Conic Relaxations of Power System Optimization: Theory and Algorithms

Fariba Zohrizadeh[b], Cedric Josz[a], Ming Jin[a], Ramtin Madani[b], Javad Lavaei[a], Somayeh Sojoudi[a]

[a] *University of California, Berkeley*
[b] *University of Texas at Arlington*

Abstract

Conic optimization has recently emerged as a powerful tool for designing tractable and guaranteed algorithms for power system operation. On the one hand, tractability is crucial due to the large size of modern electricity transmission grids. This is a result of the numerous interconnections that have been build over time. On the other hand, guarantees are needed to ensure reliability and safety for consumers at a time when power systems are growing in complexity. This is in large part due to the high penetration of renewable energy sources and the advent of electric vehicles. The aim of this paper is to review the latest literature in order to demonstrate the success of conic optimization when applied to power systems. The main focus is on how linear programming, second-order conic programming, and semidefinite programming can be used to address a central problem named the optimal power flow problem. We describe how they are used to design convex relaxations of this highly challenging non-convex optimization problem. We also show how sum-of-squares can be used to strengthen these relaxations. Finally, we present advances in first-order methods, interior-point methods, and nonconvex methods for solving conic optimization. Challenges for future research are also discussed.

Keywords: Optimal power flow, semidefinite programming, polynomial optimization, graph theory.

1. Introduction

The last several decades gave rise to formidable developments in operations research for handling nonconvexity in optimization. In the late eighties, it was recognized that one may remove all the nonconvexities in quadratically-constrained quadratic optimization so as to obtain a convex relaxation [178]. A few years later, it was observed that integer programs can likewise be relaxed to conic (and convex) optimization problems. In fact, there is a sequence of conic problems that ultimately yields the integral solutions [125, 177]. At the turn of the century, it was discovered that, hidden behind these contributions, there was a deep connection to measure theory and algebraic geometry [105, 106]. This led to the development of systematic approaches for solving polynomial optimization problems to global optimality based on sum-of-squares [166, 167]. Polynomial optimization is very general and englobes many problems arising in operations research: integer programming, linear programming, mixed integer programming, and quadratic optimization, to name a few. What the developments of the last several decades showed was that conic optimization,

Email addresses: fariba.zohrizadeh@uta.edu (Fariba Zohrizadeh), cedric.josz@gmail.com (Cedric Josz), jinming@berkeley.edu (Ming Jin), ramtin.madani@uta.edu (Ramtin Madani), lavaei@berkeley.edu (Javad Lavaei), sojoudi@berkeley.edu (Somayeh Sojoudi)

and in particular semidefinite programming, is a central tool in addressing nonconvexity. Fortunately, there exists a wide array of algorithms for solving these problems to arbitrary accuracy, some of which are highly efficient and even commercialized [5].

The development of algorithms for conic optimization dates back to the dawn of operations research with the introduction of the simplex algorithm. While this particular algorithm is well-suited for linear programming, it does not generalize to other conic programs such as semidefinite programming. Interior-point methods for linear programming [90], however, can be applied in the more general setting [16, 158]. One significant hurdle that needed to be overcome was dealing with complementarity constraints when considering the optimality conditions of a semidefinite program. This led to several numerical implementations still used today (e.g. SeDuMi [185], SDPT3 [193]). These can tackle dense problems with up to several hundred variables reliably, and can tackle larger instances by exploiting problem structure. This includes symmetries and sparsity. While they are not as mature as linear solvers, there has been great progress in recent years to develop alternative approaches, named first-order methods. They are inspired by works on differential equations in the fifties [50] and rely on operator theory.

Parallel to these advances in operations research, the largest engineering system devised by mankind came into being with the advent of the modern electric grid. With it came highly challenging problems in operation and planning for transmission system operators. This includes the alternating current optimal power flow (OPF) problem. OPF is concerned with the optimization of the steady state operating point of transmission and distribution networks in order to deliver electricity from suppliers to consumers as efficiently as possible. This fundamental problem was mathematically formulated in 1960s for the first time [34], and since then, due to the economic importance of OPF, substantial research efforts have been devoted to finding reliable solution methods to this problem [31, 126, 127].

The classical OPF formulation optimizes nodal complex voltages across a given network in terms of certain criterion (e.g., generation cost) subject to quadratic equations accounting for Kirchhoff's law, conservation of power, line thermal limits, and voltage stability limits, among other constraints. The reader is referred to [12, 31, 32] for alternative formulations of OPF. Due to the inherent complexity of physical laws that model the flow of electricity, some of these constraints are non-convex, which makes the OPF problem NP-hard in general [110, 113, 19]. One variant of this crucial problem is the security-constrained OPF where one has to account for contingencies on network components as well. In addition to minimizing the cost, security-constrained OPF concurrently plans for recourse strategies so that in case of an outage, in-service generators can compensate by adjusting their outputs. In some other variants of this problem (namely, security constrained unit commitment), unused generators are incorporated to accommodate changes in demand and to mitigate contingencies. Several other extensions of OPF have been studied under more general settings, to address considerations such as the security of operation [128, 33, 51, 8], robustness [49, 123, 209], energy storage [138, 206, 117], distributed platforms [104, 47, 70, 71, 204], uncertainty of generation [46, 200, 198], hybrid AC-DC grids [7, 199], real-time operation [191, 138], voltage stability [45, 202], and unit commitment [57, 129, 120, 35], among others [119]. System operators solve different variants of OPF on a daily basis in order to manage power system operations for which more sophisticated algorithms could lead to far more efficient and reliable decisions. Reliability is a core issue; indeed, interruptions in power incur nearly $79 billion/year in the United States alone according to the Energy Information administration [103].

The complexity of OPF is further pronounced when system operators need to determine additional binary decisions such as commitment of generating units and the on/off status of transmission lines. For instance, quantitative studies on real-world data suggest that decremental changes via switching transmission lines can simultaneously increase the operational security and reduce the cost of producing power [74, 58].

However, finding optimal switching strategies is known to be NP-hard in general [73]. From a computational perspective, exiting algorithms for determining these binary decisions rely on bounds from convex relaxations for pruning search trees and certifying closeness to global optimality. As a result, better convex relaxations and numerical methods for OPF can substantially improve the performance of search algorithms for solving mixed-integer power system problems.

Conventional methods for solving OPF include, linear approximations, local search algorithms, particle swarm optimization, fuzzy logic (see [164, 152, 153] and the references therein). A large body of literature has investigated linear programming-based relaxations and approximations of OPF [4, 179, 43, 176, 36, 52, 184, 143]. Local search algorithms for OPF have been studied in [208, 27, 141, 142, 211, 203]. However, the aforementioned methods do not guarantee the recovery of globally optimal or even feasible points [37]. In order to find globally optimal solutions, a variety of convex relaxations are studied in the literature for OPF [110, 126, 127, 213, 112, 102, 40, 18, 30, 41, 189, 20, 149]. Current industry practice for solving OPF involves non-linear optimization methods [38] and their variants; however, the need for modern approaches to address future grid challenges has been recognized by experts in the electric power industry as well as academics [156]. With the ever growing use of renewable sources, power grids are currently undergoing a revolution. New tools are needed to tackle the ever increasing complexity of system operations. By 2040, the penetration of solar power worldwide is expected to be at 29%; electrical vehicles are expected to comprise 33% of the global fleet; and utility-scale battery storage is predicted to account for 7% of the total power generation [21].

To remedy the aforementioned complexities, some of the most promising computational tools for solving OPF are developed based on conic optimization. This approach is proven to be effective in finding globally optimal points for a variety of power system optimization problems [110]. The benefits and drawbacks of conic optimization for OPF have been extensively investigated in the literature [78, 54, 55, 56, 10, 28, 114, 148], leading to various improvements in both solution quality [213, 94, 42, 39, 124, 134, 63, 22, 86, 66, 94, 42, 121, 95, 96] and computational cost [9, 80, 104, 150, 24, 6, 128, 182, 180, 134, 82]. This success story is the object of this review paper. It is an exciting time that some of the techniques developed in the operations research community could be applied to a real-world large-scale industrial problem. This has led the US Department of Energy to initiate the first-ever grid optimization competition with a $4 million cash prize whose goal is to increase power grid flexibility, reliability, safety, security, and efficiency[1].

Conic relaxation is proven to produce a globally optimal point for many special cases of OPF [64, 115, 132, 133, 91, 220, 217, 63, 22, 77, 161, 190]. However, in general, any convex relaxation may fail to produce physically meaningful solutions for OPF, due to NP-hardness. As a result, a variety of strategies are proposed in the literature to infer feasible and near-globally optimal points from inexact conic relaxations. For instance, branch-and-bound algorithms [68, 169, 39] can be used to iteratively partition search spaces in order to tighten conic relaxations. In [134, 128, 155], penalty terms are incorporated into the objective of convex relaxations in order ensure the feasibility of solutions produced by conic relaxations for OPF. In [146, 86, 147], moment-based hierarchies are used to form conic relaxations that result in globally optimal solutions for OPF. More recently, sequential and bound-tightening convex optimization method are proposed in [205] with the aim of recovering OPF feasible points [97, 205, 13, 222]. To overcome the high computational cost of solving large-scale conic optimization problems, some studies propose more efficient variants such as second-order cone programming (SOCP) [95, 129] and quadratic programming (QP) [41, 138]. Some papers have leveraged the sparsity of power networks to decompose large-scale conic constraints into lower-order ones [150, 6, 24, 128, 69].

[1] https://gocompetition.energy.gov

This paper is organized as follows. First, we formulate the optimal power flow problem and present recently proposed convex relaxations. They are significant due to the theoretical guarantees they come equipped with. They are also amenable to computations on large-scale power grids. Second, we discuss the application of sum-of-squares to polynomial optimization. We demonstrate its applicability on power systems by reviewing several recent publications. Finally, we review the latest progress on algorithms for solving conic optimization, which has seen a lot of activity in recent years.

1.1. Notations

The symbols $\mathbb{R}$, $\mathbb{C}$, $\mathbb{N}$, $\mathbb{S}^n$, $\mathbb{S}^n_+$, and $\mathbb{H}^n$ denote the sets of real numbers, complex numbers, nonnegative integer numbers, $n \times n$ real symmetric matrices, $n \times n$ symmetric positive-definite matrices, and $n \times n$ Hermitian matrices, respectively. The notations $\text{rank}\{\cdot\}$, $\text{trace}\{\cdot\}$, and $\|\cdot\|_F$ denote the rank, trace, and Frobenius norm of a matrix. The notation $X \succeq 0$ means that X is positive semidefinite. The vectorization of a matrix X is the column-stacking operation

$$\text{vec}\{X\} = [X_{1,1}, \ldots, X_{n,1}, X_{1,2}, \ldots, X_{n,2}, \ldots, X_{n,n}]^T,$$

The cardinality of a set $\mathcal{D}$ is shown as $|\mathcal{D}|$. The absolute value of a complex number z is denoted as $|z|$. Given two vectors $x, y \in \mathbb{C}^n$, the notation $|x| \leq y$ means that the absolute value of each element of x is smaller than or equal to the corresponding element of y. Similarly, x^2 denotes a vector obtained by taking the square of the entries of x element-wise. The symbol $\text{diag}\{X\}$ represents a column vector extracting the diagonal elements of a matrix X. The inner product of two matrices X and Y is denoted as $X \bullet Y$. The notation $\measuredangle x$ shows the vector of phase angles of the elements of $x \in \mathbb{C}^n$. The transpose and conjugate transpose of a complex-valued matrix X are shown as $X^\top$ and X^*, respectively. The real and imaginary parts of a complex number z are denoted as $\text{real}\{z\}$ and $\text{imag}\{z\}$. The conjugate of z is shown as $\overline{z}$. The imaginary unit is represented by i.

2. Convex relaxations of optimal power flow

In recent years, there has been a formidable progress on designing various kinds of convex relaxations. These are based on conic optimization such as semidefinite programming, second-order conic programming, and linear programming. They provide fast and guaranteed approaches for computing bounds on the global value of the nonconvex optimization problem at hand. We next discuss recent developments in this area. As some of these relaxations make use of the specific problem structure, we begin by describing the physics of the problem in detail.

2.1. Optimal power flow formulation

A power network can be represented as a graph $\mathcal{H} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ and $\mathcal{L}$ denote the set of buses and lines of the network, respectively. For each bus $k \in \mathcal{N}$, let $v_k \in \mathbb{C}$ denote the nodal complex voltage, whose magnitude $|v_k|$ and angle $\measuredangle v_k$ account for the amplitude and phase of the voltage, respectively. The complex nodal demand at bus k is denoted by $d_k \in \mathbb{C}$, whose real and imaginary parts, respectively, account for active and reactive demand. Let $\mathcal{G}$ be the set of generating units, each located at one of the buses. For each generating unit $g \in \mathcal{G}$, the values p_g and q_g, respectively, denote the amount of active and reactive powers. The unit incidence matrix $\boldsymbol{C} \in \{0,1\}^{|\mathcal{G}| \times |\mathcal{V}|}$ is defined as a binary matrix whose (g,k) entry is equal to one if and only if the generating unit g belongs to bus k. Additionally, define the pair of matrices $\vec{\boldsymbol{C}}, \overleftarrow{\boldsymbol{C}} \in \{0,1\}^{|\mathcal{E}| \times |\mathcal{V}|}$ to be the *from* and *to* incidence matrices, respectively. The (l,k) entry of $\vec{\boldsymbol{C}}$ is equal to one if and only if the line $l \in \mathcal{E}$ starts at bus k, while the (l,k) entry of $\overleftarrow{\boldsymbol{C}}$ is equal to one if and

only if line l ends at bus k. Define $\boldsymbol{Y} \in \mathbb{C}^{|\mathcal{N}|\times|\mathcal{N}|}$ as the nodal admittance matrix of the power network and $\vec{\boldsymbol{Y}}, \overleftarrow{\boldsymbol{Y}} \in \mathbb{C}^{|\mathcal{L}|\times|\mathcal{N}|}$ as the *from* and *to* branch admittance matrices, respectively. Using the notations introduced above, the OPF problem can be formulated as

$$
\begin{aligned}
\underset{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{R}^{|\mathcal{G}|}\\ \boldsymbol{v}\in\mathbb{C}^{|\mathcal{N}|}}}{\text{minimize}} \quad & h(\boldsymbol{p}) && \text{(1a)}\\
\text{subject to} \quad & \boldsymbol{d} + \operatorname{diag}\{\boldsymbol{v}\,\boldsymbol{v}^*\boldsymbol{Y}^*\} = \boldsymbol{C}^\top(\boldsymbol{p} + \mathrm{i}\boldsymbol{q}), && \text{(1b)}\\
& |\operatorname{diag}\{\vec{\boldsymbol{C}}\,\boldsymbol{v}\,\boldsymbol{v}^*\vec{\boldsymbol{Y}}^*\}| \le \boldsymbol{f}^{\max}, && \text{(1c)}\\
& |\operatorname{diag}\{\overleftarrow{\boldsymbol{C}}\,\boldsymbol{v}\,\boldsymbol{v}^*\overleftarrow{\boldsymbol{Y}}^*\}| \le \boldsymbol{f}^{\max}, && \text{(1d)}\\
& \boldsymbol{v}^{\min} \le |\boldsymbol{v}| \le \boldsymbol{v}^{\max}, && \text{(1e)}\\
& \boldsymbol{p}^{\min} \le \ \boldsymbol{p}\ \le \boldsymbol{p}^{\max}, && \text{(1f)}\\
& \boldsymbol{q}^{\min} \le \ \boldsymbol{q}\ \le \boldsymbol{q}^{\max}, && \text{(1g)}\\
& \measuredangle\operatorname{diag}\{\vec{\boldsymbol{C}}\boldsymbol{v}\boldsymbol{v}^*\overleftarrow{\boldsymbol{C}}\} \le \boldsymbol{\theta}^{\max}, && \text{(1h)}
\end{aligned}
$$

in terms of decision variables $\boldsymbol{p} \triangleq [p_g]$, $\boldsymbol{q} \triangleq [q_g]$ and $\boldsymbol{v} \triangleq [v_k]$. In the above formulation, objective function $h(\boldsymbol{p})$ is a desirable cost function, e.g. power transmission loss or generation loss. Constraint (1b) is the power balance equation, which accounts for conservation of energy at all buses of the power network. Imposing this constraint ensures that the voltages across the network are adjusted such that the overall complex power produced at each bus $k \in \mathcal{N}$ be equal to the power consumption and power exchanges of that bus. In order for the transmission lines to operate properly, the flows of power entering the lines of the network from their starting and ending buses are upper bounded by the vector of thermal/stability limits $\boldsymbol{f}_{\max} \in \mathbb{R}^{|\mathcal{L}|}$ using constraints (1c) and (1d). Similarly, the voltage magnitudes of all buses as well as the active and reactive power of the generating units are restricted by constraints (1e)–(1g), given the limits $\boldsymbol{v}_{\min}, \boldsymbol{v}_{\max} \in \mathbb{R}^{|\mathcal{N}|}$, $\boldsymbol{p}_{\min}, \boldsymbol{p}_{\max} \in \mathbb{R}^{|\mathcal{G}|}$, and $\boldsymbol{q}_{\min}, \boldsymbol{q}_{\max} \in \mathbb{R}^{|\mathcal{G}|}$. In addition, constraint (1h) is imposed to bound the phase angle difference across the lines of the network by $\boldsymbol{\theta}_{\max}$. Due to practical considerations (such as stability limits), we make the assumption that all elements of $\boldsymbol{\theta}_{\max}$ are restricted to the interval $[0^\circ, 90^\circ]$.

2.2. Semidefinite programming relaxation

The OPF problem (1a)–(1h) includes nonconvex nonlinear polynomial equations, which render the optimization problem non-convex and NP-hard [110, 19, 113]. Therefore, determining a global or near-global solution of this problem in polynomial time is a daunting computational challenge. To tackle the non-convexity of the OPF problem and reduce its computational complexity, a huge body of research has focused on convex relaxations. The relaxations provide lower bounds on the unknown globally optimal cost of the OPF problem and can assure whether a solution to the OPF problem is globally optimal. A relaxation is said to be *exact* if the globally optimal solution of the non-convex OPF problem can be recovered by the optimal solution of the relaxation. The inherent relation between the quadratic function of complex bus voltages and the complex power imposes quadratic constraint (1b) and makes the OPF problem (1a)–(1h) NP-hard. Nevertheless, all the non-convexity induced by the quadratic terms disappears if the problem is reformulated in terms of auxiliary variable $\boldsymbol{W} \in \mathbb{H}^{|\mathcal{N}|}$, accounting for $\boldsymbol{v}\boldsymbol{v}^*$ [53, 10, 79, 110]. Using $\boldsymbol{W}$, variable $\boldsymbol{v}$ can be dropped from the optimization problem by equivalently replacing the consistency constraint $\boldsymbol{W} = \boldsymbol{v}\boldsymbol{v}^*$ with two constraints (i) $\boldsymbol{W} \succeq 0$ and (ii) $\operatorname{rank}\{\boldsymbol{W}\} = 1$. Given that, problem

(1a) – (1h) can be equivalently reformulated as:

$$\begin{aligned}
\underset{\substack{\boldsymbol{p},\boldsymbol{q}\in\mathbb{R}^{|\mathcal{G}|}\\ \boldsymbol{W}\in\mathbb{H}^{|\mathcal{N}|}}}{\text{minimize}} \quad & h(\boldsymbol{p}) && \text{(2a)}\\
\text{subject to} \quad & \boldsymbol{d}+\operatorname{diag}\{\boldsymbol{W}\boldsymbol{Y}^*\}=\boldsymbol{C}^\top(\boldsymbol{p}+\mathrm{i}\boldsymbol{q}), && \text{(2b)}\\
& |\operatorname{diag}\{\vec{\boldsymbol{C}}\boldsymbol{W}\vec{\boldsymbol{Y}}^*\}|\le \boldsymbol{f}^{\max}, && \text{(2c)}\\
& |\operatorname{diag}\{\overleftarrow{\boldsymbol{C}}\boldsymbol{W}\overleftarrow{\boldsymbol{Y}}^*\}|\le \boldsymbol{f}^{\max}, && \text{(2d)}\\
& (\boldsymbol{v}^{\min})^2\le \operatorname{diag}\{\boldsymbol{W}\}\le(\boldsymbol{v}^{\max})^2 && \text{(2e)}\\
& \boldsymbol{p}^{\min}\le\ \boldsymbol{p}\ \le\boldsymbol{p}^{\max}, && \text{(2f)}\\
& \boldsymbol{q}^{\min}\le\ \boldsymbol{q}\ \le\boldsymbol{q}^{\max}, && \text{(2g)}\\
& |\operatorname{imag}\{W_{lm}\}|\le \operatorname{real}\{W_{lm}\}\tan(\theta_{lm}^{\max}), \qquad \forall(l,m)\in\mathcal{L} && \text{(2h)}\\
& \boldsymbol{W}\in\mathcal{C}_{\text{OPF}}, && \text{(2i)}
\end{aligned}$$

where $\mathcal{C}_{\text{OPF}}$ is

$$\mathcal{C}_{\text{OPF}}\triangleq\big\{\boldsymbol{H}\in\mathbb{H}^{|\mathcal{N}|}\ \big|\ \boldsymbol{H}\succeq 0,\ \operatorname{rank}\{\boldsymbol{H}\}=1\big\}.$$

Observe that in the above reformulation, constraint (2i) carries all the nonconvexity and neglecting this constraint yields a convex problem which can be solved in polynomial time. Motivated by this observation, a series of work proposed to convexify and efficiently solve the OPF problem through convex relaxation techniques including conic relaxations [81, 10, 110, 181, 42, 128, 95, 129, 222], quadratic programming (QP) [41, 222], and Mixed-Integer programming (MIP) [18] relaxations of the OPF problem. The relaxations enlarge the nonconvex feasible set $\mathcal{C}_{\text{OPF}}$ into a proper convex set, and then provide a lower bound on the globally optimal cost associated with the global solution. One of the most promising directions in solving the OPF problem is based on conic relaxation of the power flow equations. This approach to the OPF problem has attracted much attention due to its proven ability in obtaining tight lower bounds for several practical case studies [108, 109, 110]. In what follows, we will discuss the the commonly-used conic relaxation techniques for the optimal power flow problem. The SDP relaxation of the non-convex problem (1a) – (1h) can be obtained by replacing $\mathcal{C}_{\text{OPF}}$ with the cone of $|\mathcal{N}|\times|\mathcal{N}|$ Hermitian positive semidefinite matrices $\mathcal{C}_{\text{SDP}}$ defined as:

$$\mathcal{C}_{\text{SDP}}\triangleq\big\{\boldsymbol{H}\in\mathbb{H}^{|\mathcal{N}|}\ \big|\ \boldsymbol{H}\succeq 0\big\}. \tag{3}$$

The SDP relaxation is exact and finds globally optimal solution of the OPF problem (1a) – (1h) if and only if it possesses a solution $(\boldsymbol{p}^{\text{OPT}},\boldsymbol{q}^{\text{OPT}},\boldsymbol{W}^{\text{OPT}})$ such that $\operatorname{rank}\{\boldsymbol{W}^{\text{OPT}}\}=1$. It has been proven in [110] that the SDP relaxation is exact for a purely resistive network with no reactive loads. This condition holds for several test cases of OPF including IEEE benchmark systems with 14, 30, 57, 118 and 300 buses after a small perturbation in a few entries of the admittance matrix $\boldsymbol{Y}$. The work [181] showed that the success of the SDP relaxation is due to the passivity of the power infrastructure, which was formalized in [182] through the notion of sign-definite weight sets. The papers [111, 212] proved that the SDP relaxation is guaranteed to be exact if the phase angle difference across every line of the network is sufficiently close (e.g. $\boldsymbol{\theta}^{\max}\in[0,90^\circ]$). Regardless of the above-mentioned sufficient conditions, it has been shown that the SDP relaxation may not always be exact for some instances of the OPF problem [114, 28, 148]. In such cases, the solution of SDP relaxation offers a lower bound on the unknown globally optimal cost of the OPF problem. The paper [134] observed that exactness of the SDP relaxation heavily depends on the formulation of the line capacity constraint. Particularly, it has been shown that for a practical circuit, the feasible region

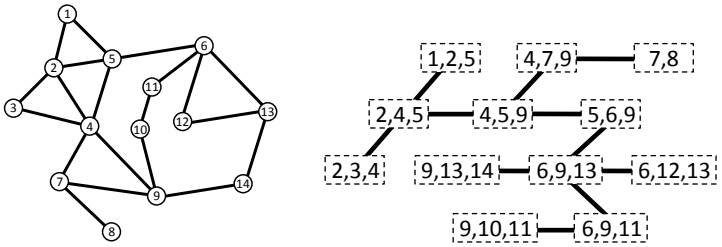

Figure 1: The IEEE 14-bus power network (left figure) and its minimal tree decomposition (right figure).

associated with constraint (1h) and $|v_l - v_m| \leq \Delta_{lm}^{\max}$ are very similar in the non-convex space, however, their relaxations in terms of auxiliary variable $\boldsymbol{W}$, i.e. constraint (2h) and $W_{ll} + W_{mm} - W_{lm} - W_{ml} \leq (\Delta_{lm}^{\max})^2$, no longer preserve this similarity and the latter constraint provides a tighter feasible set in the lifted space. Motivated by this observation [134] suggests to impose the latter constraint to strengthen the SDP relaxation and reduce its optimality gap. The papers [42] proposed to amend the feasible set of the relaxed problem via the following valid inequalities for all $(k_1, k_2) \in \mathcal{L}$:

$$v_{k_1}^{\sigma} v_{k_2}^{\sigma} \mathrm{real}\{W_{k_1k_2} e^{-i\theta_{k_1k_2}^{\max}}\} - v_{k_2}^{\max} v_{k_2}^{\sigma} W_{k_1k_1} - v_{k_1}^{\max} v_{k_1}^{\sigma} W_{k_2k_2} \geq v_{k_1}^{\max} v_{k_2}^{\max} (v_{k_1}^{\min} v_{k_2}^{\min} - v_{k_1}^{\max} v_{k_2}^{\max}) \quad \text{(4a)}$$

$$v_{k_1}^{\sigma} v_{k_2}^{\sigma} \mathrm{real}\{W_{k_1k_2} e^{-i\theta_{k_1k_2}^{\max}}\} - v_{k_2}^{\min} v_{k_2}^{\sigma} W_{k_1k_1} - v_{k_1}^{\min} v_{k_1}^{\sigma} W_{k_2k_2} \geq v_{k_1}^{\min} v_{k_2}^{\min} (v_{k_1}^{\max} v_{k_2}^{\max} - v_{k_1}^{\min} v_{k_2}^{\min}) \quad \text{(4b)}$$

where $v_k^{\sigma} = v_k^{\max} + v_k^{\min}$, for $k = \{1, \ldots, |\mathcal{N}|\}$. The papers [146, 86, 147] propose to enhance the tightness of the relaxation by employing Lasserre hierarchy of moment relaxation [107]. The first-order moment relaxation of OPF problem in the hierarchy is equivalent to the SDP relaxation (2a) – (2i) and (3). It has been shown in [146, 86, 147] that increasing the relaxation order of Lasserre hierarchy can strengthen the approximation and approaches to the global solutions of the OPF problem. Nevertheless, this comes at the cost of increasing computational complexity. The papers [68, 169] employ a Branch-and-bound algorithm [118] to iteratively partition the feasible set of the relaxation to find a solution with a smaller gap.

In addition to the exactness issues, SDP relaxation suffers from a high computational cost, which limits its applicability especially for large-scale real-world networks. This motivates researchers to alleviate the complexity of solving SDP relaxations by finding computationally cheaper alternatives such as reduced-SDP relaxation and second-order cone programming (SOCP) relaxation.

2.3. Reduced semidefinite programming relaxation

The computational complexity of SDP relaxation motivates various studies to leverage the sparsity of power networks through a graph-theoretic analysis, namely tree decomposition [9, 80, 150, 6, 24, 128, 69, 216]. An arbitrary tree decomposition of the power network, denotes as $\mathcal{B}$, decomposes $\mathcal{N}$ into several overlapping subsets $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_B \subseteq \mathcal{N}$ each serving as a node of a tree where: (i) each node $\mathcal{B}_k$ is a set containing a bag of vertices of $\mathcal{N}$, (ii) each vertex in $\mathcal{N}$ is contained in at least one bag of $\mathcal{B}$, (iii) each pair of connected vertices in $\mathcal{L}$ is contained in at least one bag, (iv) all bags containing each arbitrary vertex in $\mathcal{N}$ must be a sub-tree. The width of the tree is equal to the cardinality of the biggest $\{\mathcal{B}_k\}_{k=1}^{B}$ minus one. The tree-width of the network graph is the minimum width over all possible tree decompositions of $\mathcal{N}$. The graph of the IEEE 14-bus system and its minimal tree decomposition are depicted in Figure 1. Given $\mathcal{B}$ with B nodes, the constraint $\boldsymbol{W} \succeq 0$ can be represented with a set of small-sized conic constraints as:

$$\boldsymbol{W}\{\mathcal{B}_k\} \succeq 0, \ \ \forall k \in \{1, 2, ..., B\} \quad (5)$$

where $\boldsymbol{W}\{\mathcal{B}_k\}$ represents the $|\mathcal{B}_k| \times |\mathcal{B}_k|$ principal sub-matrix of $\boldsymbol{W}$ whose rows and columns are chosen from $\mathcal{B}_k$. The reduced-SDP relaxation can be obtained by substituting $\mathcal{C}_{\text{OPF}}$ with the cone:

$$\mathcal{C}_{\text{RSDP}} \triangleq \big\{\boldsymbol{H} \in \mathbb{H}^{|\mathcal{N}|} \;\big|\; \boldsymbol{H}\{\mathcal{B}_k\} \succeq 0, \forall k \in \{1, ..., B\}\big\}. \tag{6}$$

The reduced-SDP relaxation is exact if and only if it has a solution $(\boldsymbol{p}^{\text{OPT}}, \boldsymbol{q}^{\text{OPT}}, \boldsymbol{W}^{\text{OPT}})$ such that $\text{rank}\{\boldsymbol{W}^{\text{OPT}}\{\mathcal{B}_k\}\} = 1$ for $k = 1, \ldots, B$.

Theorem 1 (Theorem 1. [128])**.** *The optimal objective costs of the SDP relaxation and the reduced-SDP relaxation of the OPF problem are the same.*

Theorem 1 implies that the above decomposition provides an equivalent but far more tractable formulation of the SDP relaxation for large-scale power networks with a relatively small tree-width. Nevertheless, solving large-scale instances of the OPF problem on real-world systems can still be computationally challenging.

2.4. Second-order conic programming relaxation

The primary limitation of SDP-based relaxations is the rapid growth of problem dimension, which makes the problem computationally prohibitive to solve for large-scale power systems. A weaker, but computationally less demanding alternative to SDP relaxation is the second-order cone programming (SOCP) relaxation which was introduced in [81] for acyclic networks and [181] for general networks. The SOCP relaxation aims to enforce the relation between $\boldsymbol{W}$ and $\boldsymbol{v}$ through conic constraints on 2×2 principle sub-matrices of $\boldsymbol{W}$ corresponding to each line $(k_1, k_2) \in \mathcal{L}$:

$$\boldsymbol{W}\{(k_1, k_2)\} = \begin{bmatrix} W_{k_1k_1} & W_{k_1k_2} \\ W_{k_2k_1} & W_{k_2k_2} \end{bmatrix} \succeq 0, \quad \forall (k_1, k_2) \in \mathcal{L}. \tag{7}$$

which can be equivalently replaced by the following linear and a rotated second-order cone constraints:

$$\begin{aligned} &W_{k_1k_1}, W_{k_2k_2} \geq 0, && \forall (k_1, k_2) \in \mathcal{L}, \\ &W_{k_1k_1} W_{k_2k_2} \geq \left|W_{k_1k_2}\right|^2, && \forall (k_1, k_2) \in \mathcal{L}. \end{aligned} \tag{8}$$

Therefore, the SOCP relaxation of the power flow problem is obtained by optimizing problem (2a) – (2i) over the cone

$$\mathcal{C}_{\text{SOCP}} \triangleq \big\{\boldsymbol{H} \in \mathbb{H}^{|\mathcal{N}|} \;\big|\; \boldsymbol{H}\{(k_1, k_2)\} \succeq 0,\ \forall (k_1, k_2) \in \mathcal{L}\big\}. \tag{9}$$

It is evident that the SOCP relaxation is dominated by the SDP relaxation, since constraint $\boldsymbol{W} \succeq 0$ ensures that every sub-matrix of $\boldsymbol{W}$, including the 2×2 principle sub-matrices, is positive semidefinite [171]. The SOCP relaxation is exact if and only if it possesses a solution $(\boldsymbol{p}^{\text{OPT}}, \boldsymbol{q}^{\text{OPT}}, \boldsymbol{W}^{\text{OPT}})$ such that: (i) $\text{rank}\{\boldsymbol{W}^{\text{OPT}}\{(k_1, k_2)\}\} = 1$ for all $(k_1, k_2) \in \mathcal{L}$, (ii) for every directed cycle of the network, the sum of the phase of W_{lm}^{OPT} over all directed edges (l, m) of the cycle must be zero. It is theoretically proven that the SOCP relaxation is exact under certain technical assumptions on the physics of a power gird [181, 182, 111, 213]. The papers [23, 181] prove that the SOCP relaxation is exact for radial networks if load over-satisfaction is allowed, i.e.

$$\text{diag}\{\boldsymbol{W}\boldsymbol{Y}^*\} \leq \boldsymbol{C}^\top(\boldsymbol{p} + \mathrm{i}\boldsymbol{q}) - \boldsymbol{d}. \tag{10}$$

The papers [116, 61, 62, 126] guarantee the exactness of the SOCP relaxation for radial networks when there are no upper bounds on the voltage magnitudes of the buses of the network. Note that it has proven in [126, 187] that if the SOCP relaxation is exact for radial networks, then the SDP and reduced-SDP relaxations are also exact. Therefore, in the case of acyclic power network, SOCP relaxation is more appealing due to its tightness and computational tractability. Regardless of the above-mentioned restricted assumptions, the SOCP relaxation is often inexact, and its optimal solutions are not necessarily feasible for problem (1a)–(1h). To strengthen the SOCP relaxation various techniques have been suggested including valid inequalities [95], cutting planes [76, 129, 97], and convex envelope [41]. The paper [76] proposed to tighten the SOCP relaxation by imposing the conic constraints on the 3×3 principle sub-matrices of matrix variable $\boldsymbol{W}$ as:

$$\boldsymbol{W}\{(k_1,k_2,k_3)\}=\begin{bmatrix} W_{k_1k_1} & W_{k_1k_2} & W_{k_1k_3} \\ W_{k_2k_1} & W_{k_2k_2} & W_{k_2k_3} \\ W_{k_3k_1} & W_{k_3k_2} & W_{k_3k_3} \end{bmatrix} \succeq 0, \quad \forall (k_1,k_2,k_3) \in \bigcup_{k=1}^{B} \mathcal{B}_k \times \mathcal{B}_k \times \mathcal{B}_k. \tag{11}$$

where $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_B \subseteq \mathcal{N}$ are the overlapping subsets associated with an arbitrary tree decomposition of the network for which $|\mathcal{B}_k| \geq 3$, $\forall k = 1, \ldots, B$. The intuition behind imposing constraint (11) is that a nonzero, Hermitian matrix $\boldsymbol{W} \in \mathbb{H}^{|\mathcal{N}|}$ is positive semidefinite if and only if all the principal minors of $\boldsymbol{W}$ are non-negative. Based on this fact, constraint (11) is equivalent to constraint (7) together with a new constraint $\det\{\boldsymbol{W}\{(k_1,k_2,k_3)\}\} \geq 0$. The new constraints can be seen as cuts for strengthening the SOCP relaxation. The paper [97] proposed to strengthen SOCP relaxation by imposing constraints on both principal and non-principal 2×2 minors of the $\boldsymbol{W}$. These constraints are imposed based on the following characterization of rank-one positive semidefinite matrices.

Proposition 1. *A nonzero, Hermitian matrix $\boldsymbol{W} \in \mathbb{H}^{|\mathcal{N}|}$ satisfies constraints $\boldsymbol{W} \succeq 0$ and $\mathrm{rank}\{\boldsymbol{W}\} = 1$ if and only if all the 2×2 minors of $\boldsymbol{W}$ are equal to zero and $\boldsymbol{W}_{kk} \geq 0$ for $k = \{1, \ldots, |\mathcal{N}|\}$.*

It has been shown that constraints on non-principal 2×2 minors of $\boldsymbol{W}$ have physical interpretations of voltage angle differences summing to zero over cycles of three- and four-nodes in the power network.

2.5. Quadratic convex relaxation

The paper [41] proposed a Quadratic Convex (QC) relaxation that imposes quadratic constraints in terms of the variables $|v_k|^2$, $|v_{k_1}||v_{k_1}|\cos(\measuredangle v_{k_1} - \measuredangle v_{k_2})$, and $|v_{k_1}||v_{k_1}|\sin(\measuredangle v_{k_1} - \measuredangle v_{k_2})$ to enforce the relation between $\boldsymbol{W}$ and $\boldsymbol{v}$ implicitly. Under the assumption that phase angle differences between the neighboring buses in a power network are usually small, this paper computes convex envelopes of the polar representation of consistency constraint $\boldsymbol{W} = \boldsymbol{v}\boldsymbol{v}^*$ as:

$$\begin{aligned} &W_{k,k} = \langle |v_k|^2 \rangle^T, && \forall k \in \mathcal{N} \\ &\mathrm{real}\{W_{k_1k_2}\} = \langle \langle |v_{k_1}|\,|v_{k_2}| \rangle^M \langle \cos(\measuredangle v_{k_1} - \measuredangle v_{k_2}) \rangle^C \rangle^M && \forall (k_1,k_2) \in \mathcal{L} \\ &\mathrm{imag}\{W_{k_1k_2}\} = \langle \langle |v_{k_1}|\,|v_{k_2}| \rangle^M \langle \sin(\measuredangle v_{k_1} - \measuredangle v_{k_2}) \rangle^C \rangle^M && \forall (k_1,k_2) \in \mathcal{L} \end{aligned} \tag{12}$$

where $\langle f(.) \rangle^T$, $\langle f(.) \rangle^M$, $\langle f(.) \rangle^S$, and $\langle f(.) \rangle^C$ are the convex envelopes defined as:

$$\langle x^2 \rangle^T \triangleq \begin{cases} \check{x} \geq x^2, \\ \check{x} \leq (x^{\max} + x^{\min})x - x^{\max}x^{\min}, \end{cases} \quad \langle xy \rangle^M \triangleq \begin{cases} \check{xy} \geq x^{\min}y + y^{\min}x - x^{\min}y^{\min}, \\ \check{xy} \geq x^{\max}y + y^{\max}x - x^{\max}y^{\max}, \\ \check{xy} \leq x^{\min}y + y^{\max}x - x^{\min}y^{\max}, \\ \check{xy} \leq x^{\max}y + y^{\min}x - x^{\max}y^{\min}, \end{cases} \tag{13}$$

$$\langle \sin(x) \rangle^S \triangleq \begin{cases} \check{sx} \leq \cos(\frac{x^{\max}}{2})(x - \frac{x^u}{2}) + \sin(\frac{x^{\max}}{2}), \\ \check{sx} \geq \cos(\frac{x^{\max}}{2})(x + \frac{x^u}{2}) - \sin(\frac{x^{\max}}{2}), \end{cases} \quad \langle \cos(x) \rangle^C \triangleq \begin{cases} \check{cx} \leq 1 - \frac{1-\cos(x^{\max})}{(x^{\max})^2} x^2, \\ \check{cx} \geq \cos(x^{\max}), \end{cases} \tag{14}$$

where $x \in [x^{\min}, x^{\max}]$ and $y \in [y^{\min}, y^{\max}]$. These convex envelopes are further strengthened via second-order cone constraints (8). The paper [222] proposed a computationally efficient method, named the *parabolic relaxation*, which transforms non-convex problem (1a) – (1h) into a convex QCQP. Since parabolic relaxation avoids conic constraints, it requires far less computational effort and can serve as an alternative to the common practice SDP and SOCP relaxations for solving large-scale OPF problems.

2.6. Linear programming relaxation

The paper [18] develops a linear relaxation of the rectangular formulation of the OPF problem in which the non-linearity and non-convexity comes from the bi-linear equations in terms the real and imaginary parts of the complex voltages. Let us define each v_k at bus k as $v_k = e_k + \mathrm{i} f_k$ where e_k and f_k are bounded variables corresponding to the real and imaginary parts of the v_k. It has been shown that all the bi-linearity of the rectangular formulation can be removed by performing a binary expansion for suitably translated and scaled e_k and f_k. To gain more intuition, consider variables $u, v \in [0, 1]$. The variable u can be expressed as:

$$u = \sum_{j=1}^{T} 2^{-j} y_j + \delta, \tag{15}$$

where $y_j \in \{0, 1\}$, $T \geq 1$, and $0 \leq \delta \leq 2^{-T}$. Given (15), the bi-linear term uv can be relaxed as

$$\sum_{j=1}^{T} 2^{-j} w_j \leq uv \leq \sum_{j=1}^{T} 2^{-j} w_j + 2^{-T} v, \tag{16}$$

where w_j accounts for $y_j v$. To remedy the absence of non-convex equation $w_j = y_j v$ the relaxation is strengthened via the following valid inequalities:

$$\begin{aligned} & w_j \leq \min\{v, y_j\}, \\ & w_j \geq \max\{v + y_j - 1, 0\}. \end{aligned} \tag{17}$$

It has been shown that the binary expansion of the bi-linear terms results in a computationally more efficient relaxation of the OPF problem compared to the SDP and SOCP relaxations.

2.7. Heuristics and recovery of feasible points

Despite the effectiveness of the relaxation techniques in solving the OPF problem, their solutions may not lead to a physically meaningful solution for problem (1a) – (1h). One promising strategy to ameliorate this issue and implicitly enforce the rank constraint is to augment the relaxation with penalization terms [155, 134, 128, 146, 217, 222, 151, 146]. The paper [134] proposed to penalize the total reactive power generation of the network by replacing objective function (2a) with:

$$h(\boldsymbol{p}) + \varepsilon_b \sum_{g \in \mathcal{G}} q_g, \tag{18}$$

where $\varepsilon_b > 0$ is the penalty coefficient. The paper [128] proposed to penalize the apparent power loss over the series impedance of the lines of the network using the following objective function:

$$h(\boldsymbol{p}) + \varepsilon_b \sum_{g \in \mathcal{G}} q_g + \varepsilon_l \sum_{(l,m) \in \mathcal{L}} \Big| W_{ll} + W_{mm} - W_{lm} - W_{ml} \Big| |y^*_{lm}|, \tag{19}$$

where $\varepsilon_b, \varepsilon_l > 0$ are penalty coefficients and y^*_{lm} is the series impedance of line (l, m). It has been shown that employing the above-mentioned penalty terms elevates the off-diagonal entries of matrix $\boldsymbol{W}$, which indeed promotes low-rank solutions. Given $\boldsymbol{W}^{\text{OPT}}$ of the penalized SDP relaxation, an approximate feasible solution $\boldsymbol{v}$ of problem (1a) – (1g) is obtained by setting magnitude v_k equal to the square root of W^{OPT}_{kk}, for all $k \in \{1, \ldots, |\mathcal{N}|\}$ and finding the phases of the entries of $\boldsymbol{v}$ by minimizing the following optimization problem:

$$\underset{\measuredangle \boldsymbol{v} \in \mathbb{C}^{|\mathcal{N}|}}{\text{minimize}} \sum_{(l,m) \in \mathcal{L}} |\measuredangle W^{\text{OPT}}_{lm} - \measuredangle v_l - \measuredangle v_m| \tag{20a}$$

$$\text{subject to } |\measuredangle v_k| \leq \pi, \qquad \forall k \in \mathcal{N}, \tag{20b}$$

$$\measuredangle v_1 = 0. \tag{20c}$$

Note that if matrix $\boldsymbol{W}^{\text{OPT}}$ has a dominant nonzero eigenvalue, problem (20a) – (20c) seeks to find vector $\boldsymbol{v}$ such that its line angle differences are as close as possible to those suggested by $\boldsymbol{W}^{\text{OPT}}$. The paper [222, 221] proposed to incorporate linear penalty terms into the objective function (2a) where each penalty term is based on an arbitrary initial point of the relaxation. It has been proven in [92] that if the initial point is sufciently close to the feasible set of the non-convex OPF problem (1a) – (1h) and satifies a generalized linear independence constraint qualification (LICQ) condition, the penalized relaxations produce feasible solutions for the OPF problem.

3. Sum-of-squares for polynomial optimization

The motivation for using sum-of-squares for solving the optimal power flow problem is that the semidefinite programming Shor relaxation (or rank relaxation) is not strong enough to solve all problems of interest to global optimality. As was observed in [114], in the LMBM3 3-bus system (named after its authors Lesieutre, Molzahn, Borden, and Demarco) of Figure 2, the rank constraint is not always satisfied at optimality in the Shor relaxation of the optimal power flow as proposed in [110]. For instance, when the upper limit on the power flow on line 2-3 is equal to 48 MVA, the Shor relaxation provides a lower bound equal to 5819.02 \$/h. Using sum-of-squares of degree 2, one can construct a stronger convex relaxation whose optimal value is 5882.67 \$/h. The rank in the relaxation is equal to one, proving that this value is globally optimal. This is also the value found by the optimal power flow nonlinear solver `runopf` in MATPOWER. The nonlinear solver can potentially become stuck in a saddle point or local minimum, and thus returns an upper bound. Since this upper bound is equal to the lower bound found using sum-of-squares, this is another means of confirming that we have found the global value.

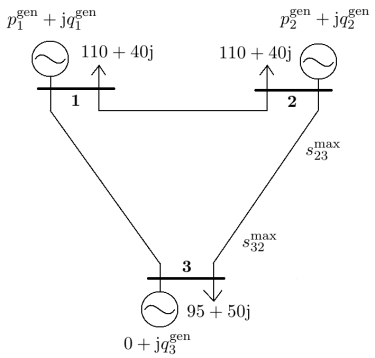

Figure 2: LMBM3 3-bus system

This example is important from a power systems perspective because it has a cycle, which is characteristic of transmission grids. In contrast, distribution grids which operate at a lower voltage level can exhibit a tree structure devoid of cycles.

3.1. Polynomial optimization

The optimal power flow problem can be seen as a special case of polynomial optimization once it is converted into real numbers. This class of problems admits a general procedure for solving them to global optimality. We next provide some of its main features and an example.

Consider a polynomial optimization problem

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \qquad f(\boldsymbol{x}) \triangleq \sum_{\alpha} f_\alpha \boldsymbol{x}^\alpha \tag{21a}$$

$$\text{subject to} \qquad g_i(\boldsymbol{x}) \triangleq \sum_{\alpha} g_{i,\alpha} \boldsymbol{x}^\alpha \geqslant 0, \quad i = 1, \ldots, m, \tag{21b}$$

where we use the multi-index notation $\boldsymbol{x}^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for $\boldsymbol{x} \in \mathbb{R}^n$, $\alpha \in \mathbb{N}^n$, and where the data are polynomials $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ so that in the above sums only a finite number of coefficients f_α and $g_{i,\alpha}$ are nonzero. We will use the notation $|\alpha| := \sum_{k=1}^n \alpha_k$.

In 2001, the Lasserre hierarchy [105, 106] (see also [166, 167]) was proposed to find global solutions to polynomial optimization problems. It is also known as moment/sum-of-squares hierarchy in reference to the primal moment hierarchy and the dual sum-of-squares hierarchy. Its global convergence is guaranteed by Putinar's Positivstellensatz [172] proven in 1993. Typically, if one of the constraints is a ball $x_1^2 + \ldots + x_n^2 \leqslant 1$, then the sequence of lower bounds provided by the hierarchy converges to the global infimum of the polynomial optimization problem. In addition, there is zero duality at all relaxation orders [85]. This is a crucial property when using path-following primal-dual interior point methods, which are some of the most efficient approaches for solving semidefinite programs.

The moment problem of order d is defined as[2]

$$\underset{\boldsymbol{y} \in \mathbb{R}^{\binom{n+2d}{2d}}}{\text{minimize}} \quad L_{\boldsymbol{y}}(f) \tag{23a}$$

$$\text{subject to} \quad \boldsymbol{y}_0 = 1 \tag{23b}$$

$$\boldsymbol{M}_d(\boldsymbol{y}) \succcurlyeq 0 \tag{23c}$$

$$\boldsymbol{M}_{d-k_i}(g_i \boldsymbol{y}) \succcurlyeq 0, \quad i = 1, \ldots, m \tag{23d}$$

where the Riesz functional, the moment matrix, and the localizing matrices are respectively defined by

$$L_{\boldsymbol{y}}(f) \triangleq \sum_{\alpha} f_\alpha y_\alpha \tag{24a}$$

$$\boldsymbol{M}_d(\boldsymbol{y}) := (y_{\alpha+\beta})_{|\alpha|,|\beta| \leqslant d} \tag{24b}$$

$$\boldsymbol{M}_{d-k_i}(g_i \boldsymbol{y}) \triangleq (\sum_{\gamma} g_{i,\gamma} y_{\alpha+\beta+\gamma})_{|\alpha|,|\beta| \leqslant d-k_i} \tag{24c}$$

$$k_i \triangleq \max\{\lceil |\alpha|/2 \rceil \text{ s.t. } g_{i,\alpha} \neq 0\}. \tag{24d}$$

Above, $\lceil . \rceil$ denotes the ceiling of a real number. The sum-of-squares problem of order d is defined as

$$\underset{\lambda, \sigma}{\text{maximize}} \quad \lambda \tag{25a}$$

$$\text{subject to} \quad f - \lambda = \sigma_0 + \sum_{k=1}^{m} \sigma_k g_k \tag{25b}$$

$$\lambda \in \mathbb{R}, \sigma_0 \in \Sigma_d[\boldsymbol{x}], \tag{25c}$$

$$\sigma_i \in \Sigma_{d-k_i}[\boldsymbol{x}], \quad i = 1, \ldots, m. \tag{25d}$$

A polynomial $\sigma(\boldsymbol{x}) = \sum_{|\alpha| \leqslant 2d} \sigma_\alpha \boldsymbol{x}^\alpha$ is a sum-of-squares, i.e. it belongs to $\Sigma_d[\boldsymbol{x}]$, if it is of the form

$$\sigma(\boldsymbol{x}) = \sum_{k} \left(\sum_{|\alpha| \leqslant d} p_{k,\alpha} \boldsymbol{x}^\alpha \right)^2 \quad \text{where} \quad p_{k,\alpha} \in \mathbb{R}. \tag{26}$$

This is equivalent to the existence of $(\varphi_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d} \succcurlyeq 0$ such that $\sum_{|\alpha| \leqslant 2d} \sigma_\alpha \boldsymbol{x}^\alpha = \sum_{|\alpha|,|\beta| \leqslant d} \varphi_{\alpha,\beta} \boldsymbol{x}^{\alpha+\beta}$.

We now focus on the result that justifies the global convergence of the Lasserre hierarchy, namely Putinar's Positivstellensatz [172]. This result provided a crucial refinement of Schmüdgen's Positivstellensatz [175] proven a few years earlier. It was crucial because it enabled numerical computations, leading to what is known today as the Lasserre hierarchy. Schmüdgen's Positivstellensatz essentially says that a polynomial that is positive on a set defined by polynomial inequalities can be decomposed as a sum of products of the polynomials multiplied by sums of squares; Putinar's removes the product from the decomposition. This

[2] We use the notation

$$\binom{n}{k} \triangleq \frac{n(n-1)\ldots(n-k+1)}{k!} \tag{22}$$

for integers $n \geqslant k \geqslant 1$ and where $k!$ stands for factorial.

can be seen in the theorem below.

Theorem 2 (Putinar's Positivstellensatz [172]). *Assume that there exist $R > 0$ and sums-of-squares $p_0, \ldots, p_m$ such that*

$$R^2 - x_1^2 - \cdots - x_n^2 = p_0 + \sum_{i=1}^{m} p_i g_i \tag{27}$$

If $f > 0$ on $\{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) \geqslant 0\, , \ldots\, ,\ g_m(\boldsymbol{x}) \geqslant 0\}$, then there exists sum-of-squares $\sigma_0, \ldots, \sigma_m$ such that

$$f = \sigma_0 + \sum_{i=1}^{m} \sigma_i g_i. \tag{28}$$

Based on this theorem, it is simple to see that whenever λ is a strict lower bound of f on the feasible set $\{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) \geqslant 0\, , \ldots\, ,\ g_m(\boldsymbol{x}) \geqslant 0\}$, the objective function minus the lower bound can be written using sums-of-squares. Thus, one may make the lower bound as closely as desired to the global infimum. In fact, in many cases, it is reached, as can be seen in the following example.

Example 1. *Consider the following polynomial optimization problem taken from [83]:*

$$\inf_{x_1, x_2 \in \mathbb{R}} \ x_1^2 + x_2^2 + 2x_1x_2 - 4x_1 - 4x_2 \ \ \textit{s.t.}\ \ x_1^2 + x_2^2 = 1$$

Its optimal value is $2 - 4\sqrt{2}$, which can be found using sums-of-squares since:

$$\begin{aligned} x_1^2 + x_2^2 + 2x_1x_2 - 4x_1 - 4x_2 \quad - \quad (2 - 4\sqrt{2}) \quad = \\ (\sqrt{2} - 1)(x_1 - x_2)^2 \ + \ \sqrt{2}(-\sqrt{2} + x_1 + x_2)^2 \\ + \ 2(\sqrt{2} - 1)(1 - x_1^2 - x_2^2) \end{aligned}$$

It can be seen from the above equation that when (x_1, x_2) is feasible, the first line must be nonnegative, proving that $2 - 4\sqrt{2}$ is a lower bound. This corresponds to the first-order Lasserre hierarchy since the polynomials inside the squares are of degree one at most.

In general, and not only in this example, the Lasserre hierarchy has finite convergence generically. This means that for a given abritary polynomial optimization problem, finite convergence will almost surely hold. It was Nie [162] who proved this result, which had been observed in practice ever since the Lasserre hierarchy was introduced. He relied on theorems of Marshall [139, 140] which attempted to answer the question: when can a nonnegative polynomial have a sum-of-squares decomposition? In Putinar's Positivstellensatz above, the assumption of positivity is made, which only guarantees asymptotic convergence. Parallel to Lasserre's contribution, Parrilo [166] pioneered the use of sum-of-squares for obtaining strong bounds on the optimal solution of nonconvex problem. He also showed how they can be used for many important problems in systems and control. In contrast to Lasserre, Parrilo's work [167] panders to Stengle's Positivstellensatz [183], which is used for proving infeasibility of systems of polynomial equations.

To sum up, the Lasserre hierarchy provides a sequence of semidefinite programs whose optimal values converge (monotonically) towards the global value of a polynomial optimization problem. This is true provided that the feasible set is compact, and that a bound R on the radius of the set is known, so that one can include a redundant ball constraint $x_1^2 + \cdots + x_n^2 \leqslant R^2$. In the case of the optimal power flow, this is not necessary because the upper bounds on voltage suffice to guarantee convergence of the Lasserre hierarchy.

3.2. Exploiting sparsity

In full generality, sum-of-squares are only applicable to problems of modest size. But with sparsity, the picture is different. Thankfully, power systems are sparse and the physical equations make them amenable to sum-of-squares.

Waki *et al.* [201] proposed to use chordal sparsity in the Lasserre hierarchy. They draw on the *correlative sparsity* graph whose vertices are the variables and whose edges signify that two variables appear simultaneously either in a constraint or in a monomial of the objective. The idea of Waki *et al.* is to restrain the variables appearing in the sum-of-squares (*a priori* all variables) to subsets of variables. The approach of Waki *et al.* reduces the computational burden of the Lasserre hierarchy for sparse problems. Concerning the optimal power flow problem, it allows one to solve some hard instances to global optimality with up to 80 variables [66] (instead of 20 without exploiting sparsity [144]). However, by using the correlative sparsity graph discussed above, a lot of the sparsity is lost. The authors of [87] thus propose a finer notion of sparsity that takes advantage of the fact that the constraints are polynomials. To that effect, they define the *monomial sparsity* graph whose vertices are the variables and whose edges signify that two variables appear simultaneously in a monomial of either the objective or a constraint. Consider the formula for complex electric power which was exposed in one of the previous sections:

$$v_k\overline{i_k} = v_k\overline{\sum_l i_{kl}} = v_k\overline{\sum_l y_{kl}(v_k - v_l)} = \left(\sum_l \overline{y}_{kl}\right)|v_k|^2 - \sum_l \overline{y}_{kl} v_k \overline{v}_l \tag{29}$$

The above computations rely on Ohm's law and Kirchhoff's laws. They lead to the main equation that appears in the optimal power flow and many of its variants. This equation leads to constraints like

$$2v_1\overline{v}_1 - (1+\mathrm{i})v_1\overline{v}_2 - (2-\mathrm{i})v_1\overline{v}_3 - (4+3\mathrm{i})v_1\overline{v}_4 = 1 - 3\mathrm{i} \tag{30}$$

where the constant terms are chosen arbitrarily. Notice that the sparsity pattern associated to this constraint

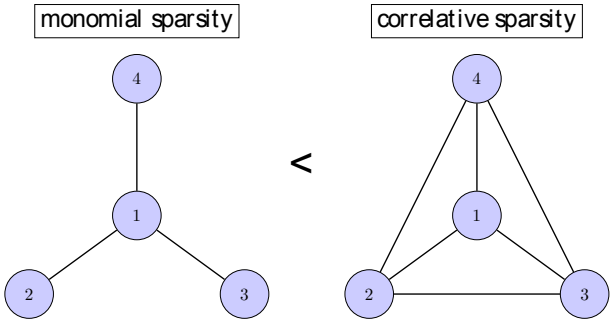

Figure 3: Two different notions of sparsity

is exactly that the one described in Figure 3. The reason why the monomial sparsity and correlative sparsity differ is due in part to the Kirchhoff. One variable ends up multiplying itself and several other variables, but those variables do not get multiplied among themselves. Accordingly, the finer notion of sparsity (monomial sparsity) leads to semidefinite constraints that are smaller. This enables an improved tractability. In addition, monomial sparsity allows one to apply different relaxation orders at each constraint.

3.3. Exploiting symmetry

Consider the problem of finding global solutions to a complex polynomial optimization problem

$$\underset{\boldsymbol{z}\in\mathbb{C}^n}{\text{minimize}} \quad f(\boldsymbol{z},\bar{\boldsymbol{z}}) \triangleq \sum_{\alpha,\beta} f_{\alpha,\beta}\boldsymbol{z}^\alpha\bar{\boldsymbol{z}}^\beta \tag{31a}$$

$$\text{subject to} \quad g_i(\boldsymbol{z},\bar{\boldsymbol{z}}) \triangleq \sum_{\alpha,\beta} g_{i,\alpha,\beta}\boldsymbol{z}^\alpha\bar{\boldsymbol{z}}^\beta \geqslant 0, \quad i=1,\ldots,m. \tag{31b}$$

We again use the multi-index notation $\boldsymbol{z}^\alpha := z_1^{\alpha_1}\cdots z_n^{\alpha_n}$ for $\boldsymbol{z}\in\mathbb{C}^n$, $\alpha\in\mathbb{N}^n$. The functions $f, g_1, \ldots, g_m$ are real-valued polynomials so that in the above sums only a finite number of coefficients $f_{\alpha,\beta}$ and $g_{i,\alpha,\beta}$ are nonzero and they satisfy $\overline{f_{\alpha,\beta}} = f_{\beta,\alpha}$ and $\overline{g_{i,\alpha,\beta}} = g_{i,\beta,\alpha}$. The feasible set is defined as $K := \{\boldsymbol{z}\in\mathbb{C}^n \,:\, g_i(\boldsymbol{z},\bar{\boldsymbol{z}})\geqslant 0,\ i=1,\ldots,m\}$. Complex numbers are used in the optimal power flow problem due to alternating current. In fact, the optimal power flow problem is an instance of a complex polynomial optimization problem.

Example 2. *In* [27, WB2, $V_2^{\max} = 1.022$ p.u.]*, an instance of the optimal power flow is proposed. It yields the following complex polynomial optimization problem*

$$\underset{z_1,z_2\in\mathbb{C}}{\text{minimize}} \quad 8|z_1-z_2|^2 \tag{32a}$$

$$\text{subject to} \quad 0.9025 \leqslant |z_1|^2 \leqslant 1.1025 \tag{32b}$$

$$0.9025 \leqslant |z_2|^2 \leqslant 1.0568 \tag{32c}$$

$$(2+10\mathrm{i})z_1\bar{z}_2 + (2-10\mathrm{i})z_2\bar{z}_1 - 4|z_2|^2 = 350 \tag{32d}$$

$$(-10+2\mathrm{i})z_1\bar{z}_2 + (-10-2\mathrm{i})z_2\bar{z}_1 + 20|z_2|^2 = -350 \tag{32e}$$

Notice that if (z_1, z_2) *is a feasible point, then so is* $(e^{\mathrm{i}\theta}z_1, e^{\mathrm{i}\theta}z_2)$ *for all* $\theta\in\mathbb{R}$ (this is the result of alternating current). *When converted to real numbers* $z_1 \triangleq x_1 + x_3\mathrm{i}$ *and* $z_2 \triangleq x_2 + x_4\mathrm{i}$*, it yields*

$$\underset{x_1,x_2,x_3,x_4\in\mathbb{R}}{\text{minimize}} \quad 8(x_1-x_2)^2 + 8(x_3-x_4)^2 \tag{33a}$$

$$\text{subject to} \quad 0.9025 \leqslant x_1^2 + x_3^2 \leqslant 1.1025 \tag{33b}$$

$$0.9025 \leqslant x_2^2 + x_4^2 \leqslant 1.0568 \tag{33c}$$

$$4x_1x_2 + 4x_3x_4 + 20x_1x_4 - 20x_3x_2 - 4x_2^2 + 4x_4^2 = 350 \tag{33d}$$

$$-20x_1x_2 - 20x_3x_4 + 4x_1x_4 - 4x_3x_2 + 20x_2^2 + 20x_4^2 = -350 \tag{33e}$$

Notice that if (x_1, x_2, x_3, x_4) *is a feasible point, then so is* $(-x_1, -x_2, -x_3, -x_4)$.

The above symmetries allow one to cancel many terms in the Lasserre hierarchy at no loss of bound quality. A complex version of the Lasserre hierarchy can be constructed by using D'Angelo and Putinar's Positivstellensatz [48], similar to Putinar's Positivstellensatz, but for Hermitian sum-of-squares rather than regular real sum-of-squares. On the examples above, the real and complex hierarchies yield the same bounds at the first, second, and third orders (888.1, 894.3, and 905.7 MW *respectively*)*, but the complex hierarchy is cheaper to compute due to the reduced size of the semidefinite constraints. The rank of the real and complex moment matrices guarantee that global convergence is reached at the third order.*

We conclude this section by noting that for general polynomial optimization problems (not necessarily complex), one may exploit symmetry using the techniques developed in [173].

3.4. Inner approximations of the sum-of-squares cone

We next discuss some restrictions one can impose on the set of sum-of-squares polynomial in order to transform it from a semidefinite cone into a second order cone. Following [135, 1], a polynomial $\sigma(x) = \sum_{|\alpha|\leqslant 2d} \sigma_\alpha x^\alpha$ is a scaled diagonally-dominant sum-of-squares (SDSOS) if it is of the form

$$\sigma(x) = \sum_k \left(p_k \boldsymbol{x}^{\alpha(k)} + q_k \boldsymbol{x}^{\beta(k)}\right)^2 \quad \text{where} \quad \begin{array}{c} \alpha(k), \beta(k) \in \mathbb{N}^n, \\ p_k, q_k \in \mathbb{R}. \end{array} \tag{34}$$

This is equivalent to the existence of $(\varphi_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d}$ such that $\sum_{|\alpha|\leqslant 2d} \sigma_\alpha \boldsymbol{x}^\alpha = \sum_{|\alpha|,|\beta|\leqslant d} \varphi_{\alpha,\beta} \boldsymbol{x}^{\alpha+\beta}$ where φ is of the form

$$\sum_{\substack{|\alpha|,\,|\beta| \leqslant d \\ \alpha \neq \beta}} \phi^{\alpha,\beta} \tag{35}$$

where each matrix $\phi^{\alpha,\beta}$ has zeros everywhere expect for four entries $\phi^{\alpha,\beta}(\alpha,\alpha)$, $\phi^{\alpha,\beta}(\alpha,\beta)$, $\phi^{\alpha,\beta}(\beta,\alpha)$, $\phi^{\alpha,\beta}(\beta,\beta) \in \mathbb{R}$ such that

$$\begin{pmatrix} \phi^{\alpha,\beta}(\alpha,\alpha) & \phi^{\alpha,\beta}(\alpha,\beta) \\ \phi^{\alpha,\beta}(\beta,\alpha) & \phi^{\alpha,\beta}(\beta,\beta) \end{pmatrix} \succcurlyeq 0. \tag{36}$$

These can be viewed as second-order conic constraints. In the sum-of-squares problem of order d, if we restrain the sum-of-squares variables $\sigma_0, \ldots, \sigma_m$ to be SDSOS, then in the moment problem of order d, we are relaxing each semidefinite constraint as follows

$$\begin{pmatrix} M_d(y)(\alpha,\alpha) & M_d(y)(\alpha,\beta) \\ M_d(y)(\beta,\alpha) & M_d(y)(\beta,\beta) \end{pmatrix} \succcurlyeq 0\,, \quad \forall |\alpha|,|\beta| \leqslant d,\ \ \alpha \neq \beta, \tag{37}$$

$$\begin{pmatrix} M_{d-k_i}(g_i y)(\alpha,\alpha) & M_{d-k_i}(g_i y)(\alpha,\beta) \\ M_{d-k_i}(g_i y)(\beta,\alpha) & M_{d-k_i}(g_i y)(\beta,\beta) \end{pmatrix} \succcurlyeq 0\,, \quad \forall |\alpha|,|\beta| \leqslant d-k_i,\ \ \alpha \neq \beta. \tag{38}$$

Naturally, these can also be viewed as second-order conic constraints. For their application to polynomial optimization, see [101, (QM-DSOSr),(QM-SDSOSr)] and [102]. In particular, they have been applied to the optimal power flow problem, enabling one to compute lower bound to medium-sized test cases (several hundreds of variables). In a similar flavor, in [145], the moment constraint is maintained as a positive semidefinite constraint, but the localizing matrices are relaxed to multiple second-order conic constraints. This guarantees that the relaxation is stronger than the first-order Lasserre relaxation. In some instances, there is a computational gain, but in others, the approach is unable to find a minimizer whereas the Lasserre hierarchy does.

4. Numerical algorithms for semidefinite programming

In the previous two sections, we have discussed general methodologies for solving nonconvex optimization problems to global optimality. These rely on solving a conic optimization efficiently on a large scale. We focus below on numerical algorithms for solving semidefinite programming, of which second-order conic programming and linear programming are subclasses. Three general approaches are presented. We begin with first-order methods which have a small cost per iteration and include for example the alternation direction method of multipliers (ADMM). Next, we explain the reasoning behind interior-point methods

which were originally designed for linear programming. Last, we discuss nonconvex methods that reformulate the semidefinite programming into a nonconvex problem by taking advantage of the fact that we are interested in low-rank solutions.

4.1. Semidefinite programming

To streamline the presentation, we focus on the standard SDP formulation:

$$\begin{aligned} \underset{\boldsymbol{X} \in \mathbb{S}^n}{\text{minimize}} \quad & \boldsymbol{C} \bullet \boldsymbol{X} && \text{(39a)} \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b} && \text{(39b)} \\ & \boldsymbol{X} \succeq 0, && \text{(39c)} \end{aligned}$$

where $\boldsymbol{C} \in \mathbb{S}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, and $\mathcal{A}(\boldsymbol{X}) = \begin{bmatrix} \boldsymbol{A}_1 \bullet \boldsymbol{X} & \cdots & \boldsymbol{A}_m \bullet \boldsymbol{X} \end{bmatrix}^\top$ is the linear matrix operator. The dual problem is given by:

$$\begin{aligned} \underset{\boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{S} \in \mathbb{S}^n}{\text{maximize}} \quad & \boldsymbol{b}^\top \boldsymbol{y} && \text{(40a)} \\ \text{subject to} \quad & \mathcal{A}^*(\boldsymbol{y}) + \boldsymbol{S} = \boldsymbol{C} && \text{(40b)} \\ & \boldsymbol{S} \succeq 0, && \text{(40c)} \end{aligned}$$

where $\mathcal{A}^*(\boldsymbol{y}) = \sum_{i=1}^m y_i \boldsymbol{A}_i$ is the adjoint operator of $\mathcal{A}$. We state some nondegeneracy assumptions, which are standard for complexity bound analysis and satisfied for a wide range of problems [3].

Assumption 1 (Nondegeneracy condition)**.** *Given the primal and dual problems, assume that:*

1. *(Linear independence) The matrix $\boldsymbol{A} = \begin{bmatrix} \operatorname{vec}\{\boldsymbol{A}_1\} & \cdots & \operatorname{vec}\{\boldsymbol{A}_m\} \end{bmatrix}$ has full column-rank (i.e., $\boldsymbol{A}^\top \boldsymbol{A}$ is nonsingular);*
2. *(Slater's condition) There exist $\boldsymbol{y}$ and $\boldsymbol{X}, \boldsymbol{S} \succ 0$ such that $\mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$ and $\mathcal{A}^*(\boldsymbol{y}) + \boldsymbol{S} = \boldsymbol{C}$ are satisfied.*

The linear independence condition requires that the number of constraints m does not exceed the degrees of freedom $\frac{1}{2}n(n+1)$. By embedding the problem within a slightly larger problem using the homogeneous self-dual embedding technique [210], Slater's condition can be also satisfied. While we focus on the standard formulations as in (39) and (40), there is no loss of generality. For instance, to incorporate an inequality $\boldsymbol{A}_i \bullet \boldsymbol{X} \geq b_i$, we can introduce $\bar{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0}^\top & x \end{bmatrix}$ that extends the dimension of $\boldsymbol{X}$ by 1, and introduce an extra equality equation that constrains the element $x = \boldsymbol{A}_i \bullet \boldsymbol{X} - b_i$. In addition, all of the complexity bounds discussed below can be generalized to conic programs posed on the Cartesian product of multiple semidefinite cones $\mathcal{K} = \boldsymbol{S}_+^{n_1} \times \boldsymbol{S}_+^{n_2} \times \cdots \times \boldsymbol{S}_+^{n_\ell}$.

4.2. First-order methods

One of the most promising first-order methods for solving large-scale SDP problems is the alternating direction method of multipliers (ADMM), which is closely related to the augmented Lagrangian method proposed in the mid-1970s [60] and [67]. The ADMM algorithm experienced a revival in the past decade, in a large part due to the publication of a popular and influential survey by Boyd et al. [26] for applications in distributed optimization and statistical learning. Two popular variations of ADMM for solving large-scale

SDPs are proposed by Wen, Goldfarb and Yin [207] and O'Donoghue et. al [163]. Following the treatment of [207], we start by defining the augmented Lagrangian function for the dual problem:

$$\mathcal{L}_\mu(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}) = -\boldsymbol{b}^\top \boldsymbol{y} + \boldsymbol{X} \bullet (\mathcal{A}^*(\boldsymbol{y}) + \boldsymbol{S} - \boldsymbol{C}) + \frac{1}{2\mu}\|\mathcal{A}^*(\boldsymbol{y}) + \boldsymbol{S} - \boldsymbol{C}\|_F^2, \tag{41}$$

where $\boldsymbol{X} \in \mathbb{S}^n$ and $\mu > 0$ is the penalty coefficient. By adding the quadratic term to the Lagrangian function, the augmented Lagrangian $\mathcal{L}_\mu(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S})$ is strongly convex in $\boldsymbol{y}$ and $\boldsymbol{X}$ and has a Lipschitz-continuous gradient. This essentially smoothes the primal problem and allows a gradient-based optimization to be effectively employed. Starting from the initial primal variable $\boldsymbol{X}^0$, the augmented Lagrangian method solves in each iteration

$$\underset{\boldsymbol{y}\in\mathbb{R}^m, \boldsymbol{S}\in\mathbb{S}^n}{\text{minimize}}\ \mathcal{L}_\mu(\boldsymbol{X}^k, \boldsymbol{y}, \boldsymbol{S}) \qquad \text{subject to} \quad \boldsymbol{S} \succeq 0, \tag{42}$$

for $\boldsymbol{y}^{k+1}$ and $\boldsymbol{S}^{k+1}$, and then updates the primal variable $\boldsymbol{X}^{k+1}$ by

$$\boldsymbol{X}^{k+1} := \boldsymbol{X}^k + \frac{\mathcal{A}^*(\boldsymbol{y}^{k+1}) + \boldsymbol{S}^{k+1} - \boldsymbol{C}}{\mu}. \tag{43}$$

Instead of jointly minimizing $\mathcal{L}_\mu(\boldsymbol{X}^k, \boldsymbol{y}, \boldsymbol{S})$ with respect to $\boldsymbol{y}$ and $\boldsymbol{S}$, which can be very time consuming for large-scale problems, ADMM minimizes the augmented Lagrangian function with respect to $\boldsymbol{y}$ and $\boldsymbol{S}$ separately, and hence no longer solving problem (42) exactly [207]:

$$\boldsymbol{y}^{k+1} := \arg\min_{\boldsymbol{y}\in\mathbb{R}^m}\ \mathcal{L}_\mu(\boldsymbol{X}^k, \boldsymbol{y}, \boldsymbol{S}^k) \tag{44a}$$

$$\boldsymbol{S}^{k+1} := \arg\min_{\boldsymbol{S}\in\mathbb{S}^n}\ \mathcal{L}_\mu(\boldsymbol{X}^k, \boldsymbol{y}^{k+1}, \boldsymbol{S}) \qquad \text{s.t.} \quad \boldsymbol{S} \succeq 0. \tag{44b}$$

The above order of (43), (44a) and (44b) is not important, and the iterates converge towards the solutions of (39) and (40) for all fixed $\mu > 0$ [207, Thm 2]. Here, the $\boldsymbol{y}$-update has a closed-form solution based on first-order optimality of the unconstrained problem

$$\boldsymbol{y}^{k+1} := -(\boldsymbol{A}^\top \boldsymbol{A})^{-1}\left(\mu(\mathcal{A}(\boldsymbol{X}^k) - \boldsymbol{b}) + \mathcal{A}\left(\boldsymbol{S}^k - \boldsymbol{C}\right)\right), \tag{45}$$

where $\boldsymbol{A}^\top \boldsymbol{A}$ is invertible by Assumption (1), and the $\boldsymbol{S}$-update is given by the projection of a symetric matrix onto the positive semidefinite cone:

$$\boldsymbol{S}^{k+1} = [\boldsymbol{V}]_+, \qquad \text{where } \boldsymbol{V} = \boldsymbol{C} - \mathcal{A}^*(\boldsymbol{y}^{k+1}) - \mu\boldsymbol{X}^k, \tag{46}$$

which also has a closed-form solution based on the spectral decomposition

$$\boldsymbol{V} = \sum_{i=1}^n \sigma_i \boldsymbol{v}_i \boldsymbol{v}_i^\top, \qquad [\boldsymbol{V}]_+ = \sum_{i=1}^n \max(\sigma_i, 0)\boldsymbol{v}_i \boldsymbol{v}_i^\top. \tag{47}$$

The above iteration is closely related to the regularization method [136] and the boundary point method [170], where $\boldsymbol{X}^k$ is fixed until $\frac{1}{\mu}(\boldsymbol{S}^{k+1} - \boldsymbol{V}^{k+1})$ is nearly feasible. Although the sequence is convergent, in practice, a heuristic based on balancing the primal and dual residuals to adjust the value of μ seems to help the numerical performance (see [207, Sec. 3.2] and [26, Sec. 3.4] for implementation details).

Since the OPF power problem has sparsity patterns that arise from the network topology, the capa-

bility of solving large-scale SDPs can be enhanced by exploiting the sparsity pattern and parallelization [59, 188, 131]. By defining a representative graph for the large-scale SDP problem, we can reduce the SDP formulation using a tree/chordal/clique decomposition technique introduced above (Theorem 1). This decomposition replaces the large-scale SDP matrix variable $\boldsymbol{X}$ with the submatrices $\boldsymbol{X}\{\mathcal{B}_k\}$ using a tree decomposition of the power network, where $\mathcal{B}_k$ is a bag of vertices of $\mathcal{N}$. Similar to (43), (44a) and (44b), a distributed ADMM-based algorithm can be derived to solve the reduced SDP problem iteratively. The main speedup is achieved in the $\boldsymbol{S}$-update, where instead of performing spectral decomposition on a large $n \times n$ matrix (performing this step using dense linear algebra requires $\Theta(n^3)$ time and $\Theta(n^2)$ memory), this expensive operation is needed for only multiple submatrices of orders bounded by the treewidth. By finding the optimal solution for the distributed SDP, one could recover the solution to the original SDP formulation using an explicit formula (see [131, Sec. IV-B] for the implementation). The idea of tree decomposition of a sparse SDP into smaller sized SDPs is first proposed in [59], which then solves the subproblems by interior point methods. Using a first-order splitting method, [188] solves the decomposed SDP problem created by [59], but the algorithm needs to solve an optimization subproblem at every iteration. The employment of ADMM to solve the reduced problem is proposed in [130, 131]. It also studies the application to OPF in real-world grids (for instance, the SDP relaxation of OPF for a European grid with 9241 buses [84] amounts to simple operations over 857 matrices of size 31 by 31 and 14035 matrices of size 2 by 2).

As for the convergence rate of ADMM, it has been shown in [72] that the sequence converges with sublinear objective error $O(1/k)$ in an ergodic sense, and the method converges to L accurate digits in $O(\exp(L))$ iterations in the worst case. In practice, it has been shown that ADMM often performs much better, converging to L accurate digits in just $O(L)$ iterations for a wide range of SDP test problems [207, 219]. The sparse structure that arises from OPF can be further exploited in the ADMM updates. The y-update is dominated by the solution of $(\boldsymbol{A}^\top \boldsymbol{A})y = r$, which has the worst-case complexity of $O(n^6)$ time and $O(n^4)$ memory using standard Cholesky factorization for fully dense matrix $\boldsymbol{A}$. However, for large and sparse $\boldsymbol{A}$, efficiency can be substantially improved by using an incomplete factorization as the preconditioner within an iterative solution algorithm like conjugate gradients [174], and by reordering the columns of $\boldsymbol{A}$ using a fill-minimizing ordering like minimum degree and nested dissection [65]. In addition to SDP reduction via tree decomposition, the $\boldsymbol{S}$-update can be further improved when the optimal solution $\boldsymbol{X}^*$ is known to be low-rank, where it may be possible to use low-rank linear algebra and an iterative spectral decomposition such as Lanczos iterations to reduce the complexity to as low as $O(n)$ per iteration.

In addition to ADMM, other first-order methods include smooth gradient methods [157], augmented Lagrangian methods [98], conjugate gradients [192, 218], which are applied either to (39) directly, or to the Hessian equation associated with an interior-point solution. Because all of these algorithms have inexpensive per-iteration costs but a sublinear worst-case convergence rate, they can compute an ϵ-accurate solution in $O(1/\epsilon)$ time. As a result, they are most commonly used to solve very large-scale SDPs with moderate accuracy requirements.

4.3. Interior-point methods

The modern study of interior-point methods was initiated by Karmarkar [90] and their extension to SDPs was due to Nesterov and Nemirovsky [158] and Alizadeh [2]. The development of primal-dual interior-point methods began with Kojima et al. [100, 99] and was further developed in early papers such as [196] and [75]. It was eventually extended to SDP and SOCP in a unified way by Nesterov and Todd [159, 160]. To solve SDPs, modern solvers, such as SDPT3 [194], SeDuMi [185, 186], and MOSEK [5], rely predominantly on interior-point methods. The interior-point method can be applied to SDP by introducing the logarithmic barrier function for positive semi-definite (PSD) cones, which is equivalent to the log-determinant penalty

for determinant maximization [197], since

$$-\sum_{j=1}^{n} \log \lambda_j(\boldsymbol{X}) = -\log \prod_{j=1}^{n} \lambda_j(\boldsymbol{X}) = -\log\det \boldsymbol{X},$$

where the eigenvalues $\lambda_j(\boldsymbol{X})$ are nonnegative for all $j \in [n]$. By replacing the PSD constraints in the primal and dual formulations (39) and (40), one can obtain a sequence of problems with only linear equality constraints:

$$\begin{aligned} \underset{\boldsymbol{X} \in \mathbb{S}^n}{\text{minimize}} \quad & \boldsymbol{C} \bullet \boldsymbol{X} - \mu \log\det \boldsymbol{X} && \text{(48a)} \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}, && \text{(48b)} \end{aligned}$$

and

$$\begin{aligned} \underset{\boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{S} \in \mathbb{S}^n}{\text{maximize}} \quad & \boldsymbol{b}^\top \boldsymbol{y} + \mu \log\det \boldsymbol{S} && \text{(49a)} \\ \text{subject to} \quad & \mathcal{A}^*(\boldsymbol{y}) + \boldsymbol{S} = \boldsymbol{C}, && \text{(49b)} \end{aligned}$$

whose solutions are denoted as $\boldsymbol{X}_\mu$ and $(\boldsymbol{y}_\mu, \boldsymbol{S}_\mu)$, respectively. It is straightforward to show that (48) and (49) form primal-dual pairs up to a constant offset. Typically, directly solving either the primal-scaled or the dual-scaled problem can suffer from accuracy and robustness issues. One popular method, called primal-dual interior method, aims at solving (48) and (49) simultaneously by resolving their joint Karush-Kuhn-Tucker (KKT) optimality condition:

$$\begin{aligned} \text{Primal feasibility:} \quad & \mathcal{A}(\boldsymbol{X}_\mu) = \boldsymbol{b} && \text{(50a)} \\ \text{Dual feasibility:} \quad & \mathcal{A}^*(\boldsymbol{y}_\mu) + \boldsymbol{S} = \boldsymbol{C} && \text{(50b)} \\ \text{Complementarity slackness:} \quad & \boldsymbol{X}_\mu \boldsymbol{S}_\mu = \mu \boldsymbol{I} && \text{(50c)} \end{aligned}$$

where the barrier parameter $\mu > 0$ is connected to the duality gap of the original SDP formulation (39) and (40) in the following way:

$$n\mu = \boldsymbol{X}_\mu \bullet \boldsymbol{S}_\mu = \boldsymbol{C} \bullet \boldsymbol{X}_\mu - \boldsymbol{b}^\top \boldsymbol{y}_\mu,$$

and

$$\begin{aligned} \boldsymbol{C} \bullet \boldsymbol{X}^* \le \boldsymbol{C} \bullet \boldsymbol{X}_\mu \le \boldsymbol{C} \bullet \boldsymbol{X}^* + n\mu \\ \boldsymbol{b}^\top \boldsymbol{y}^* - n\mu \le \boldsymbol{b}^\top \boldsymbol{y}_\mu \le \boldsymbol{b}^\top \boldsymbol{y}^*, \end{aligned}$$

where $\boldsymbol{X}^*$ and $(\boldsymbol{y}^*, \boldsymbol{S}^*)$ are the optimal solution of (39) and (40), respectively. By gradually decreasing μ towards 0, the solutions $\{\boldsymbol{X}_\mu, \boldsymbol{y}_\mu, \boldsymbol{S}_\mu\}$ form a sequence of convergent iterates, known as central path, that eventually approaches the optimal solution. To ensure theoretical guarantees of convergence, primal-dual interior-point methods for SDP often use Newton's method to solve the KKT conditions in (50), while keeping each update within a wide neighborhood of the central path

$$\mathcal{N}_\infty(\gamma) \coloneqq \left\{ \{\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{S}\} \in \mathcal{F} : \lambda_{\min}(\boldsymbol{X}\boldsymbol{S}) \ge \frac{\gamma}{n} \boldsymbol{X} \bullet \boldsymbol{S} \right\},$$

where $\mathcal{F}$ is the feasible region for the primal and dual programs in (39) and (40). The extent of the neighborhood is characterized by $\gamma \in (0,1)$, which is typically chosen with a value like 10^{-3}. The algorithm is guaranteed to converge to an approximate solution accurate to L digits after at most $O(nL)$ Newton iterations. The convergence can be often achieved within 30 to 50 iterations in practice.

Many of the real problems like ACOPF have inherent structures such as sparsity and low rank, but standard off–the–shelf interior–point solvers cannot fully exploit these structures to improve convergence speed. To capture sparsity, several modifications or reformulations have been proposed [59, 154, 93, 6, 195, 134, 128, 215].

The first category of modifications focus on the solution of the Hessian equation, which dominates the per-iteration cost:

$$\boldsymbol{H}\boldsymbol{y} = \boldsymbol{r},$$

where $\boldsymbol{H} = \big[\boldsymbol{A}^\top(\boldsymbol{W}\otimes\boldsymbol{W})\boldsymbol{A}\big]$, $\boldsymbol{A} = \big[\mathrm{vec}\{\boldsymbol{A}_1\} \quad \cdots \quad \mathrm{vec}\{\boldsymbol{A}_m\}\big]$, and $\boldsymbol{W}$ is the positive definite scaling matrix. Despite that the data matrix $\boldsymbol{A}$ are sparse, the Hessian matrix $\boldsymbol{H}$ is often fully-dense because $\boldsymbol{W}$ is fully-dense. As a result, to solve for the Hessian equation, standard approach of dense Cholesky factorization needs to be employed for both sparse & low-rank problem and dense & high-rank problem, which result in approximately the same amount of time and memory use.

To deal with this issue, one can employ the preconditioined conjugate gradient (PCG) algorithm, which requires a single matrix–vector product with the governing matrix $\boldsymbol{H}$ and a single matrix–vector product with the inverse of the preconditioner $\tilde{\boldsymbol{H}}$ in each iteration [14]. The key to numerical performance is a preconditioner that has desirable characteristics, such as being similar to $\boldsymbol{H}$ in a spectral sense while much cheaper to invert. For instance, a preconditioner of a low–rank perturbation of the sparse matrix was proposed in [214], where

$$\tilde{\boldsymbol{H}} = \tau^2\boldsymbol{A}^\top\boldsymbol{A} + \boldsymbol{K}\boldsymbol{K}^\top, \tag{51}$$

where $\boldsymbol{K}$ is a matrix with rank at most $n\,\mathrm{rank}(\boldsymbol{X}^*)$, which is constructed to improve the conditioning of the preconditioner. In particular, it has been proved that the joint condition number $\kappa = \lambda_{\max}(\tilde{\boldsymbol{H}}^{-1}\boldsymbol{H})/\lambda_{\min}(\tilde{\boldsymbol{H}}^{-1}\boldsymbol{H})$ is an absolute constant $O(1)$. This can be combined with the standard PCG convergence result to yield that PCG with the described $\tilde{\boldsymbol{H}}$ can solve the Hessian equation to L digits of accruacy in at most $O(L)$ iterations.

Another approach exploits the sparsity structure of the problem through convex reformulations. A common approach is to use clique tree conversion which reformulates a size-n semidefinite cone into a series of smaller cones. The idea was first proposed in [59, 154], and was later refined in [93].

$$\underset{\boldsymbol{W}\in\mathbb{S}^n}{\text{minimize}} \sum_{k=1}^{B} \boldsymbol{C}_k \bullet \boldsymbol{W}\{\mathcal{B}_k\} \tag{52a}$$

$$\text{subject to} \sum_{k=1}^{B} \boldsymbol{A}_{i,k} \bullet \boldsymbol{W}\{\mathcal{B}_k\} = b_i, \quad i \in \{1,...,m\} \tag{52b}$$

$$\boldsymbol{W}\{\mathcal{B}_k\} \succeq 0, \quad k \in \{1,2,...,B\} \tag{52c}$$

Since the OPF problem has inherent sparsity structures, this reformulation is able to reduce the computation time from tens of hours to tens of minutes for some large-scale benchmark systems [80, 150, 134, 128]. While empirical results indicate that many instances of the reformulated problem can be solved in near-linear time using an interior-point method, there are problem instances that could attain the worst-case cubic complexity [59, 154, 93, 134, 128]. The key issue is due to the large number of overlapping constraints imposed during the reformulation, which can significantly increase the complexity to solve the normal

equations in each interior-point iteration. Specifically, these overlapping constraints may contribute up to $O(\omega^4 n^2)$ nonzero elements to the normal matrix, thereby pushing the per-iteration cost of an interior-point method to cubic $O(n^3)$ time and quadratic $O(n^2)$ memory. Further speed-ups were obtained by dropping some of the overlapping constraints, though the reformulation may no longer be exact [6]. Nevertheless, it has been shown in [215] that one can apply the dualization technique of Löfberg to (52) before solving the problem using a general-purpose interior-point solver. In this way, the overlapping constraints contribute exactly $\Theta(\omega^4 n)$ nonzero elements to the normal matrix of the dualized problem, which makes it possible to guarantee sparsity in the normal equations. A class of SDPs which includes ACOPF as a special case, called decoupled SDPs, have been defined in [215].

Definition 1 (Decoupled SDP)**.** *Given index sets $\mathcal{B} = \{\mathcal{B}_1, \cdots, \mathcal{B}_B\}$, we say that the linear constraints $\boldsymbol{A}_i \bullet \boldsymbol{W} = b_i$ is decoupled if there exists $\mathcal{B}_k \in \mathcal{B}$ and some choice of $\boldsymbol{A}_{i,k} \in \mathbb{S}^{|\mathcal{B}_k|}$ such that*

$$\boldsymbol{A}_i \bullet \boldsymbol{W} = \boldsymbol{A}_{i,j} \bullet \boldsymbol{W}\{\mathcal{B}_k\} \qquad \forall \boldsymbol{W} \in \mathbb{S}^n.$$

We say that an instance of (39) *is decoupled if all linear constraints are decoupled.*

In the particular case of a decoupled SDP, it has been shown that there exists an algorithm that: 1) converts (39) into an instance of (52); 2) solves the dualized version of (52) to L accurate digits; and 3) recovers a corresponding solution of (39) in $O(\omega^{6.5} \cdot n^{1.5} \cdot L)$ time and $O(\omega^4 \cdot n)$ memory space (see [215, Theorem 1] for details, which also gives a constructive proof of this result). It is worthwhile to mention that the original, non-dualized version of (52) is not guaranteed to achieve near-linear time complexity due to the existence of some explicit examples of decoupled SDPs that force the problem to attain its worst-case cubic time complexity [215]. This work thus provides a theoretical guarantee for efficiency as observed in existing empirical studies on ACOPF. It has been shown that the algorithms achieved near-linear time performance for an array of benchmark OPF systems with over 10000 buses, solving each problem instances to 7 digits of accuracy within 6 minutes.

4.4. Nonconvex methods

In addition to the first-order and interior point methods, there exist some nonconvex reformulations of the original convex SDP problem 39 to exploit the low-rank and sparsity structure. One important method, also known as the Burer–Monteiro method, is based on the outer product factorization $\boldsymbol{X}^\top = \boldsymbol{R}\boldsymbol{R}^\top$, where $\boldsymbol{R} \in \mathbb{R}^{n \times r}$ is a rank–r matrix [29, 89]. This reformulation is mainly motivated by the following theoretical result for low-rank SDP, which was proved concurrently in [15] and [168].

Theorem 3 ([15, 168])**.** *Consider* (39) *with m equality constraints. Then, there exists an optimal solution $\boldsymbol{X}^*$ of* (39) *with rank $\bar{r}$ satisfying $\bar{r}(\bar{r}+1)/2 \leq m$.*

This result indicates that as long as r (i.e., the number of columns of the reformulated matrix variable $\boldsymbol{R}$) is large enough, such a reformulation will have an optimal solution set that includes some or all optimal solutions to (39). This also leads to the following nonlinear program:

$$\begin{aligned} \underset{\boldsymbol{R} \in \mathbb{R}^{n \times r}}{\text{minimize}} \quad & \boldsymbol{C} \bullet (\boldsymbol{R}\boldsymbol{R}^\top) && \text{(53a)} \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{R}\boldsymbol{R}^\top) = \boldsymbol{b}, && \text{(53b)} \end{aligned}$$

which replaces the positive semidefinite constraint in (39) with the quaddratic equality constraint $\boldsymbol{X} = \boldsymbol{R}\boldsymbol{R}^\top$. An immediate advantage is the reduction of the number of variables from n^2 to $n \times r$, which

can be substantial when $\boldsymbol{R}$ is limited to be low-rank. However, the problem is no longer convex due to the nonlinear equality constraint; therefore, it requires some effective numerical algorithms to solve. In particular, a limited BFGS augmented Lagrangian algorithm was developed in [29], which optimizes over the augmented Lagrangian function:

$$\mathcal{L}(\boldsymbol{R}, \boldsymbol{y}, \sigma) = \boldsymbol{C} \bullet (\boldsymbol{R}\boldsymbol{R}^\top) - \sum_{i=1}^{m} y_i(\boldsymbol{A}_i \bullet (\boldsymbol{R}\boldsymbol{R}^\top) - b_i) + \frac{\sigma}{2} \sum_{i=1}^{m} (\boldsymbol{A}_i \bullet (\boldsymbol{R}\boldsymbol{R}^\top) - b_i)^2, \tag{54}$$

where the variables $\boldsymbol{R} \in \mathbb{R}^{n \times r}$ and $\boldsymbol{y} \in \mathbb{R}^m$ are unrestricted, and the penalty parameter σ is positive. The basic idea behind this function is the penalization of infeasible points to deal with the equality constraints. Specifically, the last term measures the violation of equality constraints scaled by the penalty coefficient σ. Under some reasonable assumptions, an iterative scheme can be used to obtain an estimate of $(\boldsymbol{y}^k, \sigma_k)$ given the current primal iterate $\boldsymbol{R}^k$, which can be shown to be a convergent sequence. The method demonstrated favorable empirical results compared to other standard techniques such as interior point method on some standard SDP benchmark examples. Despite the empirical success, the original paper [29] only provides guarantees of local convergence due the nonconvexity nature. Recently, the theory gap started to close with a series of papers [25, 11, 165, 17]. For instance, the paper [25] provides strong performance guarantees for the Burer–Monteiro approach in the particular cases of synchronization and community detection. In particular, it established strong correlation of the ground truth with second-order critical points. The paper [17] extended the result from exact second-order stationary points to approximate second-order stationary points, which often arise in practice, and showed that with high probability, the Burer–Monteiro approach can find the global optimal in the aforementioned scenarios. However, the proof is limited in the sense that the analysis relies on the manifold structure of the rank-restricted search space, which is not a mild condition in practice.

The application of Burer–Monteiro method to the OPF problem has been investigated in [137], which employs the well-established Augmented Lagrangian approach [44] and a parallel coordinate descent with a closed-form step. The method compares favorably with SDP based methods [128] on small-scale instances, but experienced numerical difficulties on large instances, such as the Polish network with more than 2,000 nodes. This motivates the combination of first- and second-order methods [122] and the need to exploit the parallelization feature of first-order methods.

5. Conclusion

Convex relaxations have come a long way in addressing the optimal power flow problem. Only a decade ago, finding globally optimal power flows on real-world benchmarks was considered an intractable problem. Today, it is possible to make accurate computations on networks with thousands of buses and tens of thousands of constraints with global guarantees. Several challenges remain however. First, when minimizing generation costs instead of total power production, no approach is yet able to systematically find a global minimizer on large-scale networks. It is promising that very good bounds can be found on the global objective function and that near-global solutions can be extracted. New convex relaxations could be designed to better handle non-convex objectives, as opposed to non-convex constraints. Second, a finer modeling of the power system should be considered. It would account for discrete aspects such as transformer ratios and the turning on or off of generating units. Third, dynamical aspects could be considered, as in transient stability analysis. Recent work [88] has shown that conic optimization is also promising in that regard.

References

[1] Ahmadi, A. and Majumdar, A. (2019). DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(2):193–230.

[2] Alizadeh, F. (1995). Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM journal on Optimization*, 5(1):13–51.

[3] Alizadeh, F., Haeberly, J.-P. A., and Overton, M. L. (1997). Complementarity and nondegeneracy in semidefinite programming. *Mathematical programming*, 77(1):111–128.

[4] Alsac, O., Bright, J., Prais, M., and Stott, B. (1990). Further developments in LP-based optimal power flow. *IEEE Transactions on Power Systems*, 5(3):697–711.

[5] Andersen, E. D. and Andersen, K. D. (2000). The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In *High performance optimization*, pages 197–232. Springer.

[6] Andersen, M. S., Hansson, A., and Vandenberghe, L. (2014). Reduced-complexity semidefinite relaxations of optimal power flow problems. *IEEE Transactions on Power Systems*, 29(4):1855–1863.

[7] Bahrami, S., Therrien, F., Wong, V. W., and Jatskevich, J. (2017). Semidefinite relaxation of optimal power flow for AC–DC grids. *IEEE Transactions on Power Systems*, 32(1):289–304.

[8] Bai, X. and Wei, H. (2009). Semi-definite programming-based method for security-constrained unit commitment with operational and optimal power flow constraints. *IET Generation, Transmission & Distribution*, 3(2):182–197.

[9] Bai, X. and Wei, H. (2011). A semidefinite programming method with graph partitioning technique for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 33(7):1309–1314.

[10] Bai, X., Wei, H., Fujisawa, K., and Wang, Y. (2008). Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6-7):383–392.

[11] Bandeira, A. S., Boumal, N., and Voroninski, V. (2016). On the low-rank approach for semidefinite programs arising in synchronization and community detection. In *Proc. of the Conference on Learning Theory*, pages 361–382.

[12] Baran, M. and Wu, F. (1989). Optimal Capacitor Placement on Radial Distribution Systems. *IEEE Transactions on Power Delivery*, 4(1):725–734.

[13] Barati, M. and Kargarian, A. (2017). A global algorithm for AC optimal power flow based on successive linear conic optimization. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5.

[14] Barrett, R., Berry, M. W., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and Van der Vorst, H. (1994). *Templates for the solution of linear systems: building blocks for iterative methods*, volume 43. SIAM.

[15] Barvinok, A. I. (1995). Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(2):189–202.

[16] Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on Modern Convex Optimization – Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series Optim. 2. SIAM.

[17] Bhojanapalli, S., Boumal, N., Jain, P., and Netrapalli, P. (2018). Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form. *arXiv preprint arXiv:1803.00186*.

[18] Bienstock, D. and Munoz, G. (2014). On linear relaxations of OPF problems. *arXiv preprint arXiv:1411.1120*.

[19] Bienstock, D. and Verma, A. (2015). Strong NP-hardness of AC power flows feasibility. *arXiv preprint arXiv:1512.07315*.

[20] Bingane, C., Anjos, M. F., and Le Digabel, S. (2018). Tight-and-cheap conic relaxation for the AC optimal power flow problem. *IEEE Transactions on Power Systems*, 33(6):7181–7188.

[21] Bloomberg New Energy Finance, J. (2016). New energy outlook 2018. `https://about.bnef.com/new-energy-outlook/`.

[22] Bose, S., Gayme, D. F., Chandy, K. M., and Low, S. H. (2015a). Quadratically constrained quadratic programs on acyclic graphs with application to power flow. *IEEE Transactions on Control of Network Systems*, 2(3):278–287.

[23] Bose, S., Gayme, D. F., Low, S., and Chandy, K. M. (2011). Optimal power flow over tree networks. In *Annual Allerton Conference on Communication, Control, and Computing*, pages 1342–1348.

[24] Bose, S., Low, S. H., Teeraratkul, T., and Hassibi, B. (2015b). Equivalent relaxations of optimal power flow. *IEEE Transactions on Automatic Control*, 60(3):729–742.

[25] Boumal, N., Voroninski, V., and Bandeira, A. (2016). The non-convex burer-monteiro approach works on smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, pages 2757–2765.

[26] Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122.

[27] Bukhsh, W., Grothey, A., McKinnon, K., and Trodden, P. (2013a). Local Solutions of the Optimal Power Flow Problem. *IEEE Transactions on Power Systems*, 28:4780–4788.

[28] Bukhsh, W. A., Grothey, A., McKinnon, K. I., and Trodden, P. A. (2013b). Local solutions of the optimal power flow problem. *IEEE Transactions on Power Systems*, 28(4):4780–4788.

[29] Burer, S. and Monteiro, R. D. (2003). A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357.

[30] Bynum, M., Castillo, A., Watson, J.-P., and Laird, C. D. (2019). Tightening McCormick relaxations toward global solution of the ACOPF problem. *IEEE Transactions on Power Systems*, 34(1):814–817.

[31] Cain, M. B., O'Neil, R. P., and Castillo, A. (2012). History of Optimal Power Flow and Formulations (OPF Paper 1). Technical report, US Federal Energy Regulatory Commission.

[32] Capitanescu, F. (2016). Critical review of recent advances and further developments needed in AC optimal power flow. *Electric Power Systems Research*, 136:57–68.

[33] Capitanescu, F., Ramos, J. M., Panciatici, P., Kirschen, D., Marcolini, A. M., Platbrood, L., and Wehenkel, L. (2011). State-of-the-art, challenges, and future trends in security constrained optimal power flow. *Electric Power Systems Research*, 81(8):1731–1741.

[34] Carpentier, M. (1962). Contribution à l'Étude du Dispatching Économique. *Bull. de la Soc. Fran. des Élec.*, 8:431–447.

[35] Castillo, A., Laird, C., Silva-Monroy, C. A., Watson, J.-P., and ONeill, R. P. (2016a). The unit commitment problem with AC optimal power flow constraints. *IEEE Transactions on Power Systems*, 31(6):4853–4866.

[36] Castillo, A., Lipka, P., Watson, J.-P., Oren, S. S., and O'Neill, R. P. (2016b). A successive linear programming approach to solving the IV-ACOPF. *IEEE Transactions on Power Systems*, 31(4):2752–2763.

[37] Castillo, A. and O'Neill, R. (2013a). Computational Performance of Solution Techniques Applied to the ACOPF (OPF Paper 5). Technical report, US FERC.

[38] Castillo, A. and O'Neill, R. (2013b). Survey of Approaches to Solving the ACOPF (OPF Paper 4). Technical report, US Federal Energy Regulatory Commission.

[39] Chen, C., Atamtürk, A., and Oren, S. S. (2016). Bound tightening for the alternating current optimal power flow problem. *IEEE Transactions on Power Systems*, 31(5):3729–3736.

[40] Coffrin, C., Hijazi, H. L., and Van Hentenryck, P. (2015). DistFlow extensions for AC transmission systems. *arXiv preprint arXiv:1506.04773*.

[41] Coffrin, C., Hijazi, H. L., and Van Hentenryck, P. (2016). The QC relaxation: A theoretical and computational study on optimal power flow. *IEEE Transactions on Power Systems*, 31(4):3008–3018.

[42] Coffrin, C., Hijazi, H. L., and Van Hentenryck, P. (2017). Strengthening the SDP relaxation of AC power flows with convex envelopes, bound tightening, and valid inequalities. *IEEE Transactions on Power Systems*, 32(5):3549–3558.

[43] Coffrin, C. and Van Hentenryck, P. (2014). A linear-programming approximation of AC power flows. *INFORMS Journal on Computing*, 26(4):718–734.

[44] Conn, A. R., Gould, N. I., and Toint, P. (1991). A globally convergent augmented lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28(2):545–572.

[45] Cui, B. and Sun, X. A. (2018). A new voltage stability-constrained optimal power-flow model: Sufficient condition, SOCP representation, and relaxation. *IEEE Transactions on Power Systems*, 33(5):5092–5102.

[46] DallAnese, E., Baker, K., and Summers, T. (2017). Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems*, 32(5):3427–3438.

[47] Dall'Anese, E., Zhu, H., and Giannakis, G. B. (2013). Distributed optimal power flow for smart microgrids. *IEEE Transactions on Smart Grid*, 3(4):1464–1475.

[48] D'Angelo, J. and Putinar, M. (2008). Polynomial Optimization on Odd-Dimensional Spheres. In *Emerging Applications of Algebraic Geometry*. Springer New York.

[49] Dörfler, F., Simpson-Porco, J. W., and Bullo, F. (2016). Breaking the hierarchy: Distributed control and economic optimality in microgrids. *IEEE Transactions on Control of Network Systems*, 3(3):241–253.

[50] Douglas, J. and Rachford, H. H. (1956). On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, pages 421–439.

[51] Dvorkin, Y., Henneaux, P., Kirschen, D. S., and Pandzic, H. (2018). Optimizing primary response in preventive security-constrained optimal power flow. *IEEE Systems Journal*, 12(1):414–423.

[52] Eldridge, B., ONeill, R., and Castillo, A. (2018). An improved method for the DCOPF with losses. *IEEE Transactions on Power Systems*, 33(4):3779–3788.

[53] Esposito, A. G. and Ramos, E. R. (1999). Reliable load flow technique for radial distribution networks. *IEEE Transactions on Power Systems*, 14(3):1063–1069.

[54] Farivar, M., Clarke, C. R., Low, S. H., and Chandy, K. M. (2011). Inverter var control for distribution systems with renewables. In *2011 IEEE international conference on smart grid communications (SmartGridComm)*, pages 457–462.

[55] Farivar, M. and Low, S. H. (2013a). Branch flow model: Relaxations and convexification – Part I. *IEEE Transactions on Power Systems*, 28(3):2554–2564.

[56] Farivar, M. and Low, S. H. (2013b). Branch flow model: Relaxations and convexification – Part II. *IEEE Transactions on Power Systems*, 28(3):2565–2572.

[57] Fattahi, S., Ashraphijuo, M., Lavaei, J., and Atamtürk, A. (2017). Conic relaxations of the unit commitment problem. *Energy*, 134:1079–1095.

[58] Fisher, E. B., O'Neill, R. P., and Ferris, M. C. (2008). Optimal transmission switching. *IEEE Transactions on Power Systems*, 23(3):1346–1355.

[59] Fukuda, M., Kojima, M., Murota, K., and Nakata, K. (2001). Exploiting sparsity in semidefinite programming via matrix completion i: General framework. *SIAM Journal on Optimization*, 11(3):647–674.

[60] Gabay, D. and Mercier, B. (1976). A dual algorithm for the solution of non linear variational problems via finite element approximation. *Computers & Mathematics with Applications*, pages 17–40.

[61] Gan, L., Li, N., Topcu, U., and Low, S. (2012). On the exactness of convex relaxation for optimal power flow in tree networks. In *IEEE Conference on Decision and Control*, pages 465–471.

[62] Gan, L., Li, N., Topcu, U., and Low, S. H. (2013). Optimal power flow in tree networks. In *IEEE Conference on Decision and Control*, pages 2313–2318.

[63] Gan, L., Li, N., Topcu, U., and Low, S. H. (2015). Exact convex relaxation of optimal power flow in radial networks. *IEEE Transactions on Automatic Control*, 60(1):72–87.

[64] Gan, L. and Low, S. H. (2014). Optimal power flow in direct current networks. *IEEE Transactions on Power Systems*, 29(6):2892–2904.

[65] George, A. and Liu, J. W. (1981). *Computer solution of large sparse positive definite*. Prentice Hall Professional Technical Reference.

[66] Ghaddar, B., Marecek, J., and Mevissen, M. (2016). Optimal power flow as a polynomial optimization problem. *IEEE Transactions on Power Systems*, 31(1):539–546.

[67] Glowinski, R. and Marroco, A. (1975). Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(R2):41–76.

[68] Gopalakrishnan, A., Raghunathan, A. U., Nikovski, D., and Biegler, L. T. (2011). Global optimization of optimal power flow using a branch & bound algorithm. In *Annual Allerton Conference on Communication, Control, and Computing*.

[69] Guo, J., Hug, G., and Tonguz, O. K. (2017). A case for nonconvex distributed optimization in large-scale power systems. *IEEE Transactions on Power Systems*, 32(5):3842–3851.

[70] Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., and Summers, T. (2018a). Data-based distributionally robust stochastic optimal power flow, Part I: Methodologies. *IEEE Transactions on Power Systems*, pages 1483 –1492.

[71] Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., and Summers, T. (2018b). Data-based distributionally robust stochastic optimal power flow, Part II: Case studies. *IEEE Transactions on Power Systems*, pages 1493–1503.

[72] He, B. and Yuan, X. (2012). On the O(1/n) convergence rate of the Douglas–Rachford alternating direction method. *SIAM Journal on Numerical Analysis*, 50(2):700–709.

[73] Hedman, K. W., O'Neill, R. P., Fisher, E. B., and Oren, S. S. (2008). Optimal transmission switching–sensitivity analysis and extensions. *IEEE Transactions on Power Systems*, 23(3):1469–1479.

[74] Hedman, K. W., O'Neill, R. P., Fisher, E. B., and Oren, S. S. (2009). Optimal transmission switching with contingency analysis. *IEEE Transactions on Power Systems*, 24(3):1577–1586.

[75] Helmberg, C., Rendl, F., Vanderbei, R. J., and Wolkowicz, H. (1996). An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6(2):342–361.

[76] Hijazi, H., Coffrin, C., and Van Hentenryck, P. (2016). Polynomial SDP cuts for optimal power flow. In *Power Systems Computation Conference (PSCC)*, pages 1–7.

[77] Huang, S., Wu, Q., Wang, J., and Zhao, H. (2017). A sufficient condition on convex relaxation of AC optimal power flow in distribution networks. *IEEE Transactions on Power Systems*, 32(2):1359–1368.

[78] Jabr, R. (2006a). Radial Distribution Load Flow using Conic Programming. *IEEE Transactions on Power Systems*, 21:1458–1459.

[79] Jabr, R. (2008). Optimal power flow using an extended conic quadratic formulation. *IEEE Transactions on Power Systems*, 3(23):1000–1008.

[80] Jabr, R. (2012). Exploiting Sparsity in SDP Relaxations of the OPF Problem. *IEEE Transactions on Power Systems*, 27(2):1138–1139.

[81] Jabr, R. A. (2006b). Radial distribution load flow using conic programming. *IEEE Transactions on Power Systems*, 21(3):1458–1459.

[82] Jabr, R. A. and Džafić, I. (2016). A compensation-based conic OPF for weakly meshed networks. *IEEE Transactions on Power Systems*, 31(5):4167–4168.

[83] Josz, C. (2017). Counterexample to Global Convergence of DSOS and SDSOS hierarchies. `https://arxiv.org/pdf/1707.02964.pdf`.

[84] Josz, C., Fliscounakis, S., Maeght, J., and Panciatici, P. (2016). AC Power Flow Data in MATPOWER and QCQP format:

iTesla, RTE Snapshots, and PEGASE. https://arxiv.org/abs/1603.01533.
[85] Josz, C. and Henrion, D. (2015). Strong Duality in Lasserre's Hierarchy for Polynomial Optimization. *Springer Optim. Lett.*
[86] Josz, C., Maeght, J., Panciatici, P., and Gilbert, J. (2015). Application of the Moment-SOS Approach to Global Optimization of the OPF Problem. *IEEE Transactions on Power Systems*, 30(1):463–470.
[87] Josz, C. and Molzahn, D. K. (2018). Lasserre hiearchy for large scale polynomial optimization in real and complex variables. *SIAM Journal on Optimization*, pages 1017–1048.
[88] Josz, C., Molzahn, D. K., Tacchi, M., and Sojoudi, S. (2019). Transient stability analysis of power systems via occupation measures. *ISGT conference, Washington D.C.*, pages 1–6.
[89] Journée, M., Bach, F., Absil, P.-A., and Sepulchre, R. (2010). Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351.
[90] Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. In *Proc. of the annual ACM Symposium on Theory of Computing*, pages 302–311. ACM.
[91] Kekatos, V. and Giannakis, G. B. (2013). Distributed robust power system state estimation. *IEEE Transactions on Power Systems*, 28(2):1617–1626.
[92] Kheirandishfard, M., Zohrizadeh, F., and Madani, R. (2018). Convex relaxation of bilinear matrix inequalities Part I: Theoretical results. In *IEEE Conference on Decision and Control*.
[93] Kim, S., Kojima, M., Mevissen, M., and Yamashita, M. (2011). Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. *Mathematical programming*, 129(1):33–68.
[94] Kocuk, B., Dey, S. S., and Sun, X. A. (2016a). Inexactness of SDP relaxation and valid inequalities for optimal power flow. *IEEE Transactions on Power Systems*, 31(1):642–651.
[95] Kocuk, B., Dey, S. S., and Sun, X. A. (2016b). Strong SOCP relaxations for the optimal power flow problem. *Operations Research*, 64(6):1177–1196.
[96] Kocuk, B., Dey, S. S., and Sun, X. A. (2017). New formulation and strong MISOCP relaxations for AC optimal transmission switching problem. *IEEE Transactions on Power Systems*, 32(6):4161–4170.
[97] Kocuk, B., Dey, S. S., and Sun, X. A. (2018). Matrix minor reformulation and SOCP-based spatial branch-and-cut method for the AC optimal power flow problem. *Mathematical Programming*, pages 557–596.
[98] Kočvara, M. and Stingl, M. (2003). Pennon: A code for convex nonlinear and semidefinite programming. *Optimization methods and software*, 18(3):317–333.
[99] Kojima, M., Megiddo, N., Noma, T., and Yoshise, A. (1991). *A unified approach to interior point algorithms for linear complementarity problems*, volume 538. Springer Science & Business Media.
[100] Kojima, M., Mizuno, S., and Yoshise, A. (1989). A primal-dual interior point algorithm for linear programming. In *Progress in Mathematical Programming*, pages 29–47. Springer.
[101] Kuang, X., Ghaddar, B., Naoum-Sawaya, J., and Zuluaga, F. (2018). Alternative SDP and SOCP approximations for polynomial optimization. *EURO Journal on Computational Optimization*, pages 1–23.
[102] Kuang, X., Ghaddar, B., Naoum-Sawaya, J., and Zuluaga, L. F. (2017). Alternative LP and SOCP Hierarchies for ACOPF Problems. *IEEE TPS*, 32:2828–2836.
[103] LaCommare, K. H. and Eto, J. H. (1956). Cost of power interruptions to electricity consumers in the united states (u.s.). *Energy: The international Journal*.
[104] Lam, A. Y., Zhang, B., and David, N. T. (2012). Distributed algorithms for optimal power flow problem. In *Conference on Decision and Control*, pages 430–437.
[105] Lasserre, J. B. (2000). Optimisation Globale et Théorie des Moments. *C. R. Acad. Sci. Paris, Série I*, 331:929–934.
[106] Lasserre, J. B. (2001). Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11:796–817.
[107] Lasserre, J. B. (2009). *Moments, positive polynomials and their applications*, volume 1. World Scientific.
[108] Lavaei, J. and Low, S. (2010a). Convexification of optimal power flow problem. In *Annual Allerton Conference on Communication, Control, and Computing*.
[109] Lavaei, J. and Low, S. (2010b). Relationship between power loss and network topology in power systems. In *IEEE Conference on Decision and Control*.
[110] Lavaei, J. and Low, S. (2012). Zero Duality Gap in Optimal Power Flow Problem. *IEEE Transactions on Power Systems*, 27(1):92–107.
[111] Lavaei, J., Tse, D., and Zhang, B. (2014). Geometry of power flows and optimization in distribution networks. *IEEE Transactions on Power Systems*, 29(2):572–583.
[112] Lee, D., Nguyen, H. D., Dvijotham, K., and Turitsyn, K. (2018). Convex restriction of power flow feasibility set. *arXiv preprint arXiv:1803.00818*.
[113] Lehmann, K., Grastien, A., and Van Hentenryck, P. (2016). AC-feasibility on tree networks is NP-hard. *IEEE Transactions on Power Systems*, 31(1):798–801.

[114] Lesieutre, B., Molzahn, D., Borden, A., and DeMarco, C. (2011). Examining the Limits of the Application of Semidefinite Programming to Power Flow Problems. In *Annual Allerton Conference on Communication, Control, and Computing*, pages 28–30.

[115] Li, J., Liu, F., Wang, Z., Low, S. H., and Mei, S. (2018). Optimal power flow in stand-alone DC microgrids. *IEEE Transactions on Power Systems*, 33(5):5496–5506.

[116] Li, N., Chen, L., and Low, S. H. (2012). Exact convex relaxation of OPF for radial networks using branch flow model. In *IEEE SmartGridComm*, pages 7–12.

[117] Li, Q. and Vittal, V. (2017). Non-iterative enhanced SDP relaxations for optimal scheduling of distributed energy storage in distribution systems. *IEEE Transactions on Power Systems*, 32(3):1721–1732.

[118] Liberti, L. (2008). Introduction to global optimization. *Ecole Polytechnique*.

[119] Lin, J., Li, V. O., Leung, K.-C., and Lam, A. Y. (2017). Optimal power flow with power flow routers. *IEEE Transactions on Power Systems*, 32(1):531–543.

[120] Lipka, P., Oren, S. S., ONeill, R. P., and Castillo, A. (2017). Running a more complete market with the SLP-IV-ACOPF. *IEEE Transactions on Power Systems*, 32(2):1139–1148.

[121] Liu, J., Bynum, M., Castillo, A., Watson, J.-P., and Laird, C. D. (2018). A multitree approach for global solution of ACOPF problems using piecewise outer approximations. *Computers & Chemical Engineering*, 114:145–157.

[122] Liu, J., Liddell, A. C., Mareček, J., and Takáč, M. (2017). Hybrid methods in solving alternating-current optimal power flows. *IEEE Transactions on Smart Grid*, 8(6):2988–2998.

[123] Louca, R. and Bitar, E. (2018). Robust AC optimal power flow. *IEEE Transactions on Power Systems*, pages 1669–1681.

[124] Louca, R., Seiler, P., and Bitar, E. (2014). Nondegeneracy and inexactness of semidefinite relaxations of optimal power flow. *arXiv preprint arXiv:1411.4663*.

[125] Lovasz and Schrijver (1991). Cones of Matrices and Set-Functions and 01 Optimization. *SIAM J. Optim.*, 1:166–190.

[126] Low, S. H. (2014a). Convex relaxation of optimal power flow Part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27.

[127] Low, S. H. (2014b). Convex relaxation of optimal power flow Part II: Exactness. *IEEE Transactions on Control of Network Systems*, 1(2):177–189.

[128] Madani, R., Ashraphijuo, M., and Lavaei, J. (2016). Promises of conic relaxation for contingency-constrained optimal power flow problem. *IEEE Transactions on Power Systems*, 31(2):1297–1307.

[129] Madani, R., Atamturk, A., and Davoudi, A. (2017a). A scalable semidefinite relaxation approach to grid scheduling. *arXiv preprint arXiv:1707.03541*.

[130] Madani, R., Kalbat, A., and Lavaei, J. (2015a). ADMM for sparse semidefinite programming with applications to optimal power flow problem. In *IEEE Conference on Decision and Control*, pages 5932–5939.

[131] Madani, R., Kalbat, A., and Lavaei, J. (2018). A low-complexity parallelizable numerical algorithm for sparse semidefinite programming. *IEEE Transactions on Control of Network Systems*, pages 1898 – 1909.

[132] Madani, R., Lavaei, J., and Baldick, R. (2019). Convexification of power flow equations in the presence of noisy measurements. *IEEE Transactions on Automatic Control*.

[133] Madani, R., Lavaei, J., Baldick, R., and Atamtürk, A. (2017b). Power system state estimation and bad data detection by means of conic relaxation. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

[134] Madani, R., Sojoudi, S., and Lavaei, J. (2015b). Convex Relaxation for Optimal Power Flow Problem: Mesh Networks. *IEEE Transactions on Power Systems*, 30(1):199–211.

[135] Majumdar, A., Ahmadi, A. A., and Tedrake, R. (2014). Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *IEEE Conference on Decision and Control*, pages 394–401.

[136] Malick, J., Povh, J., Rendl, F., and Wiegele, A. (2009). Regularization methods for semidefinite programming. *SIAM Journal on Optimization*, 20(1):336–356.

[137] Mareček, J. and Takáč, M. (2017). A low-rank coordinate-descent algorithm for semidefinite programming relaxations of optimal power flow. *Optimization Methods and Software*, 32(4):849–871.

[138] Marley, J. F., Molzahn, D. K., and Hiskens, I. A. (2017). Solving multiperiod OPF problems using an AC-QP algorithm initialized with an SOCP relaxation. *IEEE Transactions on Power Systems*, 32(5):3538–3548.

[139] Marshall, M. (2006). Representation of non-negative polynomials with finitely many zeros. *Annales de la Faculte des Sciences Toulouse*, 15:599–609.

[140] Marshall, M. (2009). Representation of non-negative polynomials, degree bounds and applications to optimization. *Canad. J. Math.*, 61:205–221.

[141] Mehta, D., Molzahn, D. K., and Turitsyn, K. (2016a). Recent advances in computational methods for the power flow equations. In *2016 American Control Conference (ACC)*, pages 1753–1765.

[142] Mehta, D., Nguyen, H. D., and Turitsyn, K. (2016b). Numerical polynomial homotopy continuation method to locate all the power flow solutions. *IET Generation, Transmission & Distribution*, 10(12):2972–2980.

[143] Misra, S., Molzahn, D. K., and Dvijotham, K. (2018). Optimal adaptive linearizations of the AC power flow equations. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7.

[144] Molzahn, D. and Hiskens, I. (2014). Moment-Based Relaxation of the Optimal Power Flow Problem. *18th Power Syst. Comput. Conf. (PSCC)*.

[145] Molzahn, D. and Hiskens, I. (2015a). Mixed SDP/SOCP Moment Relaxations of the Optimal Power Flow Problem. In *IEEE Eindhoven PowerTech*.

[146] Molzahn, D., Josz, C., Hiskens, I., and Panciatici, P. (2015). Solution of optimal power flow problems using moment relaxations augmented with objective function penalization. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 31–38.

[147] Molzahn, D. K. and Hiskens, I. A. (2015b). Sparsity-exploiting moment-based relaxations of the optimal power flow problem. *IEEE Transactions on Power Systems*, 30(6):3168–3180.

[148] Molzahn, D. K. and Hiskens, I. A. (2016). Convex relaxations of optimal power flow problems: An illustrative example. *Transactions on Circuits and Systems I: Regular Papers*, 63(5):650–660.

[149] Molzahn, D. K., Hiskens, I. A., et al. (2019). A survey of relaxations and approximations of the power flow equations. *Foundations and Trends® in Electric Energy Systems*, 4(1-2):1–221.

[150] Molzahn, D. K., Holzer, J. T., Lesieutre, B. C., and DeMarco, C. L. (2013). Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Transactions on Power Systems*, 28(4):3987–3998.

[151] Molzahn, D. K., Josz, C., Hiskens, I. A., and Panciatici, P. (2017). A laplacian-based approach for finding near globally optimal solutions to OPF problems. *IEEE Transactions on Power Systems*, 32(1):305–315.

[152] Momoh, J., Adapa, R., and El-Hawary, M. (1999a). A Review of Selected Optimal Power Flow Literature to 1993. I. Nonlinear and Quadratic Programming Approaches. *IEEE Transactions on Power Systems*, 14(1):96–104.

[153] Momoh, J., El-Hawary, M., and Adapa, R. (1999b). A Review of Selected Optimal Power Flow Literature to 1993. II. Newton, Linear Programming and Interior Point Methods. *IEEE Transactions on Power Systems*, 14(1):105–111.

[154] Nakata, K., Fujisawa, K., Fukuda, M., Kojima, M., and Murota, K. (2003). Exploiting sparsity in semidefinite programming via matrix completion ii: Implementation and numerical results. *Mathematical Programming*, 95(2):303–327.

[155] Natarajan, K., Shi, D., and Toh, K.-C. (2013). A penalized quadratic convex reformulation method for random quadratic unconstrained binary optimization. *Optimization Online*, 6:1–26.

[156] National Academies of Sciences, Engineering, and Medicine and others (2016). *Analytic Research Foundations for the Next-Generation Electric Grid*. National Academies Press.

[157] Nesterov, Y. (2007). Smoothing technique and its applications in semidefinite optimization. *Mathematical Programming*, 110(2):245–259.

[158] Nesterov, Y. and Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*, volume 13. SIAM.

[159] Nesterov, Y. E. and Todd, M. J. (1997). Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations research*, 22(1):1–42.

[160] Nesterov, Y. E. and Todd, M. J. (1998). Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on optimization*, 8(2):324–364.

[161] Nick, M., Cherkaoui, R., Le Boudec, J.-Y., and Paolone, M. (2018). An exact convex formulation of the optimal power flow in radial distribution networks including transverse components. *IEEE Transactions on Automatic Control*, 63(3):682–697.

[162] Nie, J. (2014). Optimality Conditions and Finite Convergence of Lasserre's Hierarchy. *Mathematical Programming*, 146:97–121.

[163] ODonoghue, B., Chu, E., Parikh, N., and Boyd, S. (2016). Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068.

[164] Pandya, K. and Joshi, S. (2008). A Survey of Optimal Power Flow Methods. *Journal of Theoretical and Applied Information Technology*, 4:450–458.

[165] Park, D., Kyrillidis, A., Bhojanapalli, S., Caramanis, C., and Sanghavi, S. (2016). Provable Burer-Monteiro factorization for a class of norm-constrained matrix problems. *arXiv preprint arXiv:1606.01316*.

[166] Parrilo, P. (2000). *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, Cal. Inst. of Tech.

[167] Parrilo, P. (2003). Semidefinite Programming Relaxations for Semialgebraic Problems. *Mathematical Programming*, 96:293–320.

[168] Pataki, G. (1998). On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of operations research*, 23(2):339–358.

[169] Phan, D. T. (2012). Lagrangian duality and branch-and-bound algorithms for optimal power flow. *Operations Research*, 60(2):275–285.

[170] Povh, J., Rendl, F., and Wiegele, A. (2006). A boundary point method to solve semidefinite programs. *Computing*, 78(3):277–286.

[171] Prussing, J. (1986). The principal minor test for semidefinite matrices. *Journal of Guidance, Control, and Dynamics*, 9(1):121–122.

[172] Putinar, M. (1993). Positive Polynomials on Compact Semi-Algebraic Sets. *Indiana Univ. Math. J.*, 42:969–984.

[173] Riener, C., Theobald, T., Andrén, L. J., and Lasserre, J. B. (2013). Exploiting Symmetries in SDP-Relaxations for Polynomial Optimization. *Math. of Operations Research*, 38(1):122–141.

[174] Saad, Y. (2003). *Iterative methods for sparse linear systems*, volume 82. SIAM.

[175] Schmüdgen, K. (1991). The K-Moment Problem for Semi-Algebraic Sets. *Mathematische Annalen*, 289:203–206.

[176] Shchetinin, D., De Rubira, T. T., and Hug-Glanzmann, G. (2018). On the construction of linear approximations of line flow constraints for AC optimal power flow. *IEEE Transactions on Power Systems*, pages 1182 –1192.

[177] Sherali, H. and Adams, W. (1990). A Hierarchy of Relaxations between the Continuous and Convex Hull Representations for Zero-One Programming Problems. *SIAM J. Disc. Math.*, 3:411–430.

[178] Shor, N. (1987). Quadratic Optimization Problems. *Soviet Journal of Computer and System Sciences*, 25:1–11.

[179] Simpson-Porco, J. W. (2018). Lossy DC power flow. *IEEE Transactions on Power Systems*, 33(3):2477–2485.

[180] Sojoudi, S. and Lavaei, J. (2011). Network topologies guaranteeing zero duality gap for optimal power flow problem. *submitted for publication*.

[181] Sojoudi, S. and Lavaei, J. (2012). Physics of power networks makes hard optimization problems easy to solve. In *IEEE Power and Energy Society General Meeting*, pages 1–8.

[182] Sojoudi, S. and Lavaei, J. (2014). Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure. *SIAM Journal on Optimization*, 24(4):1746–1778.

[183] Stengle, G. (1974). A Nullstellensatz and a Positivstellensatz in Semialgebraic Geometry. *Mathematische Annalen*, 207:87–97.

[184] Stott, B., Jardim, J., and Alsaç, O. (2009). DC power flow revisited. *IEEE Transactions on Power Systems*, 24(3):1290–1300.

[185] Sturm, J. F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653.

[186] Sturm, J. F. (2002). Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. *Optimization Methods and Software*, 17(6):1105–1154.

[187] Subhonmesh, B., Low, S. H., and Chandy, K. M. (2012). Equivalence of branch flow and bus injection models. In *Annual Allerton Conference on Communication, Control, and Computing*, pages 1893–1899.

[188] Sun, Y., Andersen, M. S., and Vandenberghe, L. (2014). Decomposition in conic optimization with partially separable structure. *SIAM Journal on Optimization*, 24(2):873–897.

[189] Sundar, K., Nagarajan, H., Misra, S., Lu, M., Coffrin, C., and Bent, R. (2018). Optimization-based bound tightening using a strengthened QC-relaxation of the optimal power flow problem. *arXiv preprint arXiv:1809.04565*.

[190] Tan, C. W., Cai, D. W., and Lou, X. (2015). Resistive network optimal power flow: Uniqueness and algorithms. *IEEE Transactions on Power Systems*, 30(1):263–273.

[191] Tang, Y., Dvijotham, K., and Low, S. (2017). Real-time optimal power flow. *IEEE Transactions on Smart Grid*, 8(6):2963–2973.

[192] Toh, K.-C. and Kojima, M. (2002). Solving some large scale semidefinite programs via the conjugate residual method. *SIAM Journal on Optimization*, 12(3):669–691.

[193] Toh, K.-C., Todd, M. J., and Tütüncü, R. H. (2012). On the implementation and usage of SDPT3–a matlab software package for semidefinite-quadratic-linear programming, version 4.0. In *Handbook on semidefinite, conic and polynomial optimization*, pages 715–754. Springer.

[194] Tütüncü, R. H., Toh, K.-C., and Todd, M. J. (2003). Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical programming*, 95(2):189–217.

[195] Vandenberghe, L., Andersen, M. S., et al. (2015). Chordal graphs and semidefinite optimization. *Foundations and Trends® in Optimization*, 1(4):241–433.

[196] Vandenberghe, L. and Boyd, S. (1995). A primaldual potential reduction method for problems involving matrix inequalities. *Mathematical programming*, 69(1-3):205–236.

[197] Vandenberghe, L., Boyd, S., and Wu, S.-P. (1998). Determinant maximization with linear matrix inequality constraints. *SIAM journal on matrix analysis and applications*, 19(2):499–533.

[198] Venzke, A. and Chatzivasileiadis, S. (2018). Convex relaxations of security constrained AC optimal power flow under uncertainty. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7.

[199] Venzke, A. and Chatzivasileiadis, S. (2019). Convex relaxations of probabilistic AC optimal power flow for interconnected AC and HVDC grids. *IEEE Transactions on Power Systems*, pages 1–13.

[200] Venzke, A., Halilbasic, L., Markovic, U., Hug, G., and Chatzivasileiadis, S. (2018). Convex relaxations of chance constrained AC optimal power flow. *IEEE Transactions on Power Systems*, 33(3):2829–2841.

[201] Waki, H., Kim, S., Kojima, M., and Muramatsu, M. (2006). Sums of Squares and Semidefinite Program Relaxations for Polynomial Optimization Problems with Structured Sparsity. *SIAM J. Optim.*, 17(1):218–242.

[202] Wang, C., Cui, B., Wang, Z., and Gu, C. (2018). SDP-based optimal power flow with steady-state voltage stability constraints. *IEEE Transactions on Smart Grid*, pages 1–9.

[203] Wang, Z., Cui, B., and Wang, J. (2017). A necessary condition for power flow insolvability in power distribution systems with distributed generators. *IEEE Transactions on Power Systems*, 32(2):1440–1450.

[204] Watson, J.-P., Silva Monroy, C. A., Castillo, A., Laird, C., and O'Neill, R. (2015). Security-constrained unit commitment with linearized AC optimal power flow. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

[205] Wei, W., Wang, J., Li, N., and Mei, S. (2017). Optimal power flow of radial networks and its variations: A sequential convex optimization approach. *IEEE Transactions on Smart Grid*, 8(6):2974–2987.

[206] Wen, Y., Guo, C., Kirschen, D. S., and Dong, S. (2015). Enhanced security-constrained OPF with distributed battery energy storage. *IEEE Transactions on Power Systems*, 30(1):98–108.

[207] Wen, Z., Goldfarb, D., and Yin, W. (2010). Alternating direction augmented lagrangian methods for semidefinite programming. *Mathematical Programming Computation*, 2(3-4):203–230.

[208] Wu, D., Molzahn, D. K., Lesieutre, B. C., and Dvijotham, K. (2018a). A deterministic method to identify multiple local extrema for the AC optimal power flow problem. *IEEE Transactions on Power Systems*, 33(1):654–668.

[209] Wu, X., Conejo, A. J., and Amjady, N. (2018b). Robust security constrained ACOPF via conic programming: Identifying the worst contingencies. *IEEE Transactions on Power Systems*, 33(6):5884–5891.

[210] Ye, Y., Todd, M. J., and Mizuno, S. (1994). An $o(\sqrt{nL})$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19(1):53–67.

[211] Yu, S., Nguyen, H. D., and Turitsyn, K. S. (2015). Simple certificate of solvability of power flow equations for distribution systems. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5.

[212] Zhang, B., Lam, A. Y., Domínguez-García, A. D., and Tse, D. (2015). An optimal and distributed method for voltage regulation in power distribution systems. *IEEE Transactions on Power Systems*, 30(4):1714–1726.

[213] Zhang, B. and Tse, D. (2013). Geometry of injection regions of power networks. *IEEE Transactions on Power Systems*, 28(2):788–797.

[214] Zhang, R. Y. and Lavaei, J. (2017a). Modified interior-point method for large-and-sparse low-rank semidefinite programs. In *IEEE Conference on Decision and Control*, pages 5640–5647.

[215] Zhang, R. Y. and Lavaei, J. (2017b). Sparse semidefinite programs with guaranteed near-linear time complexity via dualized clique tree conversion. *arXiv preprint arXiv:1710.03475*.

[216] Zhang, Y., Hong, M., Dall'Anese, E., Dhople, S., and Xu, Z. (2017). Distributed controllers seeking AC optimal power flow solutions using ADMM. *IEEE Transactions on Smart Grid*, pages 4525 –4537.

[217] Zhang, Y., Madani, R., and Lavaei, J. (2018). Conic relaxations for power system state estimation with line measurements. *IEEE Transactions on Control of Network Systems*, pages 1193–1205.

[218] Zhao, X.-Y., Sun, D., and Toh, K.-C. (2010). A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM Journal on Optimization*, 20(4):1737–1765.

[219] Zheng, Y., Fantuzzi, G., Papachristodoulou, A., Goulart, P., and Wynn, A. (2017). Fast ADMM for semidefinite programs with chordal sparsity. In *Proc. of the American Control Conference*, pages 3335–3340.

[220] Zhu, H. and Giannakis, G. B. (2014). Power system nonlinear state estimation using distributed semidefinite programming. *IEEE Journal of Selected Topics in Signal Processing*, 8(6):1039–1050.

[221] Zohrizadeh, F., Kheirandishfard, M., Nasir, A., and Madani, R. (2018a). Sequential relaxation of unit commitment with AC transmission constraints. In *IEEE Conference on Decision and Control*.

[222] Zohrizadeh, F., Kheirandishfard, M., Quarm, E., and Madani, R. (2018b). Penalized parabolic relaxation for optimal power flow problem. In *IEEE Conference on Decision and Control*.