_{lm}^{max} \leq S_{lm} \leq +S_{lm}^{max} \qquad \forall (l,m) \in L \tag{2.9e}$$

$$\sum_{g \in \mathcal{G}_i} P_{Gg} - P_{Di} = V_{di} \sum_{h=1}^{n} (\mathbf{G}_{ih} V_{dh} - \mathbf{B}_{ih} V_{qh}) + V_{qi} \sum_{h=1}^{n} (\mathbf{B}_{ih} V_{dh} + \mathbf{G}_{ih} V_{qh}) \qquad \forall i \in \mathcal{N} \tag{2.9f}$$

$$\sum_{g \in \mathcal{G}_i} Q_{Gg} - Q_{Di} = V_{di} \sum_{h=1}^{n} (-\mathbf{G}_{ih} V_{dh} - \mathbf{B}_{ih} V_{qh}) + V_{qi} \sum_{h=1}^{n} (\mathbf{B}_{ih} V_{dh} - \mathbf{G}_{ih} V_{qh}) \quad \forall i \in \mathcal{N} \tag{2.9g}$$

This adaptation of the OPF problem is constructed using the voltage-based Bus Injection Model (BIM) formulation of the load flow equations. Alternative formulations include that of Distflow, which is valid for radial systems and otherwise known as the Branch Flow Model (BFM) [23]. BFM, as the name suggests, relates to quantities in branches rather than values at nodes of the network. The work in [24, 25] relate to the equivalence relation between BIM and BFM when applied to relaxations over the non trivial case of mesh networks.

2.2 OPF Non-Convexity and Convex Relaxations

2.2.1 Non-convexity

As seen earlier, the OPF problem is constructed using non-linear power flow equations. Therefore, any discussion of the convexity of the feasible set of said problem has to relate to the structure of these equations. Convexity in OPF is not the general case. What might not be so obvious in rectangular coordinates is more so in polar coordinates. Taking a look at the power flow equations in polar coordinates, one can clearly see the apparent periodic behaviour, with a period of 2π in the angles of the voltage variables [11].

In some specific cases, convexity can be established. For a constrained network the set of active and reactive power injections is convex in rectangular coordinates [26]. The same can be said for a constant voltage, small angle difference formulation of the problem [27]. As such, a lot of effort was made towards studying the feasible sets of power systems and determining the conditions for which the sets are convex. Accordingly, several techniques were developed to identify and characterize non-convexity in the OPF problem [28–32].

Due to the problem's non-convexity, the possibility for multiple local solutions exists. The number of these local solutions and their distribution with respect to one another is highly dependant on the network [16]. It follows that the OPF problem is NP-hard [14][15]. Several classical local optimization techniques are utilized in solving this problem [5, 33, 34]. The application of these techniques requires a number of simplifying assumptions. Consequently, provided that a local solutions exists for the instance of this problem being solved, the solution obtained would be local [16]. Additionally, the convergence of these techniques to a specific local solution is highly dependent on initialization. As such, local optimization techniques are ill-equipped for the global solution of the OPF problem. That is the case specifically for instances characterized by multiple local solutions.

Heuristic algorithms, such as genetic algorithms (GE) and particle swarm optimization (PSO) to name a few, are employed in globally solving the OPF problem. By design, heuristic algorithms can bypass the local optimality issue of non-convex problems. However, these algorithms suffer from poor scalability which in turn limits their practical implementation to large systems. Furthermore, whereas local optimization techniques are dependent on the starting point of the solution process, these techniques are highly influenced the choice of hyperparameters [6].

Alternatively, and since the guarantees of convex optimization techniques are not valid for a non-convex problem, the approach in [35] was developed to certify the global optimality of a given local solution.

2.2.2 Relaxations

For a given non-convex optimization problem, convex relaxations are the different convex sets utilized in the convex representation of the problem's non-convex feasible space [8]. That representation, which can be solved using convex optimization techniques, is then used to find an outer bound on the problem's global objective minimum or maximum.

Two measures of quality are used in the comparison between different relaxations. These measures are scalability and tightness. The notion of a relaxation's scalability of interest to our work is a reflection of its complexity in terms of the number and type of constraints. On the other hand, tightness relates to how true a relaxation is to the feasible space being relaxed. Tightness is therefore quantified using the difference between a relaxation's outer bound and the true objective value. However, it is not always possible to obtain the true objective value for a given instance. As such, the best solution obtained by heuristic methods (or the best bound among the different relaxation techniques) could be utilized instead. The following relation is used for that purpose

$$Optimality\,Gap = \frac{TrueObjectiveValue - Relaxation}{TrueObjectiveValue} \tag{2.10}$$

Over the past decade, a number of relaxations have been proposed for the AC-OPF problem. These relaxations were motivated by the non-convexity inherent to the OPF problem. The monograph [9], provides a detailed treatment of the subject.

A relaxation of the OPF problem aims at finding a convex representation of the problem which can give a tight enough answer in polynomial time. Another aspect relevant to the research in this area is whether a feasible solution (i.e., feasible operating variables) can be efficiently extracted from a given relaxation. Accordingly, even for a relaxation resulting in zero duality gap, the solution it provides might not be an optimal solution in the decision variables of the problem [9].

We should note that convex relaxations compete not only with their counterpart relaxations, but also with the different local and heuristic solvers which could in practice either provide more accurate solutions or scale much more favourably. Accordingly, [36] suggests an approach that uses a hybrid of a local solver and convex relaxations in solving the OPF problem.

The following is a review of the most prominent relaxations of the OPF problem. A more detailed treatment to the relaxations most relevant to our work follows in Chapters 3 and 4.

Lasserre Hierarchy: Based on the work in [37], a Polynomial Program (PP) can be solved by means of an iterative hierarchy. Each iterative step (identified as the order in the hierarchy) involves the solution of an SDP with a semidefinite constraint on a larger matrix than that of the preceding order. Successively, the relaxation is tightened in a manner that eventually converges to the optimal solution. When discussing this approach one should note the immaturity of current SDP solvers, and the exponentially increasing size of the positive semidefinite constraints on the decision variables. As a consequence of the aforementioned, this approach is limited by the dramatically increasing cost of each order in the hierarchy (relative to its predecessor), and therefore, is not scalable. For PPs that satisfy certain conditions, convergence occurs at a finite order in the hierarchy, which is the case for the OPF problems [38]. Nonetheless, since at low orders in the hierarchy the semidefinite constraints are already too large even for moderately sized systems, scalability remains to be an issue [39–42].

DSOS and SDSOS Hierarchies: The Diagonally Dominant Sum-Of-Squares (DSOS) and Scaled Diagonally Dominant Sum-Of-Squares (SDSOS) hierarchies were introduced by the authors of [10]. Each order in the Lasserre Hierarchy requires the solution of an SDP. Alternatively, DSOS and SDSOS are hierarchies of linear and second order programs which can be utilized to solve PPs. In place of the expensive SDP based Lasserre Hierarchy, these hierarchies provide a faster, yet more conservative (i.e., of higher optimality gaps) alternative. Respectively, the cones of diagonally and scaled diagonally dominant matrices are employed in the design of the linear and second order constraints of these hierarchies. Accordingly, these hierarchies scale much more favorably compared to Lasserre's [43]. DSOS and SDSOS hierarchies were applied in [1] to obtain a relaxation of the OPF problem. The work in this thesis relates to tightening the first order relaxation of the DSOS hierarchy.

Shor Relaxation: Due to the quadratic nature of the power flow equations, the OPF problem can be modeled as a Quadratically Constrained Quadratic Program (QCQP) [24]. An SDP relaxation of an QCQP was first proposed in [44]. Accordingly, the OPF problem was relaxed into an SDP in the work of [45]. The SDP relaxation is equivalent to the first order relaxation of the Lasserre Hierarchy [42]. The Lasserre Hierarchy is in turn a generalization of this relaxation [41]. Even though this relaxation is not as computationally demanding as Lasserre's, it still suffers from the immaturity of SDP solvers and scales rather poorly. Accordingly, the SDP relaxation is limited to networks with only a few hundred busses [9]. The work in this thesis utilizes the implementation of [19, 20] in creating a tighter variant of the first order DSOS relaxation of [1]. A similar approach was used to solve for the UC problem [46].

Shor Relaxation Exploiting Chordal Decomposition: Although the SDP relaxation is more tractable (compared to Lasserre's Hierarchy), it still suffers from poor scalability. Power systems are very sparse in practice. Accordingly, the entries in the positive semidefinite constraint are defined according to the graph of the network, thus reflecting whatever sparsity pattern the network has. In such cases, the Matrix Completion Theorem provides a necessary and sufficient semidefinitiness condition for the original problem in terms of smaller matrices [47]. By the Matrix Completion Theorem, the positive semidefinite constraint of the SDP relaxation is satisfied if and only if all submatrices defined based on the network's graph are positive semidefinite. Consequently, the Matrix Completion Theorem can be used to leverage the sparsity pattern in power systems, thus making the Shor relaxation more tractable. This implementation of the SDP relaxation can solve for networks sized at thousands of busses [9]. As such, chordal decomposition (i.e., a decomposition of the positive semidefinite constraint based on the maximal cliques in a chordal graph) was first applied to the OPF problem in [48]. Additional speedup was attained in [20] by providing more control over the problem's overall computational cost. This approach still does not scale as favorably as some of the other alternatives. Nonetheless, a significant speedup was achieved, thus making the almost exact SDP relaxation much more practical. This relaxation is used in investigating the effectiveness of utilizing chordal decomposition in tandem with the relaxation of this dissertation.

Second Order Conic Programming (SOCP) relaxation: First proposed by Jabr [49], another notable relaxation is achieved by formulating the problem as an SOCP. This relaxation was formulated utilizing the BIM model for radial networks. BIM is formulated in terms of a mapping in the product of the voltages. This mapping leads to forfeiting a constraint on the sum of the voltage angles in a loop. As a result, this formulation is exact for radial, but not for mesh networks. The relaxation is then achieved by means of relaxing a non-convex equality constraint on the voltages into a rotated Second Order Cone (SOC) constraint. The positive semidefinitness for all the 2×2 sub-matrices at a transmission line is a necessary, but not a sufficient condition for positive semidefinitness in the Shor relaxation [50]. It follows that the same relaxation can be attained by relaxing the tighter Shor relaxation formulated in complex variables [24]. Accordingly, the SOCP relaxation can be obtained by imposing the aforementioned necessary condition. Similarly, a weaker and less computationally tractable variant can be obtained on the real-valued formulation of the Shor relaxation [9].

The Quadratic Convex (QC) relaxation: The QC relaxation can be considered as a tightened variant of the SOCP relaxation. The work in [51] employed the relatively tight convex McCormick envelopes [52] in relaxing the non-convex quadratic, bilinear and trigonometric terms of the AC-OPF problem. The QC relaxation is formed by utilizing the convex hull of these terms, in addition to the SOC constraints of the SOCP relaxation. It was found to neither dominate nor be dominated by the SDP relaxation, a feat achieved at a significantly lower computational cost [53].

Linear relaxations: Several linear relaxations have been proposed for the OPF problem.

The advantage of linear relaxation can be derived from the maturity of the different linear programming (LP) solver technologies [7]. However, that advantage comes at the cost of higher optimality gaps when compared to those obtained by other formulations. Several linear relaxations have been suggested for the OPF problem including the network flow and copper plate relaxations [54]. Network flow relaxation is a traditional network flow [55] with additional constraints. This formulation constraints the difference in phase angles and line losses while dropping the non-linear thermal constraints from the SOC relaxation of the extended AC-OPF proposed in [54]. On the other hand, the copper plate relaxation is obtained by further relaxing the network flow to completely neglect the network. Both relaxations are valid for networks in which all series line impedances are of non-negative real and imaginary parts. Additionally, a weak linear relaxation is obtained by using McCormick envelopes to represent the power flow equations in rectangular coordinates [56]. The survey [9] provides a detailed treatment of the different linear relaxations of the OPF problem.

Introduced in [10], the DSOS hierarchy aims at providing a more tractable, yet more conservative LP alternative of Lasserre's [37]. However, this approach can be seen to be lacking when implemented to relax the OPF problem [1]. This thesis aims to improve the first order relaxation of the DSOS hierarchy applied to the OPF problem. That is achieved by utilizing the semidefinite relaxation of [19] and [20]. The merit of the sparsity exploiting technique applied to the OPF problem in [20] is also investigated.

2.3 Summary

The topic of Convex relaxations is an active area of research which deals with non-convexities in optimization problems. In this Chapter, we have presented the non-convex classical OPF problem and provided a brief discussion into its non-convexity. We then followed through by a review into the most prominent convex relaxations of the OPF problem. In the next chapter, we provide a more detailed treatment on the background necessary to our treatment of relaxing the OPF problem.

Chapter 3

Preliminaries

In this chapter, we will explore the work in the literature with direct impact on this thesis. The first two sections provide essential mathematical preliminaries. The third outlines the OPF problem formulated using voltages vectors. We then follow by a presentation on the SOS, DSOS and SDSOS optimization paradigms. The final section covers the Shor relaxation of the OPF problem.

3.1 Semidefinite Programming

Semidefinite programming is a branch of mathematical programming concerned with the minimization of an affine function over a spectrahedron (i.e., the space of square positive semidefinite matrices). Let S_n be the set of real symmetric matrices of dimension $n \times n$. $P_n \subseteq S_n$ is the cone [1] of real positive semidefinite matrices in S_n. We define the inner product $\langle ., . \rangle$ of two matrices $A, B \in S_n$ according to the following equation

$$\langle A, B \rangle = Tr(A^T B) = \sum_{i,j} A_{ij} B_{ij} \tag{3.1}$$

[1] A set $\Gamma \subseteq \mathbb{R}^n$ is a cone if $\beta\omega \in \Gamma$ holds for all $\omega \in \Gamma$ and $\beta \in \mathbb{R}^n$ such that $\beta \geq 0$ [7].

The data for the objective function is provided by a matrix $C \in S_n$. Additionally, the constraint data is defined in terms of the sets $\{A_z \in S_n | z = 1, ..., w\}$ and $\{b_z \in \mathbb{R} | z = 1, ..., w\}$. Accordingly, the primal formulation of a semidefinite program is

$$\begin{aligned}
&\min_{\mathcal{X} \in S_n} && \langle C, \mathcal{X} \rangle \\
&\text{subject to} && \\
& && \langle A_z, \mathcal{X} \rangle = b_z, z = 1, ..., w, \\
& && \mathcal{X} \in P_n.
\end{aligned} \tag{3.2}$$

We define a row vector $b = [b_1, ..., b_w]$. Exploiting Lagrangian duality, the dual SDP formulation is

$$\begin{aligned}
&\max_{\mathbf{y} \in \mathbb{R}^w} && b\ \mathbf{y} \\
&\text{subject to} && \\
& && \mathcal{A} = C - \sum_{z=1}^{w} \mathbf{y}_z A_z, \\
& && \mathcal{A} \in P_n.
\end{aligned} \tag{3.3}$$

3.2 Diagonal Dominance and Scaled Diagonal Dominance

In the work of [10], the notions of diagonal dominance and scaled diagonal dominance were utilized in providing cheaper alternatives to SOS optimization, namely the DSOS and SDSOS optimization hierarchies. Subsequently, DSOS and SDSOS relaxations of the OPF problem were introduced in [1]. Similarly, those notions were exploited to linearly inner and outer approximate the positive semidefinite cone in semidefinite programming [57]. The author of [57] further introduced an iterative technique to circumvent the gaps introduced by the Diagonally Dominant (DD) and Scaled Diagonally Dominant (SDD) representations of the positive semidefinite constraint $\mathcal{X} \in P_n$ in an SDP. The aforementioned techniques were later used to solve for the UC problem in [46]. In this section, we present the DD and SDD Cones. We will also go through the extreme ray interpretation of the cone DD and its dual cone [57].

3.2.1 The Diagonally Dominant and Scaled Diagonally Dominant Cones

We begin by defining the cones of DD and SDD matrices. We then describe the relationship between the DD, SDD and positive semidefinite cones; as the latter is integral to the semidefinite relaxation of the OPF problem.

Definition 1 *Consider a symmetric matrix $\mathcal{A} \in S_n$. $\mathcal{A}$ is said to be diagonally dominant if the following holds*

$$\mathcal{A}_{ii} \geq \sum_{j \neq i} \mathcal{A}_{ij} \qquad \forall i \tag{3.4}$$

Alternatively, $\mathcal{A}$ is said to be Scaled Diagonally Dominant (SDD) if there exists a diagonal matrix D such that $D\mathcal{A}D$ is diagonally dominant.

If a matrix $\mathcal{A}$ is diagonally dominant or scaled diagonally dominant, then $\mathcal{A}$ is positive semidefinite [58]. Let us denote the DD and SDD cones as DD_n and SDD_n such that the following inclusion holds

$$DD_n \subseteq SDD_n \subseteq P_n \tag{3.5}$$

The earlier inclusion implies that replacing the positive semidefinite constraint in an SDP by a constraint to either the DD or SDD cone constitutes a convex restriction on the original problem. Accordingly, for the purpose of relaxing an SDP we need to define the dual cones of DD_n and SDD_n. We will present the derivation for the dual cone of DD_n in the following subsections. For that purpose we will utilize the extreme ray definition of the cone DD_n [57]. The same results could be easily extended for the case of the cone SDD_n and its dual.

3.2.2 Extreme Ray Interpretation

The extreme ray interpretation provides an alternate definition of the cone DD_n previously discussed. Accordingly, let $\mathcal{U}_{n,z}$ be the set containing column vectors in $\mathbb{R}^n$ with at most z non-zero elements. A vector $u \in \mathcal{U}_{n,z}$ is defined such that every non-zero element is either $+1$ or -1. Using the vectors in $\mathcal{U}_{n,z}$, define the set of $n \times n$ matrices $U_{n,z}$ as follows

$$U_{n,z} := \{uu^T : u \in \mathcal{U}_{n,z}\} \tag{3.6}$$

For a finite set of matrices $M = \{M_1, M_2, ..., M_m\}$ defines the cone, $cone(M)$ as

$$cone(M) := \{\sum_{i=1}^{m} \alpha_i M_i : \alpha_i \geq 0 \quad \forall i\} \tag{3.7}$$

We now use the following theorem, provided here without proof (see [59] for the proof), which states the equivalence between the cones DD_n and $cone(U_{n,2})$

Theorem 1 *([59])* $DD_n = cone(U_{n,2})$

The equivalence between the diagonally dominant cone DD_n and $cone(U_{n,2})$ demonstrated by the previous theorem implies that the diagonally dominant cone has n^2 extreme rays. In the following section, we utilize this definition in determining the dual cone of DD_n by means of determining the easily attainable dual of $cone(U_{n,2})$.

3.2.3 The Diagonally Dominant Dual Cone

Let the dual cone of $cone(U_{n,2})$ be $cone^*(U_{n,2})$. Similarly, define DD_n^* as the dual cone of DD_n. Since (by Theorem 1), DD_n and $cone(U_{n,2})$ are equivalent, the equivalence between their duals naturally follows. Accordingly, DD_n^* can be obtained as shown below

$$DD_n^* = cone^*(U_{n,2}) = \{\mathcal{A} \in S_n : u_i^T \mathcal{A} u_i \geq 0, \forall u_i \in \mathcal{U}_{n,2}\} \tag{3.8}$$

A symmetric matrix $\mathcal{A} \in S_n$ is said to be positive semidefinite if and only if $u\mathcal{A}u^T \geq 0 : \forall u \in \mathcal{U}_{n,n}$. The implication that $cone^*(U_{n,n})$ is equivalent to P_n follows. Accordingly, and by virtue of the following trivial inclusion

$$cone^*(U_{n,n}) \subseteq cone^*(U_{n,2}) \tag{3.9}$$

the cone $cone^*(U^{n,2})$ is a relaxation of the positive semidefinite cone P_n. Thus, replacing the positive semidefinite constraint on the decision matrix $\mathcal{X}$ in the primal of an SDP by a constraint on the cone DD_n^* we obtain the following optimization problem which is a relaxation of the original SDP of equation (3.2) [57]

$$\begin{aligned}
&\min_{\mathcal{X} \in S_n} && \langle C, \mathcal{X} \rangle \\
&\text{subject to} && \\
& && \langle A_z, \mathcal{X} \rangle = b_z, z = 1, ..., w, \\
& && \mathcal{X} \in DD_n^*.
\end{aligned} \tag{3.10}$$

Finally, the aforementioned relaxation can be obtained by imposing a DD restriction on the positive semidefinite constraint in the dual SDP of (3.5) [57]. The earlier results are easily extendable for the case of the cone SDD_n. In a similar manner, a relaxation of the original problem could be obtained by means of imposing an SDD restriction on the dual of an SDP [57].

3.3 OPF formulated in the voltages vector X

The optimal power flow problem can be reformulated in terms of a vector of the bus voltages in a power system. This formulation will be later utilized in the different relaxations of the OPF problem. What follows is a treatment for the derivation based on the work in [19].

In an n bus power system, we may represent the voltages by the vector of length $2n$

defined as

$$\mathbf{X} = [V_{d1}, V_{d2}, ..., V_{dn}, V_{q1}, V_{q2}, ..., V_{qn}]^T \tag{3.11}$$

It follows that the problem can be rewritten in terms of the vector of decision variables $\mathbf{X}$. The resulting problem would be a program composed of polynomials of a maximum order of either 2 or 4. Based on the derivation in [42], we can rewrite the optimal power problem as a program of degree 2. That is achieved by defining the problem in terms of the decision variables of $\mathbf{X}$, as well as the active and reactive line flows in the system. We begin by defining a standard basis vector e_i in $\mathbb{R}^n$. Equation (2.5) can be written in terms of $\mathbf{X}$ in the following manner

$$V_i^2 = V_{di}^2 + V_{qi}^2 = \begin{bmatrix} e_i e_i^T & 0 \\ 0 & e_i e_i^T \end{bmatrix} \mathbf{X}\mathbf{X}^T \tag{3.12}$$

Accordingly, we define a matrix M_i for each bus i in the system as

$$M_i = \begin{bmatrix} e_i e_i^T & 0 \\ 0 & e_i e_i^T \end{bmatrix} \tag{3.13}$$

Let y_{lm} and b_{lm} be the series admittance and shunt susceptance of the line connecting nodes l and m respectively. Accordingly, we can define the matrices Y_i and Y_{lm} for each bus i and line lm as follows

$$Y_i = e_i e_i^T \mathbf{Y} \tag{3.14}$$

$$Y_{lm} = (y_{lm} + j\frac{b_{lm}}{2})e_l e_l^T - (y_{lm})e_l e_m^T \tag{3.15}$$

From the above, the complex power injection at bus i and apparent power flow in the line between nodes l and m can be found using $\mathbf{V}^* Y_i \mathbf{V}$, $\mathbf{V} Y_{lm} \mathbf{V}^*$ respectively

$$S_{i,inj}^* = V_i^* I_i = \mathbf{V}^* e_i e_i^* \mathbf{I} = \mathbf{V}^* Y_i \mathbf{V} \tag{3.16}$$

$$S_{lm}^{*} = V_l^{*}(V_l^{*}(y_{lm} + j\frac{b_{lm}}{2})) + V_l^{*}(V_l - V_m)y_{lm} = \mathbf{V}Y_{lm}\mathbf{V}^{*} \tag{3.17}$$

where $*$ is the conjugate transpose operator and $\mathbf{V}$ is a vector of the complex bus voltages.

Let $Re(.)$ and $Im(.)$ respectively denote the real and imaginary parts of a matrix in $\mathbb{C}^{n\times n}$. Accordingly, the active and reactive power injections at bus i can be formulated in terms of $\mathbf{X}$ in the following manner

$$\begin{aligned} P_{i,inj} = \quad Re\{S_i^{*}\} = \quad Re\{\mathbf{V}^{*}Y_i\mathbf{V}\} = & \quad \mathbf{X}^T \begin{bmatrix} Re(Y_i) & -Im(Y_i) \\ Im(Y_i) & Re(Y_i) \end{bmatrix} \mathbf{X} \\ = & \quad \frac{1}{2}\mathbf{X}^T \begin{bmatrix} Re(Y_i + Y_i^T) & -Im(Y_i - Y_i^T) \\ Im(Y_i - Y_i^T) & Re(Y_i + Y_i^T) \end{bmatrix} \mathbf{X} \end{aligned} \tag{3.18}$$

$$\begin{aligned} Q_{i,inj} = -Im\{S_i^{*}\} = -Im\{\mathbf{V}^{*}Y_i\mathbf{V}\} = & \quad -\mathbf{X}^T \begin{bmatrix} Im(Y_i) & -Re(Y_i) \\ Re(Y_i) & Im(Y_i) \end{bmatrix} \mathbf{X} \\ = & -\frac{1}{2}\mathbf{X}^T \begin{bmatrix} Im(Y_i + Y_i^T) & -Re(Y_i - Y_i^T) \\ Re(Y_i - Y_i^T) & Im(Y_i + Y_i^T) \end{bmatrix} \mathbf{X} \end{aligned} \tag{3.19}$$

Similarly, solving for the active and reactive power flows of line lm in terms of the voltage vector $\mathbf{X}$ yields the equations below

$$\begin{aligned} P_{lm} = Re\{\mathbf{V}Y_{lm}\mathbf{V}^{*}\} = & \quad \mathbf{X}^T \begin{bmatrix} Re(Y_{lm}) & -Im(Y_{lm}) \\ Im(Y_{lm}) & Re(Y_{lm}) \end{bmatrix} \mathbf{X} \\ = & \frac{1}{2}\mathbf{X}^T \begin{bmatrix} Re(Y_{lm} + Y_{lm}^T) & -Im(Y_{lm} - Y_{lm}^T) \\ Im(Y_{lm} - Y_{lm}^T) & Re(Y_{lm} + Y_{lm}^T) \end{bmatrix} \mathbf{X} \end{aligned} \tag{3.20}$$

$$Q_{lm} = -Im\{\mathbf{V}Y_{lm}\mathbf{V}^*\} = -\mathbf{X}^T \begin{bmatrix} Im(Y_{lm}) & -Re(Y_{lm}) \\ Re(Y_{lm}) & Im(Y_{lm}) \end{bmatrix} \mathbf{X} \tag{3.21}$$

$$= -\frac{1}{2}\mathbf{X}^T \begin{bmatrix} Im(Y_{lm} + Y_{lm}^T) & -Re(Y_{lm} - Y_{lm}^T) \\ Re(Y_{lm} - Y_{lm}^T) & Im(Y_{lm} + Y_{lm}^T) \end{bmatrix} \mathbf{X}$$

We may then define matrices $\mathbf{Y}_i, \bar{\mathbf{Y}}_i, \mathbf{Y}_{lm}$ and $\bar{\mathbf{Y}}_{lm}$ as

$$\mathbf{Y}_i = \frac{1}{2} \begin{bmatrix} Re(Y_i + Y_i^T) & -Im(Y_i - Y_i^T) \\ Im(Y_i - Y_i^T) & Re(Y_i + Y_i^T) \end{bmatrix} \tag{3.22}$$

$$\bar{\mathbf{Y}}_i = -\frac{1}{2} \begin{bmatrix} Im(Y_i + Y_i^T) & -Re(Y_i - Y_i^T) \\ Re(Y_i - Y_i^T) & Im(Y_i + Y_i^T) \end{bmatrix} \tag{3.23}$$

$$\mathbf{Y}_{lm} = \frac{1}{2} \begin{bmatrix} Re(Y_{lm} + Y_{lm}^T) & -Im(Y_{lm} - Y_{lm}^T) \\ Im(Y_{lm} - Y_{lm}^T) & Re(Y_{lm} + Y_{lm}^T) \end{bmatrix} \tag{3.24}$$

$$\bar{\mathbf{Y}}_{lm} = -\frac{1}{2} \begin{bmatrix} Im(Y_{lm} + Y_{lm}^T) & -Re(Y_{lm} - Y_{lm}^T) \\ Re(Y_{lm} - Y_{lm}^T) & Im(Y_{lm} + Y_{lm}^T) \end{bmatrix} \tag{3.25}$$

such that

$$P_{i,inj} = \mathbf{X}^T\mathbf{Y}_i\mathbf{X} = Tr(\mathbf{Y}_i\mathbf{X}^T\mathbf{X}) \tag{3.26}$$

$$Q_{i,inj} = \mathbf{X}^T\bar{\mathbf{Y}}_i\mathbf{X} = Tr(\bar{\mathbf{Y}}_i\mathbf{X}^T\mathbf{X}) \tag{3.27}$$

$$P_{lm} = \mathbf{X}^T\mathbf{Y}_{lm}\mathbf{X} = Tr(\mathbf{Y}_{lm}\mathbf{X}^T\mathbf{X}) \tag{3.28}$$

$$Q_{lm} = \mathbf{X}^T\bar{\mathbf{Y}}_{lm}\mathbf{X} = Tr(\bar{\mathbf{Y}}_{lm}\mathbf{X}^T\mathbf{X}) \tag{3.29}$$

Substituting for equations (3.27)–(3.30), we obtain the following optimization program:

$$\min \sum_{g \in \mathcal{N}_g} f_g^q(P_{Gg}) \tag{3.30a}$$

subject to

$$P_{Gi}^{min} \le Tr(\mathbf{Y}_i \mathbf{X}^T \mathbf{X}) + P_{Di} \le P_{Gi}^{max} \qquad \forall i \in \mathcal{N} \tag{3.30b}$$

$$Q_{Gi}^{min} \le Tr(\bar{\mathbf{Y}}_i \mathbf{X}^T \mathbf{X}) + Q_{Di} \le Q_{Gi}^{max} \qquad \forall i \in \mathcal{N} \tag{3.30c}$$

$$(V_i^{min})^2 \le Tr(M_i \mathbf{X}^T \mathbf{X}) \le (V_i^{max})^2 \qquad \forall i \in \mathcal{N} \tag{3.30d}$$

$$P_{lm}^2 + Q_{lm}^2 \le (S_k^{max})^2 \qquad \forall (l,m) \in L \tag{3.30e}$$

$$P_{lm} = Tr(\mathbf{Y}_{lm} \mathbf{X}^T \mathbf{X}) \qquad \forall (l,m) \in L \tag{3.30f}$$

$$Q_{lm} = Tr(\bar{\mathbf{Y}}_{lm} \mathbf{X}^T \mathbf{X}) \qquad \forall (l,m) \in L \tag{3.30g}$$

$$P_{Gg} = Tr(\mathbf{Y}_g \mathbf{X}^T \mathbf{X}) \qquad \forall g \in \mathcal{N}_g \tag{3.30h}$$

The resulting program is composed of polynomials of an order of a maximum degree of 2 in the decision variables $\mathbf{X}, P_{lm}$ and Q_{lm}. This program is to be utilized in the SOS, DSOS, SDSOS and SDP relaxations of the following sections.

3.4 Sum-Of-Squares(SOS) relaxation

3.4.1 Sum-Of-Squares Approach For Polynomial Programming

A lot of convex relaxations of the OPF problem fail at obtaining a globally optimal solution either in the value of the objective or in the decision variables. Due in part to its application in evaluating the optimality of new approaches, as well as providing an optimal solution by its own right, research effort towards the global solution of the OPF problem has been getting a lot of interest [9].

A polynomial program (PP) is that such the constraints and objective are multivariate polynomials in the decision variables of the problem. Where $f(x)$ and $g_i(x)$ are polynomials in x, the following is an example of a PP

$$\begin{aligned} &\min && f(x) \\ &\text{subject to} && \\ & && g_z(x) \geq 0, \quad z = \{1, ..., w\}. \end{aligned} \tag{3.31}$$

Solving the PP shown above boils down to certifying the non-negativity of the polynomials $g_z(x)$. Accordingly, this problem is an optimization problem over non-negative polynomials, which is unfortunately NP-hard [60]. Formulating the OPF problem as a PP affords us the advantage of utilizing PP techniques to tackle its inherent non-convexity. One example of these techniques is the Lasserre Hierarchy [37]. The Lasserre Hierarchy is a hierarchy of relaxations that converges to the optimal solution of a polynomial program. Since the OPF can be formulated as such [42], Lasserre Hierarchy can then be utilized in globally solving the OPF problem [39–42].

By definition, an SOS polynomial is non-negative. However, not all non-negative polynomials are SOS [61]. Accordingly, the set of SOS polynomials is a restriction on the set of non-negative polynomials. It follows that imposing that restriction on the dual yields a relaxation of the original problem. Accordingly, we consider the dual of (3.31) demonstrated in the following equation

$$\begin{aligned} &\max && \varphi \\ &\text{subject to} && \\ & && f(x) - \varphi \geq 0, \\ & && \forall x : g_z(x) \geq 0, z = \{1, ..., w\}. \end{aligned} \tag{3.32}$$

The aforementioned program can be equivalently formulated as the conic program [62]

$$\begin{aligned} \max \quad & \varphi \\ \text{subject to} \quad & \\ & f(x) - \varphi \in \mathcal{P}_d(S_{\mathcal{G}}) \end{aligned} \tag{3.33}$$

where $S_{\mathcal{G}}$ constitutes the set of decision variables for which the set of polynomials $\mathcal{G}$ is non-negative. Additionally, the cone $\mathcal{P}_d(S_{\mathcal{G}})$ is the cone of non-negative polynomials of degree at most d over the set $S_{\mathcal{G}}$. Despite the fact $\mathcal{P}_d(S_{\mathcal{G}})$ is convex, optimizing over it is NP-hard [10]. Accordingly, the process of solving for (3.33) translates to obtaining a tractable replacement for the set $\mathcal{P}_d(S_{\mathcal{G}})$.

By imposing an SOS restriction on the dual (3.33), the problem of optimizing over non-negative polynomials is relaxed into that of optimizing over the convex SOS polynomials. Lasserre in his work proposes a hierarchy of successive relaxations [37]. Accordingly, a relaxation of this problem can be obtained by constraining the polynomial to be SOS. That is achieved by imposing the following restriction

$$f(x) - \varphi = \sigma_0 + \sum_{i=1}^{m} g_z(x)\sigma_z \tag{3.34}$$

with $r \geq d$ being the order in the hierarchy. σ_0 is an SOS polynomial of degree r, while the degree of the SOS polynomial σ_i is $(r - degree(g_z(x))$ for $z = \{1, ..., w\}$. Imposing that a polynomial is SOS translates to an SDP, which follows from the following theorem

Theorem 2 ([63, 64]) *Let z(x,d) be a vector of monomials, in the elements in x, of degree at most d. A multivariate polynomial p := p(x) in n variables and of degree 2d is a sum of squares if and only if there exists a positive semidefinite symmetric matrix* $\boldsymbol{Q}$ *such that* $p(x){=}z(x,d)^T \boldsymbol{Q} z(x,d)$.

The matrix $\mathbf{Q}$ (i.e., the Gram matrix) is of the dimension $\binom{n+d}{d} \times \binom{n+d}{d}$, which approximately translates to $n^d \times n^d$. If the order (r) is fixed then the following problem can be cast as an SDP

$$
\begin{aligned}
&\max && \varphi \\
&\text{subject to} && \\
& && f(x) - \varphi = \sigma_0 + \sum_{i=1}^{m} g_i(x)\sigma_i && (3.35)\\
& && \sigma_0 \in SOS_r && (3.36)\\
& && \sigma_i \in SOS_{r-deg(g_i)}. && (3.37)
\end{aligned}
$$

From the above, we can see that each order in the hierarchy translates to solving an SDP. Accordingly, in addition to the order in the hierarchy, the SDP cost is dependant on the order of the polynomials being optimized over. Let $\Lambda(\mathbf{X}), \Lambda(P_{Gg})$ and $\Lambda(P_{lm}, Q_{lm})$ be SOS polynomials of degree at most 2. As such, the relaxation PP-SOS$_2$ could be defined as [42]

$$\max \quad \varphi \tag{3.38a}$$

subject to

$$
\begin{aligned}
\sum_{g\in\mathcal{N}_g} f_g^q(P_{Gg}) - \varphi = {} & \Lambda(\mathbf{X}) + \sum_{g\in\mathcal{N}_g} \Lambda(P_{Gg}) + \sum_{k\in\mathcal{L}} \Lambda(P_{lm}, Q_{lm}) \\
& + \sum_{i\in\mathcal{N}} \bar{\lambda}_i(P_{Gi}^{max} - P_{Di} - Tr(Y_i x x^T) + \sum_{i\in\mathcal{N}} \underline{\lambda_i}(-P_{Gi}^{min} + P_{Di} + Tr(Y_i x x^T) \\
& + \sum_{i\in\mathcal{N}} \bar{\gamma}_i(Q_{Gi}^{max} - Q_{Di} - Tr(\bar{Y}_i x x^T) + \sum_{i\in\mathcal{N}} \underline{\gamma_i}(-Q_{Gi}^{min} + Q_{Di} + Tr(\bar{Y}_i x x^T) \\
& + \sum_{i\in\mathcal{N}} \bar{\mu}_i((V_i^{max}) - Tr(M_i x x^T)) + \sum_{i\in\mathcal{N}} \underline{\mu_i}((-V_i^{min}) + Tr(M_i x x^T)) \\
& + \sum_{(l,m)\in L} \hat{a}_{lm}((S_{lm}^{max})^2 - P_{lm}^2 - Q_k^2) + \sum_{i\in\mathcal{N}_g} \hat{b}_i(P_{Gi} - P_{Di} - Tr(Y_i x x^T)) \\
& + \sum_{(l,m)\in L} \hat{c}_{lm}(P_{lm} - Tr(Y_{lm} x x^T) + \sum_{(l,m)\in L} \hat{d}_{lm}(Q_{lm} - Tr(\bar{Y}_{lm} x x^T)
\end{aligned}
\tag{3.38b}
$$

Following the previous discussion we can see that, for all $g \in \mathcal{N}_g$ and $(l,m) \in L$, the SOS polynomials $\Lambda(\mathbf{X})$, $\Lambda(P_{Gg})$ and $\Lambda(P_{lm}, Q_{lm})$ accept the following representations

$$\Lambda(\mathbf{x}) = \mathbf{X}\mathbf{Q}_{\mathbf{X}}\mathbf{X}^T, \qquad \mathbf{Q}_{\mathbf{X}} \in P_n \tag{3.39}$$

$$\Lambda(P_{Gg}) = \begin{bmatrix} 1 \\ P_{Gg} \end{bmatrix} \mathbf{Q}_{P_{Gg}} \begin{bmatrix} 1 & P_{Gg} \end{bmatrix}, \qquad \mathbf{Q}_{P_{Gg}} \in P_2 \qquad \forall g \in \mathcal{N}_g \tag{3.40}$$

$$\Lambda(P_{lm}, Q_{lm}) = \begin{bmatrix} 1 \\ P_{lm} \\ Q_{lm} \end{bmatrix} \mathbf{Q}_{(P_{lm},Q_{lm})} \begin{bmatrix} 1 & P_{lm} & Q_{lm} \end{bmatrix}, \quad \mathbf{Q}_{(P_{lm},Q_{lm})} \in P_3 \qquad \forall (l,m) \in L \tag{3.41}$$

The SOS relaxation is imposed via restricting the PP to the SOS cone. That restriction is imposed via the positive semidefinite constraints shown above. Similarly, a DSOS and SDSOS program could obtained via restricting the PP to either the cone of DSOS or SDSOS polynomials. These programs translate to an LP and an SOCP respectively, and are thus much more tractable. However, considering how both the DSOS and SDSOS cones are contained in the SOS cone, such a restriction produces a relaxation of the program shown above, a fact that we will utilize in our proposed relaxation [10]. More details on DSOS and SDSOS optimization follow in the next section.

3.5 DSOS and SDSOS optimization

Two useful cones inside the SOS cone are those of diagonally-dominant-sum-of-squares (DSOS) and scaled-diagonally-dominant-sum-of-squares (SDSOS) polynomials. While the SOS cone lends itself to representations in the positive semidefinite cone, the cones DSOS and SDSOS accept representations as LPs and SOCPs. By virtue of these representations, DSOS and SDSOS present more tractable alternatives to SOS in solving for PPs. We first begin by stating the definition for the DSOS and SDSOS cones.

Definition 2 ([10]) *A polynomial $p := p(x)$ is diagonally-dominant-sum-of-squares (DSOS)* *if it can be written as*

$$p(x) = \sum_i \alpha_i m_i^2(x) + \sum_{i,j} \beta_{ij}^+ (m_i(x) + m_j(x))^2 + \sum_{i,j} \beta_{ij}^- (m_i(x) - m_j(x))^2 \qquad (3.42)$$

for some monomials $m_i(x), m_j(x)$ *and some non-negative scalars* $\alpha_i, \beta_{ij}^+, \beta_{ij}^-$.

Definition 3 ([10]) *A polynomial* $p := p(x)$ *is scaled-diagonally-dominant-sum-of-squares (SDSOS) if it can be written as*

$$p(x) = \sum_i \alpha_i m_i^2(x) + \sum_{i,j} (\hat{\beta}_{ij}^+ m_i(x) + \tilde{\beta}_{ij}^+ m_j(x))^2 + \sum_{i,j} (\hat{\beta}_{ij}^- m_i(x) - \tilde{\beta}_{ij}^- m_j(x))^2 \qquad (3.43)$$

for some monomials $m_i(x), m_j(x)$ *and some scalars* $\alpha_i, \hat{\beta}_{ij}^+, \tilde{\beta}_{ij}^+, \hat{\beta}_{ij}^-, \tilde{\beta}_{ij}^-$ *with* $\alpha_i \in \mathbb{R}^+$.

Similarly to the definition of an SOS polynomial, a polynomial $p(x) = z(x,d)^T \mathbf{Q} z(x,d)$ is DSOS if and only if the Gram matrix $\mathbf{Q}$ is DD. Alternatively, $p(x)$ is SDSOS if and only if $\mathbf{Q}$ is SDD [10]. Accordingly, a restriction on a polynomial to either the DSOS or SDSOS cone translates to a restriction on the Gram matrix to be either DD or SDD respectively. Since these restrictions respectively translate to an LP and an SOCP, one can clearly see how DSOS and SDSOS programs are nothing but linear and second order conic programming problems.

We denote the sets of n variable SOS, DSOS and SDSOS polynomials with a degree of utmost $2d$ respectively as $SOS_{n,2d}$, $DSOS_{n,2d}$ and $SDSOS_{n,2d}$. Let $PSD_{n,2d}$ be the cone of n variable non-negative polynomials of degree at most 2d. Accordingly, the following inclusion holds

$$DSOS_{n,2d} \subseteq SDSOS_{n,2d} \subseteq SOS_{n,2d} \subseteq PSD_{n,2d} \qquad (3.44)$$

DSOS and SDSOS programs were applied to relax the OPF problem in [1]. From the inclusion above we can make a number of observations. First, a DSOS restriction on the dual of the OPF program constitutes a linear relaxation on the entire feasible set of the OPF problem. The utilization of an LP representation in solving for a general PP was shown to be lacking when compared to other approaches [1, 65, 66]. This coincides with the results obtained via the DSOS relaxation of the OPF problem. Accordingly, we could attribute these poor bounds to the utilization of a polyhedral set for the relaxation of the entirety of the feasible space. Second, for an n variable polynomial and at a fixed degree $2d$, restrictions on the cones $SOS_{n,2d}$ and $SDSOS_{n,2d}$ yield tighter yet less tractable relaxations of the problem [1, 10, 43]. Consider the case of an n variable quadratic PP. For such PPs, Table 3.1 illustrates constraint type, number and associated variable count for the first order SOS, DSOS and SDSOS hierarchies.

Table 3.1 Comparison of number of conic constraints in first order of the SOS, SDSOS, and DSOS hierarchies for quadratic n variable PPs. This table is taken from [1].

	Type	Number of Conic Constraints	Number of Variables Per Constraint
$SOS_{n,2}$	SDP	1	$(n+1) \times (n+1)$
$SDSOS_{n,2}$	SOCP	$(n+1)n/2$	3
$DSOS_{n,2}$	LP	$n+1$	n
	LP	n^2+n	2

Finally, and from the last two observations, we can conclude that a polyhedral restriction on only a part of the feasible set (represented by the cone $SOS_{n,2d}$) would yield a tighter relaxation. As such, we can assume that the correct utilization of a combination of polyhedral, second order conic and semidefinite representations could yield a tighter and a more scalable alternative to the $DSOS_2$ relaxation of the OPF problem. That is the case especially if we were to relax the more tractable SDP relaxation of the OPF problem instead of PP-SOS_2 [1]. In the following chapter, we utilize these assumptions in the relaxation we propose.

3.6 Shor Relaxation

An SDP relaxation of a quadratically constrained quadratic program (QCQP) was first proposed in [44]. Since the OPF accepts the formulation of a QCQP [24], an SDP relaxation of the problem can be obtained [19, 45]. Non-convexity in the OPF problem is in part an artifact of the non-linearity in the problem formulation [19]. By inspection, we can see that the matrix $\mathbf{X}^T\mathbf{X}$ contains all the non-linear terms of the problem. Accordingly, a convex relaxation could be obtained by properly handling $\mathbf{X}^T\mathbf{X}$. Introducing a mapping $\mathbf{W} = \mathbf{X}^T\mathbf{X}$ and the condition $rank(\mathbf{W}) = 2$, the power flow equations can be rewritten in terms of $\mathbf{W}$. Having done that, all constraints in terms of the quadratic and bilinear elements of the voltages are now linear with respect to $\mathbf{W}$. This isolates the non-convexity of the problem in the rank constraint introduced above. Accordingly, the SDP relaxation is achieved by replacing the rank constraint by a positive semidefinite constraint on $\mathbf{W}$ [19, 45]. If the SDP solution is of rank two (alternatively, rank one in the complex formulation of the problem) then the relaxation is exact for that particular instance. Otherwise, this relaxation provides a lower bound on the objective function. The primal formulation of the SDP relaxation of the OPF problem is displayed below

$$\min \sum_{g\in\mathcal{N}_g} f_g^q(P_{Gg}) \tag{3.45a}$$

subject to

$$P_{Gi}^{min} \leq Tr(\mathbf{Y}_i\mathbf{W}) + P_{Di} \leq P_{Gi}^{max} \qquad \forall i \in \mathcal{N} \tag{3.45b}$$

$$Q_{Gi}^{min} \leq Tr(\bar{\mathbf{Y}}_i\mathbf{W}) + Q_{Di} \leq Q_{Gi}^{max} \qquad \forall i \in \mathcal{N} \tag{3.45c}$$

$$(V_i^{min})^2 \leq Tr(M_i\mathbf{W}) \leq (V_i^{max})^2 \qquad \forall i \in \mathcal{N} \tag{3.45d}$$

$$P_{lm}^2 + Q_{lm}^2 \leq (S_k^{max})^2 \qquad \forall (l,m) \in L \tag{3.45e}$$

$$P_{lm} = Tr(\mathbf{Y}_{lm}\mathbf{W}) \qquad \forall (l,m) \in L \tag{3.45f}$$

$$Q_{lm} = Tr(\bar{\mathbf{Y}}_{lm}\mathbf{W}) \qquad \forall (l,m) \in L \tag{3.45g}$$

$$P_{Gg} = Tr(\mathbf{Y}_g\mathbf{W}) \qquad \forall g \in \mathcal{N}_g \tag{3.45h}$$

$$\mathbf{W} \in P_{2n} \tag{3.45i}$$

3.7 Summary

The OPF problem is a non-convex optimization problem. The non-convexity of the OPF problem motivates the utilization of convex relaxations in relaxing the problem's feasible set. In this chapter we presented the SDP, SOS, SDSOS and DSOS relaxations of this problem. As such, this chapter served to present the past treatments of the OPF problem relevant to our work, the relationship between them and the mathematical preliminaries necessary for their implementation. In Chapter 4, we discuss a different way of relaxation the OPF problem which builds on top of the relaxations and mathematical preliminaries presented in this chapter.

Chapter 4

Diagonally Dominant Relaxation of the OPF problem

The first order DSOS relaxation ($DSOS_2$) was implemented to relax the OPF problem in [1]. However, that implementation suffers from two limitations: poor optimality gaps and intractability. This section serves to demonstrate how a more conservative utilization of diagonal dominance could lead to tighter and more scalable relaxations. We first begin by outlining the large scale SDP relaxation of the OPF problem presented in [20], and then follow through with our proposed relaxation.

4.1 Large Scale OPF

The OPF formulation utilized in this chapter is that of [20]. This formulation was introduced to provide a representation of the problem more suitable for large networks. Instead of restricting the generation at a node to one generator, this formulation allows for the existence of multiple generators at a node, each defined by its individual generation and respective cost function. It further allows for the inclusion of generators with piece-wise linear cost functions. Parallel transmission lines are also integrated into the definition of the OPF problem. Following is a presentation of this formulation and its dual.

4.1.1 Formulation

In a network, consider an arbitrary bus i and line k (connecting busses l_k and m_k). The set $\mathcal{L}$ is composed of all such lines in the network, thus permitting multiple lines between busses l and m. As such, this definition allows for parallel connections between two busses in the power system. $\mathbb{G}_i^{pw}$ and $\mathbb{G}_i^{q}$ are, respectively, the sets of generators with piece-wise linear cost and generators with quadratic cost at bus i.

Let $\mathcal{N}$ be the set of all busses in the system. It follows that the sets $\mathcal{N}$,$\mathbb{G}^{pw}$,$\mathbb{G}^{q}$ and $\mathbb{G}$ are defined as

$$\mathcal{N} := \{1, 2, 3, ..., n\} \tag{4.1}$$

$$\mathbb{G}_i := \mathbb{G}_i^{pw} \bigcup \mathbb{G}_i^{q} \quad \forall i \in \mathcal{N} \tag{4.2}$$

$$\mathbb{G}^{pw} := \bigcup_{i \in \mathcal{N}} \mathbb{G}_i^{pw} \tag{4.3}$$

$$\mathbb{G}^{q} := \bigcup_{i \in \mathcal{N}} \mathbb{G}_i^{q} \tag{4.4}$$

The phasors S_{Gg}, S_{di} and V_i follow the definitions of equations (2.6)–(2.8). However, contrary to the earlier formulation of the problem, the parameter S_{Gg} is defined for each generator $g \in \mathbb{G}_i$, instead of $g \in \mathcal{N}_g$. For each generator g such that $g \in \mathbb{G}^q$, the cost of generation is calculated using the cost function of (2.8), defined specifically for each such generator g. Similarly, for each generator $g \in \mathbb{G}^{pw}$, we define the following cost function

$$f_g^{pw}(P_{Gg}) = \begin{cases} m_{g1}(P_{Gg} - a_{g1}) + b_{g1} & P_{Gg} \leq a_{g1} \\ m_{g2}(P_{Gg} - a_{g2}) + b_{g2} & a_{g1} \leq P_{Gg} \leq a_{g2} \\ . & . \\ . & . \\ m_{gr}(P_{Gg} - a_{gr}) + b_{gr} & a_{gr} \leq P_{Gg} \end{cases} \tag{4.5}$$

where r_g is the number of line segments. Each segment $t \in \{1, ..., r_g\}$ is specified by a slope m_{gt} and a point (a_{gt}, b_{gt}). Following the earlier definitions, the large-scale OPF problem of [20] is cast as

$$\min \sum_{g \in \mathbb{G}_{pw}} f_g^{pw}(P_{Gg}) + \sum_{g \in \mathbb{G}_q} f_g^q(P_{Gg}) \tag{4.6a}$$

subject to

$$P_{Gg}^{min} \leq P_{Gg} \leq P_{Gg}^{max} \qquad \forall g \in \mathbb{G} \tag{4.6b}$$

$$Q_{Gg}^{min} \leq Q_{Gg} \leq Q_{Gg}^{max} \qquad \forall g \in \mathbb{G} \tag{4.6c}$$

$$|V_i^{min}| \leq |V_i| \leq |V_i^{max}| \qquad \forall i \in \mathcal{N} \tag{4.6d}$$

$$-S_k^{max} \leq S_k \leq +S_k^{max} \qquad \forall k \in \mathcal{L} \tag{4.6e}$$

$$\sum_{g \in \mathbb{G}_i} P_{Gg} - P_{Di} = V_{di} \sum_{h=1}^{n} (\mathbf{G}_{ih} V_{dh} - \mathbf{B}_{ih} V_{qh}) + V_{qi} \sum_{h=1}^{n} (\mathbf{B}_{ih} V_{dh} + \mathbf{G}_{ih} V_{qh}) \qquad \forall i \in \mathcal{N} \tag{4.6f}$$

$$\sum_{g \in \mathcal{G}_i} Q_{Gg} - Q_{Di} = V_{di} \sum_{h=1}^{n} (-\mathbf{G}_{ih} V_{dh} - \mathbf{B}_{ih} V_{qh}) + V_{qi} \sum_{h=1}^{n} (\mathbf{B}_{ih} V_{dh} - \mathbf{G}_{ih} V_{qh}) \quad \forall i \in \mathcal{N} \tag{4.6g}$$

4.1.2 SDP relaxation Primal and Dual Forms

We first begin by outlining the matrices used in formulating the SDP relaxation of the optimization problem in (4.6). The SDP relaxation of this program employs the matrices $\mathbf{Y}_i, \bar{\mathbf{Y}}_i$ and M_i, which were introduced in Section 3.3. A single connection at an edge of a network in this formulation encompasses both the transmission line and transformer of that connection. Following the work in [67], a connection k is therefore modeled by a common branch model of a π transmission line in series with an ideal transformer of tap ratio magnitude τ_k and phase shift angle θ_k.

Let f_i be the ith standard basis vector in $\mathbb{R}^{2N}$, and define the parameters c_l, c_m, s_l and s_m as

$$c_l = \left(\frac{1}{2\tau_k}\right) \left(g_k \cos(\theta_k) + b_k \cos\left(\theta_k + \frac{\pi}{2}\right)\right) \tag{4.7}$$

$$c_m = \left(\frac{1}{2\tau_k}\right) \left(g_k \cos(-\theta_k) + b_k \cos\left(-\theta_k + \frac{\pi}{2}\right)\right) \tag{4.8}$$

$$s_l = \left(\frac{1}{2\tau_k}\right)\left(g_k \sin(\theta_k) + b_k \sin\left(\theta_k + \frac{\pi}{2}\right)\right) \tag{4.9}$$

$$s_m = \left(\frac{1}{2\tau_k}\right)\left(g_k \sin(-\theta_k) + b_k \sin\left(-\theta_k + \frac{\pi}{2}\right)\right) \tag{4.10}$$

Following the earlier definitions, the line flow constraints matrices can be cast as

$$\begin{aligned}\mathbf{Z}_{k_l} =& \frac{g_k}{\tau_k^2}(f_{l_k}f_{l_k}^T + f_{l_k+n}f_{l_k+n}^T) \\ &- c_l(f_{l_k}f_{m_k}^T + f_{m_k}f_{l_k}^T + f_{l_k+n}f_{m_k+n}^T + f_{m_k+n}f_{l_k+n}^T) \\ &+ s_l(f_{l_k}f_{m_k+n}^T + f_{m_k+n}f_{l_k}^T - f_{l_k}f_{m_k+n}^T - f_{m_k+n}f_{l_k}^T)\end{aligned} \tag{4.11}$$

$$\begin{aligned}\mathbf{Z}_{k_m} =& g_k(f_{m_k}f_{m_k}^T + f_{m_k+n}f_{m_k+n}^T) \\ &- c_m(f_{l_k}f_{m_k}^T + f_{m_k}f_{l_k}^T + f_{l_k+n}f_{m_k+n}^T + f_{m_k+n}f_{l_k+n}^T) \\ &+ s_m(f_{l_k+n}f_{m_k}^T + f_{m_k}f_{l_k+n}^T - f_{l_k}f_{m_k+n}^T - f_{m_k+n}f_{l_k}^T)\end{aligned} \tag{4.12}$$

$$\begin{aligned}\bar{\mathbf{Z}}_{k_l} =& -\left(\frac{2b_k + b_{sh,k}}{2\tau_k^2}\right)(f_{l_k}f_{l_k}^T + f_{l_k+n}f_{l_k+n}^T) \\ &+ c_l(f_{l_k}f_{m_k+n}^T + f_{m_k+n}f_{l_k}^T - f_{l_k}f_{m_k+n}^T - f_{m_k+n}f_{l_k}^T) \\ &+ s_l(f_{l_k}f_{m_k}^T + f_{m_k}f_{l_k}^T + f_{l_k+n}f_{m_k+n}^T + f_{m_k+n}f_{l_k+n}^T)\end{aligned} \tag{4.13}$$

$$\begin{aligned}\bar{\mathbf{Z}}_{k_m} =& -\left(b_k + \frac{b_{sh,k}}{2}\right)(f_{m_k}f_{m_k}^T + f_{m_k+n}f_{m_k+n}^T) \\ &+ c_m(f_{l_k+n}f_{m_k}^T + f_{m_k}f_{l_k+n}^T - f_{l_k}f_{m_k+n}^T - f_{m_k+n}f_{l_k}^T) \\ &+ s_m(f_{l_k}f_{m_k}^T + f_{m_k}f_{l_k}^T + f_{l_k+n}f_{m_k+n}^T + f_{m_k+n}f_{l_k+n}^T)\end{aligned} \tag{4.14}$$

We draw on the formulation outlined before for the SDP relaxation of the OPF problem. In the same manner displayed in Section 3.6, the active and reactive power injections as well as the voltages at bus i are defined in terms of the matrices $\mathbf{Y}_i, \bar{\mathbf{Y}}_i$ and M_i. However, the transformer inclusion of the common branch model produces an asymmetry in power flows in a line $k \in \mathcal{L}$, connecting an edge (l_k, m_k). Accordingly, the active and reactive power flows for the line are defined separately for each terminal u as $Tr(\mathbf{Z}_{k_u}\mathbf{W})$ and $Tr(\bar{\mathbf{Z}}_{k_u}\mathbf{W})$ respectively. The quadratic cost functions and apparent line flow limits are formulated using

Schur's complement [7]. On the other hand, the constrained cost variable method of [67] is utilized in formulating the piece-wise linear cost functions. The SDP relaxation obtained by relaxing the rank constraint on $\mathbf{W}$ is

$$\min \sum_{g \in \mathbb{G}_{pw}} \beta_g + \sum_{g \in \mathbb{G}_q} \alpha_g \qquad \text{subject to} \tag{4.15a}$$

$$P_{Gg}^{min} \leq P_{Gg} \leq P_{Gg}^{max} \qquad \forall g \in \mathbb{G} \tag{4.15b}$$

$$Q_{inj,i}^{min} \leq Tr(\bar{\mathbf{Y}}_i \mathbf{W}) \leq Q_{inj,i}^{max} \qquad \forall i \in \mathcal{N} \tag{4.15c}$$

$$P_{inj,i} = \sum_{g \in \mathcal{G}_\rangle} P_{Gg} - P_{Di} = Tr(\mathbf{Y}_i W) \qquad \forall i \in \mathcal{N} \tag{4.15d}$$

$$Q_{inj,i}^{max} = \sum_{g \in \mathbb{G}_i} Q_{Gg}^{max} - Q_{Di} \qquad \forall i \in \mathcal{N} \tag{4.15e}$$

$$Q_{inj,i}^{min} = \sum_{g \in \mathbb{G}_i} Q_{Gg}^{min} - Q_{Di} \qquad \forall i \in \mathcal{N} \tag{4.15f}$$

$$|V_i^{min}|^2 \leq Tr(M_i W) \leq |V_i^{max}|^2 \qquad \forall i \in \mathcal{N} \tag{4.15g}$$

$$\begin{bmatrix} -(S_k^{max})^2 & Tr(\mathbf{Z}_{k_l}\mathbf{W}) & Tr(\bar{\mathbf{Z}}_{k_l}\mathbf{W}) \\ Tr(\mathbf{Z}_{k_l}\mathbf{W}) & -1 & 0 \\ Tr(\bar{\mathbf{Z}}_{k_l}\mathbf{W}) & 0 & -1 \end{bmatrix} \preceq 0 \qquad \forall k \in \mathcal{L} \tag{4.15h}$$

$$\begin{bmatrix} -(S_k^{max})^2 & Tr(\mathbf{Z}_{k_m}\mathbf{W}) & Tr(\bar{\mathbf{Z}}_{k_m}\mathbf{W}) \\ Tr(\mathbf{Z}_{k_m}\mathbf{W}) & -1 & 0 \\ Tr(\bar{\mathbf{Z}}_{k_m}\mathbf{W}) & 0 & -1 \end{bmatrix} \preceq 0 \qquad \forall k \in \mathcal{L} \tag{4.15i}$$

$$\begin{bmatrix} c_{g1}P_{Gg} + c_{g0} - \alpha_g & \sqrt{c_{g2}}P_{Gg} \\ \sqrt{c_{g2}}P_{Gg} & -1 \end{bmatrix} \preceq 0 \qquad \forall g \in \mathbb{G}^q \tag{4.15j}$$

$$\{\beta_g \geqslant m_{gt}(P_{Gg} - a_{gt} + b_{gt}), \forall t = 1, ..., r_g\} \qquad \forall g \in \mathbb{G}^{pw} \tag{4.15k}$$

$$\mathbf{W} \succeq 0 \tag{4.15l}$$

Following the work of [19, 20], a Lagrangian dual of the earlier program can be obtained. For that purpose, we define ψ_k, γ_k and μ_k to be the vector Lagrangian multipliers of inequalities relating to bounds on active power, reactive power and quadratic voltage terms respectively. A bar is used to denote multipliers on upper bounds whereas an underlined denotes those on lower bounds. An unconstrained Lagrangian multiplier is defined as λ_i. The multiplier λ_i is the locational marginal price (LMP) at node i. For each line and at each terminal u, define a 3×3 symmetric matrix $\mathbf{H}_{k_u}$[1]. Similarly, define a 2×2 symmetric matrix $\mathbf{R}_g$ for each generator with a quadratic cost function[2]. Alternatively, a Lagrangian multiplier ζ_{gt} is defined for each line segment t in the cost function of generators with piecewise-linear cost functions. It should be noted that $\mathbf{H}_{k_u}^{cd}$ and $\mathbf{R}_g^{cd}$ denote the (c, d) element of the lagrangian multiplier matrices $\mathbf{H}_{k_u}$ and $\mathbf{R}_g$ respectively. The dual of (4.15) can be therefore written as

$$\min \qquad (-\rho) \tag{4.16a}$$

subject to

$$\mathbf{A} \succeq 0 \tag{4.16b}$$

$$\mathbf{H}_{k_l} \succeq 0, \mathbf{H}_{k_m} \succeq 0 \qquad \forall k \in \mathcal{L} \tag{4.16c}$$

$$\mathbf{R}_g \succeq 0, \mathbf{R}_g^{11} = 1 \qquad \forall g \in \mathbb{G}^q \tag{4.16d}$$

$$\sum_{t=1}^{r_g} \zeta_{gt} = 1 \qquad \forall g \in \mathbb{G}^{pw} \tag{4.16e}$$

$$\left\{\lambda_i = c_{gi} + 2\sqrt{c_{g2}}\mathbf{R}_g^{12} + \bar{\psi}_g - \underline{\psi}_g \quad , \forall g \in \mathbb{G}_i^q\right\} \qquad \forall i \in \mathcal{N} \tag{4.16f}$$

$$\left\{\lambda_i = \sum_{t=1}^{r_g} \zeta_{gt} m_{gt} \quad , \forall g \in \mathbb{G}_i^{pw}\right\} \qquad \forall i \in \mathcal{N} \tag{4.16g}$$

$$\underline{\psi}_g, \bar{\psi}_g, \underline{\gamma}_i, \bar{\gamma}_i, \underline{\mu}_i, \bar{\mu}_i, \zeta_{gt} \geq 0 \tag{4.16h}$$

[1] $\mathbf{H}_{k_u}$ is the Lagrangian Multiplier corresponding to equations (4.15h) and (4.15i) at terminal u (of line k) equal to l_k and m_k respectively.

[2] $\mathbf{R}_g$ is the Lagrangian Multiplier corresponding to equation (4.15j).

where the scalar function ρ and the matrix valued function $\mathbf{A}$ are

$$\begin{aligned}
\rho = & \sum_{i \in \mathcal{N}} \Bigg\{ \lambda_i P_{Di} + \underline{\gamma}_i Q_i^{min} - \bar{\gamma}_i Q_i^{max} + \underline{\mu}_i \left(V_i^{min}\right)^2 - \bar{\mu}_i \left(V_i^{max}\right)^2 \\
& + \sum_{g \in \mathbb{G}_i^q} \left(\underline{\psi}_g P_{Gg}^{min} - \bar{\psi}_g P_{Gg}^{max} + c_{g0} - \mathbf{R}_g^{22} \right) \\
& - \sum_{g \in \mathbb{G}_q^{pw}} \sum_{t=1}^{r_g} (\zeta_{gt}(m_{gt} a_{gt} - b_{gt})) \Bigg\} \\
& - \sum_{k \in \mathcal{L}} \left\{ (S_k^{max})^2 \left(\mathbf{H}_{k_l}^{11} + \mathbf{H}_{k_m}^{11} \right) + \mathbf{H}_{k_l}^{22} + \mathbf{H}_{k_m}^{22} + \mathbf{H}_{k_l}^{33} + \mathbf{H}_{k_m}^{33} \right\}
\end{aligned} \tag{4.17}$$

$$\begin{aligned}
\mathbf{A} = & \sum_{i \in \mathcal{N}} \left\{ \lambda_i \mathbf{Y}_i + (\bar{\gamma}_i - \underline{\gamma}_i) \bar{\mathbf{Y}}_i + (\bar{\mu}_i - \underline{\mu}_i) \mathbf{M}_i \right\} \\
& + 2 \sum_{k \in \mathcal{L}} \left\{ \mathbf{H}_{k_l}^{12} \mathbf{Z}_{k_l} + \mathbf{H}_{k_m}^{12} \mathbf{Z}_{k_m} + \mathbf{H}_{k_l}^{13} \bar{\mathbf{Z}}_{k_l} + \mathbf{H}_{k_m}^{13} \bar{\mathbf{Z}}_{k_m} \right\}
\end{aligned} \tag{4.18}$$

4.2 Diagonal Dominance and the Optimal Power Flow Problem

This section presents the formulation and results of a relaxation that utilizes diagonal dominance in relaxing the OPF problem. First, the formulation and rationale are introduced. A comparison between the results of this relaxation and the SDP, $SDSOS_2$, $DSOS_2$ relaxations of the OPF problem follows.

4.2.1 Shortcomings of the DSOS relaxation and proposed approach

A DSOS polynomial can be represented by restrictions on matrices to be diagonally dominant. The program accordingly translates to an LP [10]. Consequently, the DSOS relaxation of the problem is a linear relaxation of the entirety of the feasible space. The performance of such a restriction can be seen to be lacking when implemented to relax the OPF problem [1].

The DSOS relaxation of the OPF problem can be seen to suffer from high optimality gaps. Additionally, as diagonally dominant restriction on a matrix of dimension $n \times n$ requires n^2

constraints, the DSOS relaxation requires an exploding number of linear inequalities. The resulting memory limitations of such a requirement prevent the application of this relaxation into networks with bus count as large as 300 [1]. It should be noted that by exploiting sparsity, this relaxation could be made to be much more tractable. However, the high optimality gaps do not encourage improving the scalability of that relaxation prior to producing a tighter variant.

The first order DSOS relaxation is a relaxation of the first order SOS relaxation obtained by imposing a polyhedral restriction of the entirety of the feasible set of the problem. It therefore follows that a tighter relaxation could be obtained if only a part of the feasible space is restricted in a linearly representable space. Since a relaxation highly concerns itself with the notions of optimality gap, speed and scalability it therefore makes sense to impose the relaxation on the most expensive segments of the problem.

The shortcomings of the $DSOS_2$ relaxation are characteristic of LP convex relaxations of general PPs [65, 66]. Accordingly, in this thesis, we propose that the problem is decomposed into expensive and non-expensive segments, and that only the most expensive segment is relaxed linearly. It follows that instead of relaxing the entire feasible space of the problem, we end up with a tighter relaxation targeting only the most computationally demanding constraints. In the following subsection, we present our tightened relaxation.

4.2.2 Formulation of the Proposed Relaxation

We attempt a decomposition of our problem into expensive and non-expensive segments. The decomposition is attempted by means of utilizing the Shor relaxation which is equivalent to, yet more tractable than, the first order SOS relaxation [1, 42].

The first order DSOS relaxation is a relaxation of the SOS formulation of the OPF problem. The proposed relaxation of this thesis is a relaxation on the SDP relaxation. The SDP relaxation is equivalent to, and more tractable than, the first order in the SOS hierarchy. As such, the proposed relaxation is, by proxy, a relaxation of a problem that is equivalent to, yet more tractable than, the SOS problem relaxed by the DSOS relaxation. A detailed proof for the equivalence between the two can be found in [42].

The main rationale behind this approach is that semidefinite solvers scale poorly when semidefinite constraints of a large dimension are being solved for. On the other hand, SDP solvers perform extremely well for constraints of small dimension. Accordingly, we reason that replacing these constraints by a quadratic number of linear inequalities could be counter-productive from the perspectives of both optimality gap and scalability. Considering Shor's relaxation, the semidefinite constraint on the $2n \times 2n$ matrix $\mathbf{W}$ of the primal accounts for the brunt of the computational demand of the optimization problem. The other semidefinite constraints on matrices of dimensions 2×2 and 3×3 are computationally of no concern. That is especially the case when we consider their contribution to our relaxation's tightness.

The work in [1] relaxes the OPF problem into an LP. That naturally follows from how a DSOS polynomial restriction is representable using a restriction to the diagonally dominant cone [10].

Reference [57] presents an approach for relaxing an SDP using the set of diagonally dominant matrices. A relaxation of the like is achieved by a diagonally dominant restriction on the dual of an SDP as outlined in Section 3.2. We propose using this approach in relaxing the Shor relaxation of the OPF problem.

Accordingly, we propose that a polyhedral relaxation is imposed only on the positive semidefinite constraint of the matrix $\mathbf{W}$. We achieve this by imposing a diagonally dominant restriction on the dual's positive semidefinite constraint $\mathbf{A}$, as defined in (4.18). The first order DSOS relaxation is a polyhedral relaxation of the entirety of the first order SOS relaxation. Accordingly, relaxing a part of the first order SOS relaxation is by proxy a tighter relaxation of the OPF problem than that presented by the DSOS variant. Due to the equivalence relation between the Shor relaxation and the first order SOS relaxation, we can conclude that this approach would produce a tighter variant of the first order DSOS relaxation. A similar approach was utilized in relaxing a UC problem in [46].

4.2.3 Experimental Setup

The different relaxations this thesis proposes are implemented using `Matlab`. We utilize the package `spotless` in formulating the different optimization programs (including the SDP Relaxation [20]) [68]. These different formulations are then solved using the solver `MOSEK` [69]. Default solver settings are utilized in the solution of all instances. The instances used in the evaluation of the different relaxations are those of distributed in the MatPower research platform [67]. For a consistent comparison with the results from [1], the same test cases and best bound values were utilized [42]. The optimality gaps are calculated utilizing equation (2.10), replacing the true objective value with the best bound for each case respectively. The experiments were conducted on a 2.9 GHz Intel Core i5 processor, 8 GB DDR3 RAM, 2016 13-inch Macbook Pro. The same experimental setup is used throughout this dissertation.

4.2.4 Experimental Results

This section outlines the results of the relaxation proposed in this chapter. The relaxation is achieved by means of a DD restriction on only the positive semidefinite constraint $\mathbf{A}$, as defined in (4.18). The remaining positive semidefinite constraints of equation (4.16) are maintained. Accordingly, we use the notation DD-SDP to represent this relaxation. Table 4.1 displays the resulting bounds of our relaxation relative to the SDP, PP-DSOS_2 and PP-SDSOS_2 relaxations of the OPF problem. Chordal decomposition is utilized in significantly improving the tractability of the SDP relaxation of the OPF problem. As such, and for a better frame of reference we outline the results for the SDP relaxation utilizing chordal decomposition in this subsection. A more detailed treatment of that relaxation follows in the next chapter.

We should note that the optimality gaps obtained are different from those in [1]. This could be attributed to a number of factors including but not limited to software and hardware differences. For example, operating system, solver edition, and processor architecture are some of the aspects that could have influenced these differences in bounds.

Table 4.1 Comparison of the DD-SDP Relaxation Bounds and Gaps

	Best	SDP Relaxation		PP-DSOS$_2$		PP-SDSOS$_2$		DD-SDP	
Instance	Bound	Bound	Gap (%)	Bound	Gap (%)	Bound	Gap (%)	Bound	Gap (%)
case9Q	5297.4	5297.41	0.00	4448	16.03	5220	1.46	5216	1.54
case14	8081.7	8081.7	0.00	6548.2	18.98	7660.1	5.22	7642.5	5.44
case30	576.8	576.9	0.02	355.1	38.44	567.75	1.57	565.21	2.01
case39	41889.1	41889.05	0.00	13378	68.06	41291	1.43	41217	1.60
case57	41738.3	41738.17	0.00	32518	22.09	41098	1.53	41003	1.76
case118	129372.4	129668.21	0.23	98698	23.71	126009.58	2.60	125934.12	2.66
case300	720031.0	719763.68	0.04	533496.97	25.91	706731.48	1.85	705710.15	1.99

As shown in Table 4.1, the DD-SDP relaxation provides significantly better bounds for the OPF problem when compared to those obtained by PP-DSOS$_2$ and comparable bounds to those of PP-SDSOS$_2$. The relaxation DD-SDP can be implemented on even larger cases contrary to the PP-DSOS$_2$ relaxation. The scaling behaviour of DD-SDP with respect to the other relaxations is demonstrated in Table 4.2.

Table 4.2 DD-SDP Relaxation, CPU Run-Time Comparison

Instance	SDP Relaxation	PP-DSOS$_2$	PP-SDSOS$_2$	DD-SDP
	CPU Time (s)			
case9Q	0.07	0.01	0.04	0.04
case14	0.04	0.02	0.07	0.03
case30	0.37	0.07	0.26	0.13
case39	0.48	0.09	1.10	0.45
case57	0.61	0.15	1.93	0.23
case118	1.24	0.52	10.42	0.58
case300	3.84	2.22	117.23	1.81

The results in Table 4.2 demonstrates the superiority of PP-DSOS$_2$ for the smaller test cases. For cases 9Q, 14, and 30, the relaxation DD-SDP has run times that are comparable to the relaxation PP-SDSOS$_2$. For larger test cases, DD-SDP outperforms PP-SDSOS$_2$. For the largest two test cases DD-SDP shows the fastest run times among the three relaxations. It can be clearly seen how for larger test cases, specifically the 118 and 300 bus systems, the run time of DD-SDP outperforms the other relaxations under consideration.

As such, we can conclude that it scales much more favorably compared to first order DSOS relaxation. However, some issues still remain relating to the scalability of the DD-SDP relaxation. That is the case considering how the number of linear inequalities is proportional $4n^2$. Similarly to PP-DSOS$_2$, we can reasonably conclude the exploding number of linear inequalities to be an issue for larger instances.

Another important observation is that our relaxation provides bounds that are relatively close to those of PP-SDSOS$_2$. This leads us to the conclusion that the inherent weakness of PP-DSOS$_2$ is not derived from the polyhedral representation of the portion of the feasible space defined by the positive semidefinite constraint on $\mathbf{A}$. Furthermore, since the second order conic representation of PP-SDSOS$_2$ provides relatively close bounds to DD-SDP, we can assume the remaining positive semidefinite constraints to be representable via second order conic constraints. We will utilize this assumption in later relaxing the remaining positive semidefinite constraints into SOC based constraints.

4.2.5 Summary

The DD-SDP relaxation can be seen to provide an improvement over the original PP-DSOS$_2$ relaxation. The bounds obtained via the DD-SDP relaxation are comparable to the far superior PP-SDSOS$_2$ relaxation. However, due to the number of linear inequalities this relaxation entails, the computational demand of DD-SDP still needs to be addressed.

The scalability issue of the DD-SDP relaxation and the potential of utilizing second order conic representations motivate the work of the following chapter. In the next chapter, we investigate the utilization of chordal decomposition in circumventing the scalability and tractability issues of this relaxation. Additionaly, we propose a couple of new restrictions on the dual to further improve the tractability of our relaxation.

Chapter 5

Sparsity and Structure

In this chapter, we discuss how to best exploit the inherent sparsity of power systems to improve the tractability of the relaxation proposed in Chapter 4. For that purpose, we investigate the utilization of chordal decomposition. We further explore the inherent problem structure and empirical results in further improving the tractability of our relaxation.

5.1 Chordal Sparsity

Sparsity is an inherent property of power networks. As such, sparsity exploiting techniques can and do in fact help in reducing the complexity of the different optimization programs in the power industry [9]. Chordal decomposition is one of the most prominent techniques for exploiting the sparsity of power systems. This approach provides a significant speedup for the SDP relaxation of the problem and has yielded impressive results in improving its tractability [20]. Accordingly, and since the relaxation of the previous chapter is based on the SDP relaxation of the problem, this section serves to investigate the exploitation of chordal decomposition in improving its tractability. In the first subsection, a review of the chordal decomposition exploiting SDP relaxation of [20] is provided. The second subsection follows with an implementation and discussion on the exploitation of chordal decomposition with respect to our relaxation.

5.1.1 SDP Relaxation Exploiting Chordal Sparsity

The brunt of the computational demand the SDP relaxation is in the $2n \times 2n$ positive semidefinite constraint $\mathbf{A}$, as defined in (4.18). Accordingly, several formulations were suggested to decompose this constraint into positive semidefinite constraints on smaller matrices, thus significantly improving the tractability of the SDP relaxation [20, 48, 70, 71].

This section outlines the application of a graph theoretic approach utilized in reducing the computational burden of the Shor relaxation. We begin by defining the relevant terms essential for the proper understanding of this approach. That is then followed by an explanation of its underlying theory and relevant application to the Shor relaxation [20, 48, 71].

In an un-directed graph, a cycle is a set of the edges in the graph forming a path which concludes at the first vertex of the path. A chord in the cycle is an edge connecting two non-adjacent nodes in the path forming the cycle. Subsequently, a graph is said to be chordal if there exists a chord for each of its cycles of a length larger than three. A clique in the graph is a subset of the vertices in the graph such that there exists an edge connecting every vertex pair in the subset. A maximal clique is a clique in the graph for which no inclusion in another clique of the graph holds. A clique tree is a maximum weight spanning tree of a graph in which its maximal cliques are represented by nodes in the tree. Concurrently, the number of shared buses between cliques is denoted by the weight on the edge forming a connection between their respective nodes in the tree.

Let $\mathcal{A}$ be a symmetric matrix with an underlying un-directed graph such that its entries are not completely determined. We define submatrices of $\mathcal{A}$ associated with each of the maximal cliques of the underlying graph of the matrix. The Matrix Completion Theorem states that the positive semidefinitness of the aforementioned matrices is both a sufficient and necessary condition for the existence of a positive semidefinite completion of $\mathcal{A}$ [72].

It follows from the above that a positive semidefinite constraint on a given matrix can be substituted by constraints on the sub-matrices associated with the maximal cliques of the matrix's underlying graph. Considering the highly sparse nature of power systems, significant computational savings were achieved by implementing this approach to the SDP relaxation of the OPF problem [20, 48].

Jabr in his work [48], utilizes this theorem in decomposing the large $2n \times 2n$ SDP constraint of the OPF problem. By definition, a decomposition of the like requires the determination of maximal cliques in a given graph. Maximal clique identification is achievable in linear time for chordal graphs [73]. Whereas power networks are not chordal, a chordal completion of a power network's graph is needed for the determination of maximal cliques. Following the proposition of Jabr [48], a chordal completion could be obtained by utilizing the minimum fill-in Cholesky decomposition approach of [74].

Requiring Cholesky decomposition on the absolute value of the imaginary part of the network's admittance matrix, we can see how this approach is limited to networks for which the aforementioned matrix is positive definite. However, that is not always the case. An alternate approach which circumvents this limitation was proposed in [20]. Jabr's chordal extension depends on the location of non-zero entries in the matrix in question. It follows that a different matrix which exhibits that same structure would provide for an equivalent extension. Accordingly, defining the matrix of the same structure, such that it is always positive definite would resolve the limitation of the earlier approach.

Another relevant breakthrough in [20] is the investigation of linking constraints and their influence on the computational complexity of the program. Maximal cliques in the graph do intersect. These intersections, which are represented by weights in the clique tree, require that equality be enforced on the shared elements of a maximal clique pair. Henceforth, equality constraints are introduced into the program. It follows that the larger the number of cliques, the larger the number of linking constraints needed. Accordingly, a clique merger algorithm was introduced in [20] to investigate the relationship between the number of linking constraints introduced in the program and the number of maximal cliques (i.e. size of the semidefinite constraints corresponding to the resulting maximal cliques).

To summarize, the utilization of chordal sparsity over the dual of the Shor relaxation translates to replacing the constraint of (4.16b) with the following set of constraints

$$\mathcal{A}_{\mathcal{C}_q} \succeq 0, \forall q \in \{1, 2, ..., p\} \tag{5.1}$$

where $\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_p$ are the maximal cliques of the chordal completion for the underlying graph of $\mathcal{A}$. It follows that $\mathcal{A}_{\mathcal{C}_q}$ is the sub-matrix corresponding to the maximal clique q. In the following section, we investigate the fidelity of using this approach on the relaxation proposed in this dissertation.

5.1.2 DD Relaxation and Chordal Sparsity

This section serves to investigate the utilization of the earlier technique in tandem with the relaxation proposed in Chapter 4. A clear limitation of the relaxation proposed in Chapter 4 is the exploding number of equality constraints used in relaxing the program [57]. That is apparent in how $4n^2$ inequality constraints are needed to restrict the $2n \times 2n$ matrix $\mathbf{A}$ in equation (4.18) to the cone DD_n. This section serves to investigate how chordal decomposition could be leveraged in reducing the number of inequality constraints needed for such a restriction on A.

Let ϵ be the set of edges in the underlying graph of $\mathcal{A}$ which conveniently represent the sparsity pattern in our problem. Cardinality of a set is defined by card(). The (i, j) element of $E_{\mathcal{C}_q} \in \mathbb{R}^{card(\mathcal{C}_q) \times 2n}$ is

$$E^{ij}_{\mathcal{C}_q} := \begin{cases} 1, \mathcal{C}_q(i) = j \\ 0, \text{Otherwise} \end{cases} \tag{5.2}$$

Similarly, to discuss how chordal decomposition could be utilized in tandem with the relaxation DD-SDP, we employ the following proposition for DD restrictions.

Proposition 1 (Miller, Zheng, Sznaier and Papachristodoulou [75])

$$\mathcal{A} \in DD_n(\epsilon, 0) \iff \exists \mathcal{A}_{\mathcal{C}_q} \in DD_{card(\mathcal{C}_q)} \;\; s.t. \;\; \mathcal{A} = \sum_{q=1}^{p} E_{\mathcal{C}_q}^T \mathcal{A}_{\mathcal{C}_q} E_{\mathcal{C}_q}.$$

From the earlier proposition, it can be seen how a diagonal dominant restriction on the matrix $\mathcal{A}$ can be achieved by imposing it on the sub-matrices $\mathcal{A}_{\mathcal{C}_q}$. The number of inequality constraints for a diagonally dominant restriction on a matrix is quadratically proportional to the dimension of the matrix. It therefore follows that enforcing this restriction on the sub-matrices corresponding to the maximal cliques of the chordal completion of the underlying graph of $\mathbf{A}$ in equation (4.18) could produce a more favorable result. The earlier proposition follows from

$$4n^2 \geq \sum_{q=1}^{p} \text{card}(\mathcal{C}_q)^2 \tag{5.3}$$

where $\sum_{q=1}^{p} \text{card}(\mathcal{C}_q)^2$ is the number inequality constraints required for a diagonal restriction on all sub-matrices corresponding to each of the maximal cliques of $\mathcal{A}$. It naturally follows that the sparser the matrix, the greater the apparent advantage of using this approach. Accordingly, the earlier proposal translates to

$$\mathcal{A}_{\mathcal{C}_q} \in DD_{\text{card}(\mathcal{C}_q)}, \forall q \in \{1, 2, ..., p\} \tag{5.4}$$

However, we should note that decomposing the constraint into those corresponding to the maximal cliques requires the introduction of linking constraints at the intersection of these cliques. In turn, the intersections add to the complexity of the program.

The work of [57] presents an iterative procedure by which the linear inequalities are modified. To obtain a more accurate approximation of the objective, the Cholesky factor of the answer is used in tightening the inner approximation of the positive semidefinite cone. This approach was implemented on the decomposed problem considering how Cholesky is of order $O(n^3)$ [75]. Improvement in the objective is guaranteed under the assumption that the matrix being factored is positive definite. Additionally, the convergence is highly problem dependent [57]. Initial investigation of this approach in tandem with chordal decomposition demonstrated slow convergence in addition to numerical difficulties for this problem.

5.2 Sparsity and Structure

In this section, we consider the properties (i.e. sparsity and inherent structure) of the diagonal dominance-based relaxation of this thesis. To better exploit sparsity, we take a deeper look at the implications accompanying the chordal decomposition of the earlier section. In the settings of the OPF problem and other problems of similar structures, we also investigate the possible reductions allowed by the problem structure.

5.2.1 Sparsity

Sparsity is a prevalent property in power systems. As a result, optimization problems in this domain inherit this property as part of their structure. Accordingly, significant computational gains could be achieved if sparsity is to be exploited properly. We are to investigate the proper use of sparsity in the DD exploiting relaxation of this thesis. We begin our investigation by a study of the chordal decomposition of the earlier section.

Tractability of semidefinite programs is significantly improved by decomposing a large sparse semidefinite constraint into smaller, yet equivalent, set of positive semidefinite constraints. Additionally, a diagonally dominant restriction on an $n \times n$ matrix requires n^2 constraints. Accordingly, the merit of chordal decomposition to a method that scales quadratically to matrix size could reasonably be inferred. Intuitively, a diagonally dominant restriction on these smaller matrices would require a smaller number of constraints when compared to the original matrix. Consequently, one might be tempted to utilize chordal decomposition in improving the tractability of a sparse DD restriction.

As counter-intuitive as it may seem, that however is not the case. In this section our goal is to show how a chordal decomposition, when used in tandem with diagonal dominance, is actually counterproductive. For that purpose, consider a symmetric matrix $\mathcal{A} \in S_4$ with a sparsity pattern demonstrated by the following chordal graph

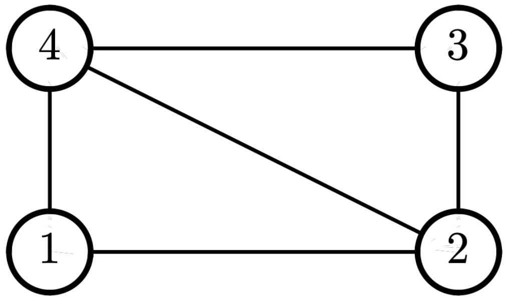

Fig. 5.1 Underlying graph of $\mathcal{A}$

Our main result is demonstrated through an example of a chordal decomposition on a sparse dual SDP which was provided in [76]. The maximal cliques defined by the graph in Fig. 5.1 are $\mathcal{C}_1 = \{1, 3, 4\}$ and $\mathcal{C}_2 = \{2, 3, 4\}$. Accordingly, the matrix $\mathcal{A}$ and the coupled sub-matrices obtained by chordal decomposition are as follows

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & 0 & \mathcal{A}_{14} \\ \mathcal{A}_{12} & \mathcal{A}_{22} & \mathcal{A}_{23} & \mathcal{A}_{24} \\ 0 & \mathcal{A}_{23} & \mathcal{A}_{33} & \mathcal{A}_{34} \\ \mathcal{A}_{14} & \mathcal{A}_{24} & \mathcal{A}_{34} & \mathcal{A}_{44} \end{bmatrix} \in DD_4 \Leftrightarrow \begin{cases} \mathcal{A}_{\mathcal{C}_1} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{A}_{14} \\ \mathcal{A}_{12} & a_1 & a_3 \\ \mathcal{A}_{14} & a_3 & a_2 \end{bmatrix} \mathcal{A}_{\mathcal{C}_2} = \begin{bmatrix} b_1 & \mathcal{A}_{23} & b_3 \\ \mathcal{A}_{23} & \mathcal{A}_{33} & Z_{34} \\ b_3 & \mathcal{A}_{34} & b_2 \end{bmatrix} \\ \mathcal{A}_{\mathcal{C}_1} \in DD_3, \mathcal{A}_{\mathcal{C}_2} \in DD_3 \\ a_1 + b_1 = \mathcal{A}_{22} \\ a_2 + b_2 = \mathcal{A}_{44} \\ a_3 + b_3 = \mathcal{A}_{24} \end{cases} \tag{5.5}$$

From the chordal decomposition demonstrated above, we can make a number of observations. Enforcing the restriction $\mathcal{A} \in DD_4$ normally requires $4^2 = 16$ constraints. However, we can represent the restriction by only 14 inequalities due to the zero off diagonal elements in $\mathcal{A}$.

Alternatively, enforcing an equivalent restriction on the sub-matrices $\mathcal{A}_{\mathcal{C}_q}$ requires $(3)^2 = 9$ constraints each. The computational demand of the program is further increased by the linking constraints coupling the two sub-matrices. The net program requires 18 inequality and 3 equality constraints.

The increase in constraints can be attributed to two factors. First, we notice how additional inequality constraints are introduced at the intersection of maximal cliques. For each diagonal element at an intersection node, one additional inequality constraint is introduced. Alternatively, two inequality constraints are required for each off diagonal element at the intersection of maximal cliques. It should be noted that since the matrices are symmetric, imposing the off diagonal constraints on either the upper or lower half is sufficient. Second, at the intersection of cliques, the coupling constraints introduce another complicating factor which further increases the complexity of our problem. That factor comes as a result of the non-chordal nature characteristic of a power system's underlying graph.

For the example above we have assumed the underlying graph of $\mathcal{A}$ to be chordal. However, that is not always the case in power systems [20]. Accordingly, the graph that we used in the decomposition outlined above is a chordal completion of a non-chordal graph better suited to represent a power system. As such, the process of chordal completion on a non-chordal graph which enables the utilization of chordal decomposition imposes additional constraints which can, for all intents and purposes, be avoided. Taking that into account, we notice how even being a chordal completion, the matrix $\mathcal{A} \in S_4$ of Fig 5.1 can be restricted to the cone DD_n using only 14 inequality constraints. Furthermore, if we were to consider an original non-chordal graph which was completed into that of Fig. 5.1, the diagonally dominant restriction would be of even less expense (a maximum of 12 constraints). In summary, the process of finding a chordal completion to a graph, in addition to the linking constraints result in adding new constraints to our problem.

The earlier example of Fig. 5.1 is that for a case of practically negligible sparsity when compared to applications such as power systems. However, the fact remains that by means of imposing the DD restriction on the non zero entries of the matrix, a more efficient implementation is obtained than that using chordal decomposition. Let n be the number of buses in the system and m the number of lines. It turns out that $2n + 8m$ linear inequalities are sufficient for the diagonally dominant restriction on $\mathbf{A}$ of equation (4.18). Accordingly, the size of the program scales linearly with respect to system size.

Alternatively, the iterative approach in [57, 75] depends on Cholesky factorizing of the matrix being restricted in tightening the restriction to the cone DD_n. Considering how that decomposition is of the order $O(n^3)$, the matrix dimensions are of a much detrimental concern for such an implementation. Accordingly, chordal decomposition might still be of use. However, similar to the work done in [20], a recombination algorithm might play a vital role in producing a variant of optimal tractability. For such an approach, our relaxation could provide for a fast solution of the first iteration in the iterative approach proposed in [57, 75]. In the following section, we aim at utilizing the problem structure in further improving the scalability of the linearly scaling sparsity exploiting relaxation of this subsection.

5.2.2 Structure

The steady state analysis of power systems operation is done in terms of phasor quantities. As discussed earlier, the OPF problem is an optimization problem which relates to the steady state line flows and operating parameters of a power system. Therefore, the OPF problem is originally an optimization problem in the complex domain [2]. However, most currently available SDP solvers are designed for real valued SDPs. Accordingly, the SDP relaxation of the problem is that of an equivalent formulation in the real domain [9]. The equivalent real valued reformulation of the SDP relaxation holds several structural properties which prove to be advantageous to the work of this thesis. This section serves to investigate those properties and how to best leverage them in the context of our relaxation.

We first begin by outlining the relationship between a positive semidefinite constraint for an SDP in the complex domain and its equivalent reformulation in the real domain. Let $\hat{b}_z$

be a predefined real scalar. Define the matrices $\hat{A}_z$ and $\hat{C}$ as the data matrices for a complex SDP. Accordingly, an SDP in the complex domain could be defined as follows

$$\begin{aligned} &\min && \langle \hat{C}, \Psi \rangle \\ &\text{subject to} && \\ & && \langle \hat{A}_z, \Psi \rangle = \hat{b}_z, z = 1, ..., w, \\ & && \Psi \succeq 0. \end{aligned} \tag{5.6}$$

where $\Psi \in \mathbb{C}^{n \times n}$ is a Hermitian matrix.

The matrix Ψ is positive semidefinite if and only if

$$x^H \Psi x \geq 0 \qquad \forall x \in \mathbb{C}^n \tag{5.7}$$

This equation can be reformulated into

$$\begin{bmatrix} Re(x) & Im(x) \end{bmatrix} \begin{bmatrix} Re(\Psi) & -Im(\Psi) \\ Im(\Psi) & Re(\Psi) \end{bmatrix} \begin{bmatrix} Re(x) \\ Im(x) \end{bmatrix} \geq 0 \qquad \forall \begin{bmatrix} Re(x) \\ Im(x) \end{bmatrix} \in \mathbb{R}^{2n} \tag{5.8}$$

Therefore, the positive semidefinite constraint on the Hermitian matrix $\Psi \in \mathbb{C}^{n \times n}$ can be replaced by a positive semidefinite constraint on a matrix in $\mathbb{R}^{2n \times 2n}$. Accordingly, it can be seen how Ψ is positive semidefinite if and only if the following holds [7, 77]

$$\begin{bmatrix} Re(\Psi) & -Im(\Psi) \\ Im(\Psi) & Re(\Psi) \end{bmatrix} \in P_{2n} \tag{5.9}$$

Similarly, and by Lagrangian duality, the same could be proven for the positive semidefinite constraint of the dual of the semidefinite program shown in equation (5.6) (i.e., the positive semidefinite constraint of the dual has the structure displayed in equation (5.8)).

A matrix is diagonally dominant if it satisfies the relation of equation (3.4). For any matrix of the structure shown in (5.8), several reductions could be made when imposing a DD restriction. We can clearly see how imposing a diagonally dominant restriction on the matrix of the structure displayed above could be done by satisfying the relation on either the rows on the top or bottom half of the matrix. Additionally, and since Ψ is a Hermitian matrix, the imaginary part of Ψ is skew-symmetric. The absolute value in equation (3.4) requires two inequalities and one new variable. Due the skew-symmetric nature of $Im(\Psi)$, the same definition of the absolute value over the entries in the top off diagonal elements of $Im(\Psi)$ applies for the lower off diagonal entries in that skew-symmetric matrix. As such, instead of defining the absolute value for all off diagonal entries, numbering $(n^2 - n)$ in $Im(\Psi)$, it is sufficient to define them for either the upper or lower off diagonal entries in the matrix. This only requires $((n^2 - n)/2)$ variables and $((n^2 - n))$ linear inequalities instead of $(n^2 - n)$ variables and $2(n^2 - n)$ linear inequalities.

The matrix we are interested in is matrix $\mathbf{A}$ of the dual formulation of the SDP relaxation shown in equation (4.16). In addition to the derivation above, the earlier structural properties we discussed can be also inferred from the definition of $\mathbf{A}$ in equation (4.18). These structural properties can be exploited in the DD restriction on the matrix $\mathbf{A}$.

As a consequence of matrix $\mathbf{A}$ having the structure shown in (5.8), it is sufficient to apply the relation in (3.4) to only the first n rows of $\mathbf{A}$ for $\mathbf{A} \in DD_{2n}$ to hold. As such, the number of inequalities is reduced to $n + 6m$.

Furthermore, Matrix $\mathbf{A}$ is a matrix in $\mathbb{R}^{2n \times 2n}$. This matrix is the real value formulation of a semidefinite constraint on a Hermitian matrix in $\mathbb{C}^{n \times n}$. As such, the off diagonal block matrices of $\mathbf{A}$ correspond to the imaginary part of the aforementioned Hermitian matrix. Therefore, it follows that the off diagonal block matrices of matrix $\mathbf{A}$ will be skew-symmetric. By virtue of their skew-symmetric nature, the number of linear inequalities is further reduced by $2m$. In summary, by exploiting these structural properties of $\mathbf{A}$, the number of linear inequalities reduces from $(2n+8m)$ to $(n+4m)$.

5.3 Empirical Relaxation

This section discusses a relaxation aiming at further increasing the speed of the relaxation outlined earlier at no apparent increase in optimality gaps.

The diagonally dominant restriction on the matrix $\mathbf{A}$ appears to operate in a manner that would allow for a reduction in the size of the relaxation. Empirically, the entries corresponding to the imaginary block matrices in the structure shown in (5.8), are being set to very small values. As discussed earlier, we represent each entry in either the upper or lower off diagonal part of $\mathbf{A}$ using two linear inequalities. Accordingly, having the prior knowledge that certain off-diagonal entries are going to be close to zero allows for a reduction in the size of the relaxation. We propose forcing such entries to zero. That would entail replacing $2m$ inequalities by m equalities. As such, we end up replacing $2m$ inequalities of the form $x \leq a$ and $x \geq -a$ with $2m$ inequalities of the form $x \leq 0$ and $x \geq 0$. It therefore follows that the feasible set of our relaxation is reduced.

We move to discuss how this would yield a relaxation of the original problem. To prove that the earlier procedure produces a valid relaxation of the OPF problem, consider the following dual SDP

$$\begin{aligned}
&\max_{\mathbf{y}\in\mathbb{R}^w} && b\ \mathbf{y} \\
&\text{subject to} && \\
& && \mathcal{A} = C - \sum_{z=1}^{w} \mathbf{y}_z A_z, \qquad (5.10) \\
& && \mathcal{A} \in P_{2n}.
\end{aligned}$$

where b, C and A_z are defined as in Section 3.1.

Let $\mathcal{A}$ be defined by the following block matrix

$$\hat{\mathcal{A}} = \begin{bmatrix} \hat{\mathcal{A}}_1 & \hat{\mathcal{A}}_2 \\ \hat{\mathcal{A}}_3 & \hat{\mathcal{A}}_4 \end{bmatrix} \tag{5.11}$$

where $\hat{\mathcal{A}}_1, \hat{\mathcal{A}}_2, \hat{\mathcal{A}}_3$ and $\hat{\mathcal{A}}_4$ are matrices in $\mathbb{R}^{n \times n}$. We want to prove that setting the entries of block matrices $\hat{\mathcal{A}}_2$, $\hat{\mathcal{A}}_3$ to $\mathbf{0}$ and restricting the matrices $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_4$ to be diagonally dominant would yield a valid relaxation of the primal SDP shown in equation (3.2). By setting the matrices $\hat{\mathcal{A}}_2$ and $\hat{\mathcal{A}}_3$ to zero, the resulting matrix $\mathcal{A}$ is

$$\hat{\mathcal{A}} = \begin{bmatrix} \hat{\mathcal{A}}_1 & \mathbf{0} \\ \mathbf{0} & \hat{\mathcal{A}}_4 \end{bmatrix} \tag{5.12}$$

From (5.12), it can be clearly seen that $\hat{\mathcal{A}}$ is diagonally dominant if and only if the block matrices $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_4$ are diagonally dominant. It therefore follows that this procedure results in a restriction on the matrix $\mathcal{A}$ to part of the diagonally dominant cone where the entries of the off diagonal block matrices are 0. Accordingly, and by the extreme ray interpretation of Section 3.2, this procedure can be seen to produce a valid relaxation of the primal.

We further attempt to relax the problem by imposing an SDD restriction on the remaining positive semidefinite constraints. Based on the similarity between the bounds obtained via the DD-SDP and PP-SDSOS$_2$ relaxations, we assume that this relaxation would not suffer from a significant increase in optimality gaps.

5.3.1 Experimental Results

Section 5.2.1 provided a detailed explanation of the disadvantages of using chordal decomposition. In summary, this technique requires the introduction of linking constraints between the sub-matrices corresponding to the maximal cliques in a graph. Additionally, the application of chordal decomposition in power systems necessitates the chordal completion of the non-chordal power system graph. This introduces new variables to the matrix in question. Consequently, those variables translate to an increase in the number of constraints needed for a DD restriction on the resulting sub-matrices.

Following the work in [20], an upper bound was imposed on the number of maximal cliques utilized in decomposing the positive semidefinite constraint of the SDP relaxation in Chapter 4. Accordingly, a bound of 65% of the number of maximal cliques was utilized over all test cases. Better run times may be obtained using a bound specifically tailored for each instance. Since the matrix combination algorithm is out of the scope of this dissertation, please consult [20] for more details on the topic. In the case of the DD-SDP relaxation, the increase in computational demand chordal decomposition introduces comes at no advantage. The remainder of this subsection is dedicated to the results of the different relaxations this chapter proposes. Accordingly, the numerical results are displayed in Tables 5.1 and 5.2.

Table 5.1 Proposed relaxations, gap comparison

Instance	Best Bound	SDP Relaxation	DD-SDP	SAS-DD-SDP	O-SAS-DD-SDP	SAS-DD-SDD	O-SAS-DD-SDD
		Gap (%)					
case9Q	5297.4	0.013	1.537	1.537	1.537	1.537	1.537
case14	8081.7	0.0025	5.435	5.435	5.435	5.435	5.435
case30	576.8	0	2.009	2.009	2.009	2.009	2.009
case39	41889.1	0.082	1.605	1.605	1.605	1.605	1.605
case57	41738.3	0.001	1.762	1.762	1.762	1.762	1.762
case118	129372.4	0.218	2.658	2.658	2.659	2.658	2.66
case300	720031.0	0.044	1.989	1.989	2.019	1.989	2.019

[1] DD-SDP: Relaxation Obtained by a DD restriction on matrix **A** in equation (4.18).

[2] SAS-DD-SDP: Relaxation obtained by exploiting sparsity and removing redundant and duplicate constraints from DD-SDP.

[3] O-SAS-DD-SDP: Relaxation obtained by setting the off diagonal block matrices of **A** to zero in SAS-DD-SDP.

[4] SAS-DD-SDD: Relaxation obtained by an SDD restriction on the 3×3 and 2×2 postivie semidefinite constraints in SAS-DD-SDP.

[5] O-SAS-DD-SDD: Relaxation obtained by setting the off diagonal block matrices of **A** to zero and an SDD restriction on the 3×3 and 2×2 postivie semidefinite constraints in SAS-DD-SDP.

Table 5.2 Proposed relaxations, CPU run-time comparison

Instance	SDP Relaxation [20]	DD-SDP	SAS-DD-SDP	O-SAS-DD-SDP	SAS-DD-SDD	O-SAS-DD-SDD
	CPU Time (s)					
case9Q	0.07	0.04	0.03	0.04	0.02	0.02
case14	0.04	0.03	0.03	0.01	0.02	0.01
case30	0.37	0.13	0.09	0.13	0.03	0.02
case39	0.48	0.45	0.29	0.39	0.10	0.07
case57	0.61	0.23	0.06	0.03	0.07	0.03
case118	1.24	0.58	0.20	0.15	0.12	0.10
case300	3.84	1.81	0.40	0.30	0.31	0.25

We denote the sparsity and structure exploiting relaxation as SAS-DD-SDP. We obtain this relaxation by means of removing redundant constraints in the DD-SDP relaxation of Chapter 4. The process is outlined in Section 5.2. As can be seen in Table 5.1, the bounds obtained via this relaxation are identical to its parent relaxation. Those bounds are achieved at a significantly lower computational cost as shown in Table 5.2. The speedup comes as a natural consequence of the reduction in the number of constraints and variables (with respect to system size) required for imposing a DD restriction.

The number of constraints in DD-SDP scales quadratically with the number of busses in the system. Consequently, this would hinder the implementation of this relaxation to larger test cases. On the other hand, the number of constraints needed for the relaxation SAS-DD-SDP scales linearly with the number of busses and lines in the system. First, the sparsity of the matrix is utilized in removing redundant constraints on zero elements of the matrix $\mathbf{A}$. This achieves linear scaling in the number of the busses and lines in the system. Second, the structure of $\mathbf{A}$ is utilized in halving the number of necessary constraints. To achieve this reduction, duplicates in the linear inequalities necessary for the diagonally dominant restriction are remove. As such, the speedup obtained by the relaxation SAS-DD-SDP comes at no cost in optimality gap.

Empirically, the values of the off diagonal block matrices at the solution of DD-SDP are set to small, close to zero, values. Therefore, we implement a relaxed variant of SAS-DD-SDP which sets those matrices to zero and denote it by O-SAS-DD-SDP. This relaxation replaces two inequalities constraints by an equality constraint in addition to eliminating the need for additional m variables in the relaxation formulation. This results in reducing the feasible set of the dual and thus further relaxing the problem. Since these entries are already set to small values in SAS-DD-SDP, imposing this restriction on the off diagonal entries has little to no influence on the optimally gaps.

The relaxation O-SAS-DD-SDP is slower for small test cases. Alternatively, the solution time for larger instances is an improvement over that of SAS-DD-SDP. For small test cases, the gaps resulting from this relaxation are identical to those of SAS-DD-SDP. However, a marginal increase in the gaps can be noticed for the 118 and 300 bus test cases.

Building on the observations over the results of Chapter 4, we replace the remaining positive semidefinite constraints in the SAS-DD-SDP relaxation by second order conic constraints. That is achieved by replacing the positive semidefinite constraints of equations (4.16c) and (4.16d) by an SDD restriction. We identify this relaxation as SAS-DD-SDD. This approach yields a more consistent speed-up than that of the O-SAS-DD-SDP at lower cost in the gaps for larger test cases. However, one point of comparison between these two relaxations is that the SAS-DD-SDD relaxation introduces an extremely small increase in the optimality gaps for small test cases, contrary to O-SAS-DD-SDP.

In the test cases utilized for the evaluation of our relaxations, only cases 9Q, 30 and 39 include apparent power line flow limits. Accordingly, only these cases could serve for evaluating the impact of the SDD restriction on the matrices $\mathbf{H}_{k_l}$ and $\mathbf{H}_{k_m}$. From the earlier results, we can reasonably conclude the reason behind the poor bounds of the PP-DSOS$_2$ relaxation to be in its attempt at linearly representing the constraint $\mathbf{R}_g \succeq 0$. The aforementioned constraint corresponds to the quadratic cost function on generator g. It naturally follows for a linear relaxation to have such a negative impact on the bounds.

In the O-SAS-DD-SDD relaxation we combine the aforementioned techniques in creating the fastest and most scalable implementation of our relaxation. Accordingly, we impose the SDD restriction of the SAS-DD-SDD, as well as set the elements of the off diagonal block matrices in **A** to zero. Similar to its parent relaxations, the change in optimality gaps introduced by this relaxation can be seen to be negligible.

The resulting objective bounds seem to be proportional to the upper bound on the absolute values of the off diagonal entries. The weakness of this relaxation can be attributed to its failure in reasonably relating the imaginary values (off-diagonal block matrices) to their real counterparts (diagonal block matrices) of the matrix **A**. Therefore, we can reasonably assume that finding valid upper bounds on these imaginary values can therefore provide for a tighter relaxation.

Overall, SAS-DD-SDP provides identical bounds to those of DD-SDP. This comes by virtue of their identical representation of the relaxed feasible set. The number of linear inequalities entailed by SAS-DD-SDP scales linearly, contrary to the quadratically scaling DD-SDP. The diagonal dominance restriction forces the entries of the off diagonal block matrices in **A** to small values. As such, forcing those entries to zero provides a very small increase in the optimality gaps while improving the run time for larger test cases. An SDD restriction of the remainder positive semidefinite constraints in the problem results in a speedup over all the cases. That is achieved at the cost of a very small increase to the optimality gaps.

We have demonstrated that the least expensive DD restriction could be used in obtaining optimality gaps on bar with those obtained via an SDD restriction on the entirety of the feasible set of the problem (namely the SDSOS relaxation). The sparsity, structure and general behavior of the relaxation could be utilized in further reducing the computational requirements of the relaxation to yeild a better compromise between scalability and obtimality gaps than those of the first order DSOS and SDSOS relaxations.

The SAS-DD-SDP relaxation is an improvement over the first order DSOS relaxation in terms of both scalability and optimality gap. However, there exists several relaxations which provide a good balance between scalability and optimality gap such as the SOCP and QC relaxations. The SOCP and QC relaxations employ second order cone constraints in relaxing the positive semidefinite constraint of the semidefinite relaxation of the OPF problem. Alternatively, the relaxations proposed in this thesis relax that constraint utilizing linear inequalities. As such, the second order cone representation inherent to the QC and SOCP relaxations results in lower optimality gaps. We must also bear in mind that to fully assess how the different relaxation of this thesis perform in comparison to other relaxations, the difference in scalability needs to be properly addressed.

Setting the off diagonal elements to zero as well as replacing the 3×3 and 2×2 semidefinite constraints of equations (4.16c) and (4.16d) by second order cone based constraints improves the scalability of the relaxation SAS-DD-SDP. For the test cases under study, this improvement is obtained at almost no cost in the optimality gap. However, this implementation could yield poorer bounds for test cases of other power systems. Furthermore, the semidefinite relaxation produces higher optimality gaps for the challenging test cases of power systems operating under congestion [78]. We can assume that applying this thesis's relaxations on such challenging test cases will result in higher optimality gaps. That is the case since the different relaxations of this thesis are relaxations of the semidefinite relaxation of the OPF problem. We should also note that the influence of setting the off diagonal elements to zero and replacing the 3×3 and 2×2 semidefinite constraints by second order cone constraints could be of higher consequence for such test cases.

5.4 Summary

The relaxation DD-SDP exploits diagonal dominance in only relaxation the semidefinite constraint on the matrix $\mathbf{A}$. This relaxation achieved a significant improvement in the bounds obtained by means of using diagonal dominance to relax this problem. Due to the nature of power systems, sparsity could be utilized in improving the tractability of this relaxation. As such, this chapter served to investigate the optimal utilization of sparsity in tandem with the relaxation DD-SDP. We demonstrated the superiority of direct sparsity exploitation when compared to chordal decomposition. We further investigated the influence of replacing the remaining semidefinite constraints by an SDD restriction, thus creating a hybrid DD and SDD based relaxation. We concluded this chapter by a relaxation based on empirical observations. This relaxation forces the off diagonal block matrices of $\mathbf{A}$ to zero thus reducing the feasible set of our problem. Utilizing DD and SDD restrictions, in addition to forcing the off diagonal block matrices of $\mathbf{A}$ to zero prove to be the most scalable implementation of our relaxation at relatively no cost in the optimality gap. The next chapter provides a summary for the work presented in this thesis in addition to some future directions.

Chapter 6

Conclusion and Future Work

In this thesis, we examined the utilization diagonal dominance in relaxing the optimal power flow problem. We made use of the pre-existing first order DSOS relaxation of the OPF problem, as well as developed a much more tractable and a tighter variant exploiting the same underlying principle (i.e., diagonal dominance). This chapter summarizes the main findings of our work and the questions they raise in the context of future research.

6.1 Conclusions

In Chapter 2 we provided a survey of the literature surrounding the OPF problem. Chapter 3 followed with a detailed discussion on the mathematical preliminaries, as well as the most relevant OPF relaxations. Chapter 4 included an alternative way of utilizing diagonal dominance whereas Chapter 5 tackled the tractability of our proposed relaxation.

By definition, the PP-DSOS$_2$ relaxation is obtained via a polyhedral restriction on the entire feasible space of the dual of the original OPF problem. Accordingly, the PP-DSOS$_2$ relaxation is an LP. In Chapter 4, we discussed the apparent disadvantages of this relaxation in the context of the OPF problem. As demonstrated by a previous implementation, this relaxation has two significant limitations. First, the optimality gaps this relaxation displays are poor. Second, despite being an LP, the PP-DSOS$_2$ relaxation suffers from poor scalability.

The poor scalability of this relaxation comes as a consequence of the large number of linear inequalities it requires [1].

We then proposed a tighter variant of the PP-DSOS$_2$ relaxation. Our variant utilizes the same underlying principle of diagonal dominance. The DSOS$_2$ relaxation is a relaxation of the SOS_2 relaxation of the OPF problem. Accordingly, we use the SDP relaxation (which is equivalent to the SOS_2 relaxation) in creating a tighter variant of PP-DSOS$_2$. The computational demand of the SDP relaxation of the OPF problem can be attributed to a segment of the entire program, specifically the positive semidefinite constraint on the $2n \times 2n$ matrix $\mathbf{A}$. Whereas the PP-DSOS$_2$ relaxation relaxes the entirety of the problem, we suggested that only the positive semidefinite constraint on the matrix $\mathbf{A}$ is restricted to the cone DD_{2n}. A relaxation of the other parts (obtained by a restriction on the matrices $\mathbf{H}_{k_u}$ and $\mathbf{R}_g$) introduces linear inequalities which hinder the scalability and performance of the relaxation. Accordingly, the constraints concerned with these matrices are maintained as they introduce no significant computational overhead. As a result, our relaxation provides significant improvements in the bounds over the objective function.

In Chapter 5, we explored how to best utilize the inherent sparsity of power systems in improving the tractability of our relaxation. We then demonstrated how structure could be utilized in further improving our final relaxation. Chordal decomposition demonstrated exceptional results when applied to improve the tractability of the SDP relaxation of the OPF problem [20]. Therefore, we investigate its application with respect to our relaxation. We demonstrate how the direct exploitation of sparsity would prove to be most beneficial. That is the case considering how chordal decomposition introduces additional constraints, therefore increasing the complexity of the program. Having established that a direct exploitation of sparsity is optimal, we move to exploit the structure of the problem. Exploiting the Hermitian and skew-symmetric structures in our problem, we are able to further reduce the number of necessary linear inequalities from $(2n + 8m)$ to $(n + 4m)$, where n and m are the number of busses and lines in the power system, respectively.

Empirical evidence suggested that the off-diagonal block matrices of $\mathbf{A}$, which correspond to the imaginary part of the Hermititan semidefinite constraint, are being set to zero. Finally, we explored how forcing these matrices to zero would affect the optimality gaps and scalability of our relaxation. By virtue of setting these block matrices to zero, $2m$ linear inequalities are replaced by m equalities. The optimality gaps do not seem to significantly deviate from those obtained by our original relaxation of the problem. This approach results in an increase in solution time for smaller instances but a decrease for the larger test cases. It should be noted that the value of the upper bound imposed on the absolute value of the entries in the block off diagonal matrices is directly proportional to the objective value.

Considering how the bounds of our relaxation are close to those obtained by the $SDSOS_2$ relaxation, we proposed an SDD restriction on the remaining positive semidefinite constraints. No noticeable change in the bounds was observed, whereas a notable and consistent speedup is demonstrated for all the test cases.

The relaxation, combining the past two restrictions, represents the fastest variant of the relaxations proposed in this dissertation. No significant change in the bounds was noticed in the case of this relaxation. Ultimately, a relaxation is evaluated based on the lower limits they provide, their speed and scalability. In that regard, we believe that our implementation efficiently utilizes diagonal dominance in providing a relaxation that reasonably compromises between scalability and tightness.

6.2 Future Work

In this section we aim to provide a number of paths and unanswered questions that could be factored into tightening the relaxations of this thesis. This relaxation provides bounds similar to those provided by copper plate relaxation of [54]. As such, one path forward could be an investigation into the similarities between these two relaxations. Furthermore, an interesting avenue could be the investigation into the optimality gaps and run times with respect to other relaxations for larger test cases and for test cases emulating specific operating conditions.

Out empirical investigation demonstrated that the diagonally dominant restriction on the matrix **A** of equation (4.18) forces the entries of the off-diagonal block matrices in **A** to small values. Accordingly, it can be inferred that this relaxation does not properly account for the relationship between the imaginary and real parts in the original Hermitian positive semidefinite constraint of our problem. This can be empirically verified by means of restricting the diagonal block matrices in the original SDP to be diagonal dominanat while maintaining a positive semidefinite constraint on **A**. This shows how such a restriction does not account for much of the optimality gaps of our DD based SDP relaxation. Accordingly, finding a way to better relate the imaginary and real block matrices would, in essence, yield a significantly tighter relaxation of the OPF problem.

Another interesting direction would be an investigation into the upper bounds of the absolute values in the off diagonal block matrices. The upper bound on the absolute values of the off-diagonal block matrices is directly proportional to the value of the objective function. Accordingly, the proper determination of these upper bounds such that the program is still a relaxation of the primal could yield a tighter relaxation. These upper bounds could be also specified in a manner that the resulting program is instead a restriction or an approximation of the OPF problem.

Lastly, the work in [79] serves to provide a fast method by which sparse sum of squares program can be solved. Comparison to the more conservative DSOS and SDSOS relaxations demonstrate promising results. However, the performance of such an approach on power systems remains to be seen.

References

[1] X. Kuang, B. Ghaddar, J. Naoum-Sawaya, and L. Zuluaga, "Alternative LP and SOCP hierarchies for ACOPF problems," *IEEE Transactions on Power Systems*, vol. 32, pp. 2828–2836, 2016.

[2] A. Gomez-Exposito, A. Conejo, and C. Cañizares, *Electric Energy Systems: Analysis and Operation, 2nd ed.* Boca Raton: FL: CRC Press, 2018.

[3] J. Shea, "Electric power system applications of optimization [book reviews]," *IEEE Electrical Insulation Magazine*, vol. 17, p. 62, 2001.

[4] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin de la Societe Francoise des Electriciens*, vol. 8, no. 3, p. 431–447, 1962.

[5] A. Castillo and R. P. O'Neill, "Survey of approaches to solving the ACOPF (OPF paper 4)," tech. rep., US Federal Energy Regulatory Commission, 2013.

[6] M. Niu, C. Wan, and Z. Xu, "A review on applications of heuristic optimization algorithms for optimal power flow in modern power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 2, no. 4, pp. 289–297, 2014.

[7] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[8] J. A. Taylor, *Convex Optimization of Power Systems*. Cambridge University Press, 2015.

[9] D. K. Molzahn and I. A. Hiskens, *A Survey of Relaxations and Approximations of the Power Flow Equations*. Now Foundations and Trends, 2019.

[10] A. Ahmadi and A. Majumdar, "DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization," *SIAM Journal on Applied Algebra and Geometry*, vol. 3, pp. 193–230, 2019.

[11] B. C. Lesieutre and I. A. Hiskens, "Convexity of the set of feasible injections and revenue adequacy in FTR markets," *IEEE Transactions on Power Systems*, vol. 20, pp. 1790–1798, 2005.

[12] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, pp. 745 – 751, 06 1974.

[13] A. Soroudi and T. Amraee, “Decision making under uncertainty in energy systems: State of the art,” *Renewable and Sustainable Energy Reviews*, vol. 28, pp. 376–384, 12 2013.

[14] D. Bienstock and A. Verma, “Strong NP-hardness of AC power flows feasibility,” *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.

[15] K. Lehmann, A. Grastien, and P. Van Hentenryck, “AC-feasibility on tree networks is NP-hard,” *IEEE Transactions on Power Systems*, vol. 31, pp. 798–801, 10 2014.

[16] W. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, “Local solutions of the optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 28, pp. 4780–4788, 11 2013.

[17] B. Stott, J. Jardim, and O. Alsac, “DC power flow revisited,” *IEEE Transactions on Power Systems*, vol. 24, pp. 1290 – 1300, 09 2009.

[18] J. Glover, M. Sarma, and T. Overbye, *Power System Analysis and Design*. Thompson Learning, 01 2008.

[19] J. Lavaei and S. Low, “Zero duality gap in optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 27, pp. 92–107, 02 2012.

[20] D. Molzahn, J. Holzer, B. Lesieutre, and C. Demarco, “Implementation of a large-scale optimal power flow solver based on semidefinite programming,” *IEEE Transactions on Power Systems*, vol. 28, pp. 3987–3998, 11 2013.

[21] D. K. Molzahn, *Application of semidefinite optimization techniques to problems in electric power systems*. PhD thesis, Ph. D. Dissertation, University of Wisconsin-Madison, 2013.

[22] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*, vol. 7. New York: McGraw-hill, 01 1994.

[23] M. Baran and F. F. Wu, “Optimal sizing of capacitors placed on a radial distribution system,” *IEEE Transactions on Power Delivery*, vol. 4, pp. 735–743, 1 1989.

[24] S. H. Low, “Convex relaxation of optimal power flow—Part I: Formulations and equivalence,” *IEEE Transactions on Control of Network Systems*, vol. 1, pp. 15–27, 03 2014.

[25] S. H. Low, “Convex relaxation of optimal power flow—Part II: Exactness,” *IEEE Transactions on Control of Network Systems*, vol. 1, pp. 177–189, 06 2014.

[26] J. Jarjis and F. D. Galiana, “Quantitative analysis of steady state stability in power networks,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, pp. 318–326, 1 1981.

[27] H.-P. Chao and S. Peck, “A market mechanism for electric power transmission,” *Journal of Regulatory Economics*, vol. 10, pp. 25–59, 07 1996.

[28] D. K. Molzahn, "Identifying and characterizing non-convexities in feasible spaces of optimal power flow problems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, pp. 672–676, 5 2018.

[29] B. Polyak and E. Gryazina, "Convexity/nonconvexity certificates for power flow analysis," in *Advances in Energy System Optimization*, pp. 221–230, Springer, 2017.

[30] D. K. Molzahn, "Computing the feasible spaces of optimal power flow problems," *IEEE Transactions on Power Systems*, vol. 32, pp. 4752–4763, 11 2017.

[31] H. Chiang and C. Jiang, "Feasible region of optimal power flow: Characterization and applications," *IEEE Transactions on Power Systems*, vol. 33, pp. 236–244, 1 2018.

[32] M. R. Narimani, D. K. Molzahn, D. Wu, and M. L. Crow, "Empirical investigation of non-convexities in optimal power flow problems," in *2018 Annual American Control Conference (ACC)*, pp. 3847–3854, IEEE, 2018.

[33] J. A. Momoh, R. Adapa, and M. E. El-Hawary, "A review of selected optimal power flow literature to 1993. I. Nonlinear and quadratic programming approaches," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 96–104, 1999.

[34] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. II. Newton, linear programming and interior point methods," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 105–111, 1999.

[35] D. K. Molzahn, B. C. Lesieutre, and C. L. DeMarco, "A sufficient condition for global optimality of solutions to the optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 29, pp. 978–979, 03 2014.

[36] A. Barzegar, D. K. Molzahn, and R. Su, "A method for quickly bounding the optimal objective value of an OPF problem using a semidefinite relaxation and a local solution," *Electric Power Systems Research*, vol. 177, p. 105954, 12 2019.

[37] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[38] J. Nie, "Optimality conditions and finite convergence of Lasserre's hierarchy," *Mathematical programming*, vol. 146, no. 1-2, pp. 97–121, 2014.

[39] C. Josz, J. Maeght, P. Panciatici, and J. C. Gilbert, "Application of the moment-SOS approach to global optimization of the OPF problem," *IEEE Transactions on Power Systems*, vol. 30, pp. 463–470, 01 2015.

[40] D. K. Molzahn and I. A. Hiskens, "Moment-based relaxation of the optimal power flow problem," in *2014 Power Systems Computation Conference*, pp. 1–7, 08 2014.

[41] D. K. Molzahn and I. A. Hiskens, "Sparsity-exploiting moment-based relaxations of the optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 30, pp. 3168–3180, 11 2015.

[42] B. Ghaddar, J. Marecek, and M. Mevissen, "Optimal power flow as a polynomial optimization problem," *IEEE Transactions on Power Systems*, vol. 31, pp. 539–546, 2016.

[43] X. Kuang, B. Ghaddar, J. Naoum-Sawaya, and L. F. Zuluaga, "Alternative sdp and socp approximations for polynomial optimization," *EURO Journal on Computational Optimization*, vol. 7, no. 2, pp. 153–175, 2019.

[44] N. Z. Shor, "Quadratic optimization problems," *Soviet Journal of Computer and Systems Sciences*, vol. 25, p. 1–11, 1987.

[45] X. Bai, H. Weihua, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, pp. 383–392, 07 2008.

[46] E. Salgado, A. Scozzari, F. Tardella, and L. Liberti, "Alternating current optimal power flow with generator selection," in *International Symposium on Combinatorial Optimization*, pp. 364–375, Springer, 2018.

[47] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz, "Positive definite completions of partial Hermitian matrices," *Linear algebra and its applications*, vol. 58, pp. 109–124, 1984.

[48] R. Jabr, "Exploiting sparsity in SDP relaxations of the OPF problem," *IEEE Transactions on Power Systems*, vol. 27, pp. 1138–1139, 05 2012.

[49] R. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, pp. 1458 – 1459, 09 2006.

[50] J. Prussing, "The principal minor test for semidefinite matrices," *Journal of Guidance Control and Dynamics*, vol. 9, pp. 121–122, 01 1986.

[51] H. Hijazi, C. Coffrin, and P. Van Hentenryck, "Convex quadratic relaxations of mixed-integer nonlinear programs in power systems," 09 2013. http://www.optimization-online.org/DB_HTML/2013/09/4057.html.

[52] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems," *Mathematical Programming*, vol. 10, pp. 147–175, 1976.

[53] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "The QC relaxation: A theoretical and computational study on optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, pp. 3008–3018, 02 2015.

[54] C. Coffrin, H. Hijazi, and P. Van Hentenryck, "Network flow and copper plate relaxations for ac transmission systems," in *2016 Power Systems Computation Conference (PSCC)*, pp. 1–8, IEEE, 2016.

[55] R. K. Ahyja, J. B. Orlin, and T. L. Magnanti, *Network flows: Theory, algorithms, and applications*. Prentice-Hall, 1993.

[56] B. Kocuk, S. Dey, and X. Sun, "Strong SOCP relaxations for the optimal power flow problem," *Operations Research*, vol. 64, pp. 1177–1196, 05 2016.

[57] G. Hall, "Optimization over nonnegative and convex polynomials with and without semidefinite programming," 06 2018.

[58] S. A. Gershgorin, "Uber die abgrenzung der eigenwerte einer matrix," no. 6, pp. 749–754, 1931.

[59] G. Barker and D. Carlson, "Cones of diagonally dominant matrices," *Pacific Journal of Mathematics*, vol. 57, no. 1, pp. 15–32, 1975.

[60] K. G. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Mathematical Programming*, vol. 39, pp. 117–129, 1987.

[61] D. V. Hilbert, "Ueber die darstellung definiter formen als summe von formenquadraten," *Mathematische Annalen*, vol. 32, pp. 342–350.

[62] B. Ghaddar, J. C. Vera, and M. F. Anjos, "A dynamic inequality generation scheme for polynomial programming," *Mathematical Programming*, vol. 156, no. 1-2, pp. 21–57, 2016.

[63] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, 2000.

[64] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, vol. 96, pp. 293–320, 05 2003.

[65] J. Pena, J. Vera, and L. Zuluaga, "Computing the stability number of a graph via linear and semidefinite programming," *SIAM Journal on Optimization*, vol. 18, pp. 87–105, 01 2007.

[66] J.-B. Lasserre, "Semidefinite programming vs. LP relaxations for polynomial programming," *Mathematics of operations research*, vol. 27, pp. 347–360, 05 2002.

[67] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, pp. 12–19, 02 2011.

[68] A. Megretski, "Spotless package," 2010. https://github.com/ spot-toolbox/spotless.

[69] MOSEK ApS, "The MOSEK optimization software," 2010. http://www.mosek.com.

[70] S. Sojoudi and J. Lavaei, "Physics of power networks makes hard optimization problems easy to solve," in *2012 IEEE Power and Energy Society General Meeting*, pp. 1–8, 07 2012.

[71] X. Bai and H. Weihua, "A semidefinite programming method with graph partitioning technique for optimal power flow problems," *International Journal of Electrical Power Energy Systems*, vol. 33, pp. 1309–1314, 09 2011.

[72] M. Fukuda, M. Kojima, K. Murota, and K. Nakata, "Exploiting sparsity in semidefinite programming via matrix completion I: General framework," *SIAM Journal on Optimization*, vol. 11, pp. 647–674, 02 1970.

[73] R. E. Tarjan and M. Yannakakis, "Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs," *SIAM Journal on computing*, vol. 13, no. 3, pp. 566–579, 1984.

[74] P. R. Amestoy, T. A. Davis, and I. S. Duff, "Algorithm 837: Amd, an approximate minimum degree ordering algorithm," *ACM Transactions on Mathematical Software (TOMS)*, vol. 30, no. 3, pp. 381–388, 2004.

[75] J. Miller, Y. Zheng, M. Sznaier, and A. Papachristodoulou, "Decomposed structured subsets for semidefinite and sum-of-squares optimization," *arXiv preprint arXiv:1911.12859*, 2019.

[76] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn, "Chordal decomposition in operator-splitting methods for sparse semidefinite programs," *Mathematical Programming*, vol. 180, no. 1, pp. 489–532, 2020.

[77] MOSEK ApS, "Modeling cookbook," http://docs.mosek.com/MOSEKModelingCookbook-letter.pdf.

[78] C. Coffrin, D. Gordon, and P. Scott, "NESTA, The NICTA energy system test case archive.," *CoRR*, vol. abs/1411.0359, 2014, [Online]. Available: https://arxiv.org/abs/1411.0359.

[79] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, "Sparse sum-of-squares (SOS) optimization: A bridge between DSOS/SDSOS and SOS optimization for sparse polynomials," in *2019 American Control Conference (ACC)*, pp. 5513–5518, IEEE, 2019.