



Invited Review

A survey on conic relaxations of optimal power flow problem

Fariba Zohrizadeh^a, Cedric Josz^b, Ming Jin^c, Ramtin Madani^a, Javad Lavaei^{c,*}, Somayeh Sojoudi^c

^a The University of Texas at Arlington, USA

^b Columbia University, USA

^c University of California, Berkeley, USA

ARTICLE INFO

Article history:

Received 25 April 2019

Accepted 14 January 2020

Available online 27 January 2020

Keywords:

Optimal power flow

Semidefinite programming

Polynomial optimization

Graph theory

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 built 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 cone 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.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

The last several decades gave rise to substantial 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 (Shor, 1987). 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 (Lovasz & Schrijver, 1991; Sherali & Adams, 1990). 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 (Lasserre, 2000; 2001). This led to the development of systematic approaches for solving polynomial optimization problems to global optimality based on sum-of-squares (Parrilo, 2000; 2003). 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, 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 (Andersen & Andersen, 2000).

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 (Karmarkar, 1984), however, can be applied in the more general setting (Ben-Tal & Nemirovski, 2001; Nesterov & Nemirovskii, 1994). One significant hurdle that needed to be overcome was dealing with complementarity constraints associated with the optimality conditions of a semidefinite program. This led to several numerical implementations still used today (e.g. SeDuMi, Sturm, 1999; SDPT3, Toh, Todd, & Tütüncü, 2012). These can tackle dense problems with up to several hundred variables reliably, as well as larger instances by exploiting problem structures. 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, namely first-order

* Corresponding author.

E-mail addresses: fariba.zohrizadeh@uta.edu (F. Zohrizadeh), jinming@berkeley.edu (M. Jin), ramtin.madani@uta.edu (R. Madani), lavaei@berkeley.edu (J. Lavaei), sojoudi@berkeley.edu (S. Sojoudi).

methods. They are inspired by works on differential equations in the fifties (Douglas & Rachford, 1956) 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 systems, most notably 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 (Carpentier, 1962), and since then, due to the economic importance of OPF, substantial research efforts have been devoted to finding reliable solution methods to this problem (Cain, O'Neil, & Castillo, 2012; Low, 2014a; 2014b).

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 Baran and Wu (1989), Cain et al. (2012), and Capitanescu (2016) 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 (Bienstock & Verma, 2015; Lavaei & Low, 2012; Lehmann, Grastien, & Van Hentenryck, 2016). 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 (Bai & Wei, 2009; Capitanescu et al., 2011; Dvorkin, Henneaux, Kirschen, & Pandzic, 2018; Madani, Ashraphijuo, & Lavaei, 2016), robustness (Dörfler, Simpson-Porco, & Bullo, 2016; Louca & Bitar, 2018; Wu, Conejo, & Amjadi, 2018), energy storage (Li & Vittal, 2017; Marley, Molzahn, & Hiskens, 2017; Wen, Guo, Kirschen, & Dong, 2015), distributed platforms (Dall'Anese, Zhu, & Giannakis, 2013; Guo, Baker, Dall'Anese, Hu, & Summers, 2018a; 2018b; Lam, Zhang, & David, 2012; Watson, Silva Monroy, Castillo, Laird, & O'Neill, 2015), uncertainty of generation (Dall'Anese, Baker, & Summers, 2017; Venzke & Chatzivasileiadis, 2018; Venzke, Halilbasic, Markovic, Hug, & Chatzivasileiadis, 2018), data-driven optimal power flow (Baker, Dall'Anese, & Summers, 2016; Halilbašić, Thams, Venzke, Chatzivasileiadis, & Pinson, 2018; Mieth & Dvorkin, 2018; Zhang, Shen, & Mathieu, 2015), hybrid AC-DC grids (Bahrami, Therrien, Wong, & Jatskevich, 2017; Venzke & Chatzivasileiadis, 2019), real-time operation (Marley et al., 2017; Tang, Dvijotham, & Low, 2017), voltage stability (Cui & Sun, 2018; Wang, Cui, Wang, & Gu, 2018), discrete aspects such as transformer tap positions (Bingane, Anjos, & Le Digabel, 2018; Bingane, Anjos, & Le Digabel, 2019) and unit commitment (Castillo, Laird, Silva-Monroy, Watson, & O'Neill, 2016; Fat-tahi, Ashraphijuo, Lavaei, & Atamtürk, 2017; Lipka, Oren, O'Neill, & Castillo, 2017; Madani, Atamtürk, & Davoudi, 2017), among others (Lin, Li, Leung, & Lam, 2017). 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 (LaCommare & Eto, 2006).

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 power production (Fisher, O'Neill, & Ferris, 2008; Hedman, O'Neill, Fisher, & Oren, 2009). However, finding optimal switching strategies is known to be NP-hard in general (Hedman, O'Neill, Fisher, & Oren, 2008). 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, neural networks, ant colony optimization, genetic algorithms, fuzzy logic (see Bakirtzis, Biskas, Zoumas, & Petridis, 2002; Bouktir & Slimani, 2005; Gutierrez-Martinez, Cañizares, Fuerte-Esquivel, Pizano-Martinez, & Gu, 2010; Momoh, Adapa, & El-Hawary, 1999; Momoh, El-Hawary, & Adapa, 1999; Pandya & Joshi, 2008 and the references therein). A large body of literature has investigated linear programming-based relaxations and approximations of OPF (Alsac, Bright, Prais, & Stott, 1990; Castillo, Lipka, Watson, Oren, & O'Neill, 2016; Coffrin & Van Hentenryck, 2014; Eldridge, O'Neill, & Castillo, 2018; Misra, Molzahn, & Dvijotham, 2018; Shchetinin, De Rubira, & Hug-Glanzmann, 2018; Simpson-Porco, 2018; Stott, Jardim, & Alsaç, 2009). Local search algorithms for OPF have been studied in Wu, Molzahn, Lesieutre, and Dvijotham (2018), Bukhsh, Grothey, McKinnon, and Trodden (2013a), Mehta, Molzahn, and Turitsyn (2016a), Mehta, Nguyen, and Turitsyn (2016b), Yu, Nguyen, and Turitsyn (2015), and Wang, Cui, and Wang (2017). However, the aforementioned methods do not guarantee the recovery of globally optimal or even feasible points (Castillo & O'Neill, 2013a). In order to find globally optimal solutions, a variety of convex relaxations are studied in the literature for OPF (Bienstock & Munoz, 2014; Bingane et al., 2018; Bynum, Castillo, Watson, & Laird, 2019; Coffrin, Hijazi, & Van Hentenryck, 2015; 2016; Kuang, Ghaddar, Naoum-Sawaya, & Zuluaga, 2017; Lavaei & Low, 2012; Lee, Nguyen, Dvijotham, & Turitsyn, 2018; Low, 2014a; 2014b; Molzahn, Hiskens et al., 2019; Sundar et al., 2018; Zhang & Tse, 2013). Current industry practice for solving OPF involves non-linear optimization methods (Castillo & O'Neill, 2013b) 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 (National Academies of Sciences, Engineering, & Medicine & others, 2016). 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 (Bloomberg New Energy Finance, 2016).

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 (Lavaei & Low, 2012). The benefits and drawbacks of conic optimization for OPF have been extensively investigated in the literature (Bai, Wei, Fujisawa, & Wang, 2008; Bukhsh, Grothey, McKinnon, & Trodden, 2013b; Farivar, Clarke, Low, & Chandy, 2011; Farivar & Low, 2013a; 2013b; Jabr, 2006a; Lesieutre, Molzahn, Borden, & DeMarco, 2011; Molzahn & Hiskens, 2016),

leading to various improvements in both solution quality (Bose, Gayme, Chandy, & Low, 2015; Chen, Atamtürk, & Oren, 2016; Coffrin, Hijazi, & Van Hentenryck, 2017; 2017; Gan, Li, Topcu, & Low, 2015; Ghaddar, Marecek, & Mevissen, 2016; Josz, Maeght, Panciatici, & Gilbert, 2015; Kocuk, Dey, & Sun, 2016a; 2016a; 2016b; 2017; Liu, Bynum, Castillo, Watson, & Laird, 2018; Louca, Seiler, & Bitar, 2014; Madani, Sojoudi, & Lavaei, 2015; Zhang & Tse, 2013) and computational cost (Andersen, Hansson, & Vandenberghe, 2014; Bai & Wei, 2011; Bose, Low, Teeraratkul, & Hassibi, 2015; Jabr, 2012; Jabr & Džafić, 2016; Lam et al., 2012; Madani et al., 2016; Madani, Sojoudi et al., 2015; Molzahn, Holzer, Lesieutre, & DeMarco, 2013; Sojoudi & Lavaei, 2011; 2014). 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 efficiency¹.

We refer readers to Molzahn et al. (2019) for a thorough review of power flow equations as well as their commonly-used approximations, relaxations and restrictions. In this paper, we devote a particular focus to the application of conic programming to solving optimal power flow and the resulting hierarchy of relaxations, as well as the related numerical algorithms. Conic relaxation is proven to produce a globally optimal point for many special cases of OPF (Bose, Gayme et al., 2015; Gan et al., 2015; Gan & Low, 2014; Huang, Wu, Wang, & Zhao, 2017; Kekatos & Giannakis, 2013; Li, Liu, Wang, Low, & Mei, 2018; Madani, Lavaei, & Baldick, 2019; Madani, Lavaei, Baldick, & Atamtürk, 2017; Nick, Cherkaoui, Le Boudec, & Paolone, 2018; Tan, Cai, & Lou, 2015; Zhang, Madani, & Lavaei, 2018; Zhu & Giannakis, 2014). However, in general, an arbitrary/generic convex relaxation may fail to produce physically meaningful solutions for OPF due to NP-hardness. Despite that, in the presence of integer variables, the use of convex relaxation in place of linearization can lead to far better solutions and more efficient branch-and-bound search. Additionally, a variety of strategies are proposed in the literature to infer feasible and near-globally optimal points from inexact conic relaxations. Moreover, several papers have shown that OPF has some hidden structure that makes conic relaxations to solve the problem to global optimality in one shot for a large class of systems. For instance, branch-and-bound algorithms (Chen et al., 2016; Gopalakrishnan, Raghunathan, Nikovski, & Biegler, 2011; Phan, 2012) can be used to iteratively partition search spaces in order to tighten conic relaxations. In Madani et al. (2016); Madani, Sojoudi et al. (2015) and Natarajan, Shi, and Toh (2013), penalty terms are incorporated into the objective of convex relaxations in order to ensure the feasibility of solutions produced by conic relaxations for OPF. In Molzahn, Josz, Hiskens, and Panciatici (2015), Josz et al. (2015), and Molzahn and Hiskens (2015c), 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 methods are proposed in Wei, Wang, Li, and Mei (2017) with the aim of recovering OPF feasible points (Barati & Kargarian, 2017; Kocuk, Dey, & Sun, 2018; Wei et al., 2017; Zohrizadeh, Kheirandishfard, Quarm, & Madani, 2018). 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) (Kocuk, Dey, & Sun, 2016b; Madani, Atamtürk et al., 2017) and quadratic programming (QP) (Coffrin, Hijazi, & Van Hentenryck, 2016; Marley et al., 2017). Some papers have leveraged the sparsity of power networks to decompose

large-scale conic constraints into lower-order ones (Andersen et al., 2014; Bose, Low et al., 2015; Guo, Hug, & Tonguz, 2017; Madani et al., 2016; Molzahn et al., 2013).

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 complex polynomial optimization. Complex numbers are used in optimal power flow problem in order to quantify steady-state alternating currents. In fact, the optimal power flow problem is an instance of complex 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

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

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 $x \in \mathbb{C}^n$ and $y \in \mathbb{R}^n$, 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 , i.e., for each $k \in \{1, 2, \dots, n\}$, we have $\sqrt{\text{real}\{x_k\}^2 + \text{imag}\{x_k\}^2} \leq y_k$. 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 $\angle x$ 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 \bar{z} . The imaginary unit is represented by i .

2. Convex relaxations of optimal power flow

A variety of convex relaxations have been proposed for OPF in recent years. These are based on conic optimization such as semidefinite programming, second-order cone 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 details. Table 1 categorizes the references cited in this section.

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 $\angle v_k$ account for the amplitude and phase of the voltage, respectively. The OPF problem is most commonly formulated in polar coordinates with respect to $|v|$ and $\angle v$. However, for the purpose of devising convex

¹ <https://gocompetition.energy.gov>.

Table 1

References cited on convex relaxation of OPF.

NP-hardness of OPF	Lavaei and Low (2012), Bienstock and Verma (2015), Lehmann et al. (2016)
Semidefinite programming relaxations	Jabr (2006b), Bai et al. (2008), Jabr (2008), Lavaei and Low (2012), Sojoudi and Lavaei (2012), Coffrin et al. (2017), Madani et al. (2016), Kocuk et al. (2016b), Madani, Atamturk et al. (2017), Zohrizadeh, Kheirandishfard, Quarm et al. (2018)
Second-order cone programming	Jabr (2006b), Sojoudi and Lavaei (2012), Sojoudi and Lavaei (2014), Lavaei et al. (2014), Zhang and Tse (2013)
Quadratic programming	Coffrin et al. (2016), Zohrizadeh, Kheirandishfard, Quarm et al. (2018)
Lasserre hierarchy for OPF	Molzahn et al. (2015), Josz et al. (2015), Molzahn and Hiskens (2015c), Lasserre (2009)
Mixed-integer programming	Bienstock and Munoz (2014)
Branch-and-bound	Gopalakrishnan et al. (2011), Phan (2012), Liberti (2008)
Tree decomposition	Bai and Wei (2011), Jabr (2012), Molzahn et al. (2013), Andersen et al. (2014), Bose, Low et al. (2015), Madani et al. (2016), Guo et al. (2017), Zhang et al. (2017)
Valid inequalities	Kocuk et al. (2016b), Hijazi et al. (2016), Madani, Atamturk et al. (2017), Kocuk et al. (2018), Coffrin et al. (2016), Coffrin et al. (2017)
Exactness results	Bose et al. (2011), Sojoudi and Lavaei (2012), Li et al. (2012), Gan et al. (2012), Gan, Li, Topcu, and Low (2013), Low (2014a), Subhoni mesh et al. (2012), Lavaei and Low (2010a), Lavaei and Low (2010b), Lavaei and Low (2012), Sojoudi and Lavaei (2012), Sojoudi and Lavaei (2014), Lavaei et al. (2014), Zhang, Lam et al. (2015), Madani, Sojoudi et al. (2015)
Negative results and counter examples	Lesieutre et al. (2011), Bukhsh et al. (2013b), Molzahn and Hiskens (2016)
Objective penalization	Natarajan et al. (2013), Madani et al. (2016); Madani, Sojoudi et al. (2015), Molzahn et al. (2015), Zhang, Madani et al. (2018), Zohrizadeh, Kheirandishfard, Quarm et al. (2018), Molzahn et al. (2017), Molzahn et al. (2015), Zohrizadeh, Kheirandishfard, Nasir et al. (2018), Kheirandishfard et al. (2018)

relaxations, it is more convenient to cast OPF in rectangular coordinates so that trigonometric functions are avoided. Thus, we adopt the rectangular formulation of OPF. Readers are referred to Cain et al. (2012) for other equivalent formulations. 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 $\mathbf{C} \in \{0, 1\}^{|\mathcal{G}| \times |\mathcal{V}|}$ 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 $\vec{\mathbf{C}}, \vec{\mathbf{C}} \in \{0, 1\}^{|\mathcal{E}| \times |\mathcal{V}|}$ to be the *from* and *to* incidence matrices, respectively. We designate a direction for each line in the network. The (l, k) entry of $\vec{\mathbf{C}}$ is equal to one if and only if the line $l \in \mathcal{E}$ starts at bus k , while the (l, k) entry of $\vec{\mathbf{C}}$ is equal to one if and only if line l ends at bus k . Define $\mathbf{Y} \in \mathbb{C}^{|\mathcal{V}| \times |\mathcal{V}|}$ as the nodal admittance matrix of the power network and $\vec{\mathbf{Y}}, \vec{\mathbf{Y}} \in \mathbb{C}^{|\mathcal{E}| \times |\mathcal{V}|}$ as the *from* and *to* branch admittance matrices, respectively. Using the above notations, the OPF problem can be formulated as

$$\begin{aligned} & \text{minimize} && h(\mathbf{p}) \\ & \mathbf{p}, \mathbf{q} \in \mathbb{R}^{|\mathcal{G}|} && \\ & \mathbf{v} \in \mathbb{C}^{|\mathcal{V}|} && \end{aligned} \quad (1a)$$

$$\text{subject to} \quad \mathbf{d} + \text{diag}\{\mathbf{v}\mathbf{v}^* \mathbf{Y}^*\} = \mathbf{C}^\top(\mathbf{p} + i\mathbf{q}), \quad (1b)$$

$$|\text{diag}\{\vec{\mathbf{C}} \mathbf{v}\mathbf{v}^* \vec{\mathbf{Y}}^*\}| \leq \mathbf{f}^{\max}, \quad (1c)$$

$$|\text{diag}\{\vec{\mathbf{C}} \mathbf{v}\mathbf{v}^* \vec{\mathbf{Y}}^*\}| \leq \mathbf{f}^{\max}, \quad (1d)$$

$$\mathbf{v}^{\min} \leq \mathbf{v} \leq \mathbf{v}^{\max}, \quad (1e)$$

$$\mathbf{p}^{\min} \leq \mathbf{p} \leq \mathbf{p}^{\max}, \quad (1f)$$

$$\mathbf{q}^{\min} \leq \mathbf{q} \leq \mathbf{q}^{\max}, \quad (1g)$$

$$\angle \text{diag}\{\vec{\mathbf{C}} \mathbf{v}\mathbf{v}^* \vec{\mathbf{C}}\} \leq \theta^{\max}, \quad (1h)$$

in terms of decision variables $\mathbf{p} \triangleq [p_g]$, $\mathbf{q} \triangleq [q_g]$ and $\mathbf{v} \triangleq [v_k]$. In the above formulation, objective function $h(\mathbf{p})$ is the cost function, e.g. power transmission loss or generation cost. 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 $\mathbf{f}^{\max} \in \mathbb{R}^{|\mathcal{E}|}$ 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) and (1g) given the limits $\mathbf{v}_{\min}, \mathbf{v}_{\max} \in \mathbb{R}^{|\mathcal{V}|}$, $\mathbf{p}_{\min}, \mathbf{p}_{\max} \in \mathbb{R}^{|\mathcal{G}|}$, and $\mathbf{q}_{\min}, \mathbf{q}_{\max} \in \mathbb{R}^{|\mathcal{G}|}$. In addition, constraint (1h) is imposed to bound the phase angle differences across the lines of the network by θ_{\max} . Due to practical considerations (such as stability limits), we make the assumption that all elements of θ_{\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 (Bienstock & Verma, 2015; Lavaei & Low, 2012; Lehmann et al., 2016). Therefore, determining a global or near-global solution to 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 $\mathbf{W} \in \mathbb{H}^{|\mathcal{V}|}$, accounting for $\mathbf{v}\mathbf{v}^*$ (Bai et al., 2008; Esposito & Ramos, 1999; Jabr, 2008; Lavaei & Low, 2012). Using \mathbf{W} , variable \mathbf{v} can be dropped from the optimization problem by equivalently replacing the

consistency constraint $\mathbf{W} = \mathbf{v}\mathbf{v}^*$ with two constraints (i) $\mathbf{W} \succeq 0$ and (ii) $\text{rank}\{\mathbf{W}\} = 1$. Given that, problem (1a)–(1h) can be equivalently reformulated as:

$$\begin{aligned} & \text{minimize} && h(\mathbf{p}) \\ & \mathbf{p}, \mathbf{q} \in \mathbb{R}^{|\mathcal{N}|} \\ & \mathbf{W} \in \mathbb{H}^{|\mathcal{N}|} \end{aligned} \quad (2a)$$

$$\text{subject to} \quad \mathbf{d} + \text{diag}\{\mathbf{W}\mathbf{Y}^*\} = \mathbf{C}^\top(\mathbf{p} + i\mathbf{q}), \quad (2b)$$

$$|\text{diag}\{\vec{\mathbf{C}} \mathbf{W} \vec{\mathbf{Y}}^*\}| \leq \mathbf{f}^{\max}, \quad (2c)$$

$$|\text{diag}\{\vec{\mathbf{C}} \mathbf{W} \vec{\mathbf{Y}}^*\}| \leq \mathbf{f}^{\max}, \quad (2d)$$

$$(\mathbf{v}^{\min})^2 \leq \text{diag}\{\mathbf{W}\} \leq (\mathbf{v}^{\max})^2 \quad (2e)$$

$$\mathbf{p}^{\min} \leq \mathbf{p} \leq \mathbf{p}^{\max}, \quad (2f)$$

$$\mathbf{q}^{\min} \leq \mathbf{q} \leq \mathbf{q}^{\max}, \quad (2g)$$

$$|\text{imag}\{W_{lm}\}| \leq \text{real}\{W_{lm}\} \tan(\theta_{lm}^{\max}), \quad \forall (l, m) \in \mathcal{L} \quad (2h)$$

$$\mathbf{W} \in \mathcal{C}_{\text{OPF}}, \quad (2i)$$

where \mathcal{C}_{OPF} is

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

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 (Bai et al., 2008; Coffrin et al., 2017; Jabr, 2006b; Kocuk et al., 2016b; Lavaei & Low, 2012; Madani et al., 2016; Madani, Atamturk et al., 2017; Sojoudi & Lavaei, 2012; Zohrizadeh, Kheirandishfard, Quarm et al., 2018), quadratic programming (QP) (Coffrin et al., 2016; Zohrizadeh, Kheirandishfard, Quarm et al., 2018), and Mixed-Integer programming (MIP) (Bienstock & Munoz, 2014) relaxations of the OPF problem. The relaxations enlarge the nonconvex feasible set \mathcal{C}_{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 to obtain tight lower bounds for several practical case studies (Lavaei & Low, 2010a; 2010b; 2012). In what follows, we will discuss 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}_{OPF} with the cone of $|\mathcal{N}| \times |\mathcal{N}|$ Hermitian positive semidefinite matrices \mathcal{C}_{SDP} defined as:

$$\mathcal{C}_{\text{SDP}} \triangleq \{\mathbf{H} \in \mathbb{H}^{|\mathcal{N}|} \mid \mathbf{H} \succeq 0\}. \quad (3)$$

The SDP relaxation is exact and finds globally optimal solution of the OPF problem (1a)–(1h) if and only if it possesses a solution $(\mathbf{p}^{\text{OPT}}, \mathbf{q}^{\text{OPT}}, \mathbf{W}^{\text{OPT}})$ such that $\text{rank}\{\mathbf{W}^{\text{OPT}}\} = 1$. It has been proven in Lavaei and Low (2012) 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 \mathbf{Y} . The work (Sojoudi & Lavaei, 2012) showed that the success of the SDP relaxation is due to the passivity of the power infrastructure, which was formalized in Sojoudi and Lavaei (2014) through the notion of sign-definite weight sets. The papers (Lavaei, Tse, & Zhang, 2014; Zhang, Lam, Domínguez-García, & Tse, 2015) 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. $\theta^{\max} \in [0, 90^\circ]$). The paper (Tan et al., 2015) offers theoretical guarantees on the exactness of SOCP relaxation for resistive networks.

However, in the absence of the aforementioned conditions, it has been shown through counterexamples (Bukhsh et al., 2013b; Lesieutre et al., 2011; Molzahn & Hiskens, 2016) and NP-hardness results (Bienstock & Verma, 2015; Lehmann et al., 2016) that SDP relaxation is not exact in general. In Lesieutre et al. (2011), a three-bus OPF benchmark system is presented in which one of the lines is congested and as a consequence, SDP relaxation turns out to be inexact. In Bukhsh et al. (2013b), a number of other instances are given by imposing lower bounds on generator reactive power outputs. In such cases, the solution of SDP relaxation offers a lower bound on the unknown globally optimal cost of the OPF problem. The paper (Madani, Sojoudi et al., 2015) observed that exactness of the SDP relaxation heavily depends on the formulation of the line capacity constraints. Particularly, it has been shown that for a practical system, the feasible region associated with constraint (1h) and $|\nu_l - \nu_m| \leq \Delta_{lm}^{\max}$ are very similar in the non-convex space, however, their relaxations in terms of auxiliary variable \mathbf{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 (Madani, Sojoudi et al., 2015) suggests to impose the latter constraint to strengthen the SDP relaxation and reduce its optimality gap. The papers (Coffrin et al., 2017) proposed to amend the feasible set of the relaxed problem via the following valid inequalities for all $(k_1, k_2) \in \mathcal{L}$:

$$\begin{aligned} & v_{k_1}^\sigma v_{k_2}^\sigma \text{real}\left\{W_{k_1 k_2} e^{-i\theta_{k_1 k_2}^{\max}}\right\} - v_{k_2}^{\max} v_{k_2}^\sigma W_{k_1 k_1} - v_{k_1}^{\max} v_{k_1}^\sigma W_{k_2 k_2} \\ & \geq v_{k_1}^{\max} v_{k_2}^{\max} (v_{k_1}^{\min} v_{k_2}^{\min} - v_{k_1}^{\max} v_{k_2}^{\max}) \end{aligned} \quad (4a)$$

$$\begin{aligned} & v_{k_1}^\sigma v_{k_2}^\sigma \text{real}\left\{W_{k_1 k_2} e^{-i\theta_{k_1 k_2}^{\max}}\right\} - v_{k_2}^{\min} v_{k_2}^\sigma W_{k_1 k_1} - v_{k_1}^{\min} v_{k_1}^\sigma W_{k_2 k_2} \\ & \geq v_{k_1}^{\min} v_{k_2}^{\min} (v_{k_1}^{\max} v_{k_2}^{\max} - v_{k_1}^{\min} v_{k_2}^{\min}) \end{aligned} \quad (4b)$$

where $v_k^\sigma = v_k^{\max} + v_k^{\min}$, for $k = \{1, \dots, |\mathcal{N}|\}$. The papers (Josz et al., 2015; Molzahn et al., 2015; Molzahn & Hiskens, 2015c) propose to enhance the tightness of the relaxation by employing Lasserre hierarchy of moment relaxation (Lasserre, 2009). The first-order moment relaxation of the OPF problem in the hierarchy is equivalent to the SDP relaxation (2a)–(2i) and (3). It has been shown in Molzahn et al. (2015), Josz et al. (2015), Molzahn and Hiskens (2015c) 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 (Gopalakrishnan et al., 2011; Phan, 2012) employ a Branch-and-bound algorithm (Liberti, 2008) 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 (beyond 500 buses). This motivates researchers to alleviate the complexity of solving SDP relaxations by finding computationally cheaper alternatives such

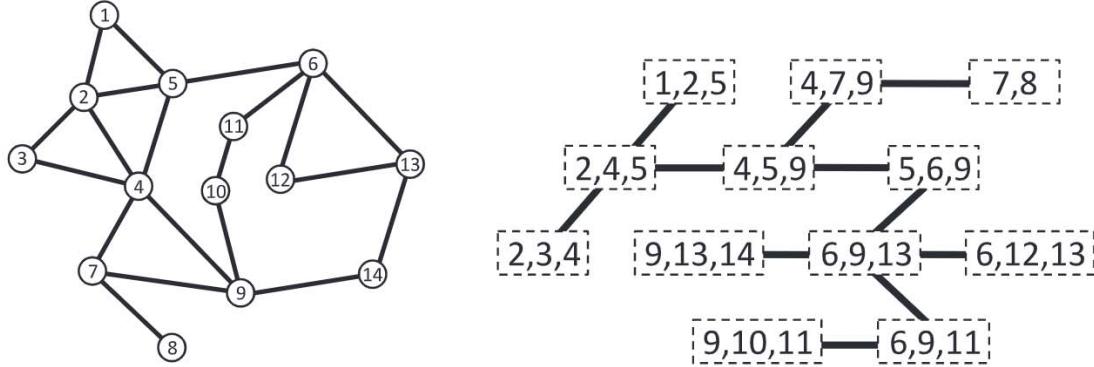


Fig. 1. The IEEE 14-bus power network (left figure) and its minimal tree decomposition (right figure).

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 (Andersen et al., 2014; Bai & Wei, 2011; Bose, Low et al., 2015; Guo et al., 2017; Jabr, 2012; Madani et al., 2016; Molzahn et al., 2013; Zhang, Hong, Dall'Anese, Dhole, & Xu, 2017). An arbitrary tree decomposition of the power network, denoted as \mathcal{B} , decomposes \mathcal{N} into several overlapping subsets $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_B \subseteq \mathcal{N}$ 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 Fig. 1. Given \mathcal{B} with B nodes, the constraint $\mathbf{W} \succeq 0$ can be represented with a set of small-sized conic constraints as:

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

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

$$\mathcal{C}_{\text{RSDP}} \triangleq \{\mathbf{H} \in \mathbb{H}^{|\mathcal{N}|} \mid \mathbf{H}\{\mathcal{B}_k\} \succeq 0, \forall k \in \{1, \dots, B\}\}. \quad (6)$$

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

Theorem 1 (Madani et al., 2016). *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 cone 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 Jabr (2006b) for acyclic networks and Sojoudi and Lavaei (2012) for general networks. The SOCP relaxation aims to enforce the relation between \mathbf{W} and \mathbf{v} through conic constraints on 2×2 principle sub-matrices of \mathbf{W} corresponding to each line $(k_1, k_2) \in \mathcal{L}$:

$$\mathbf{W}\{(k_1, k_2)\} = \begin{bmatrix} W_{k_1 k_1} & W_{k_1 k_2} \\ W_{k_2 k_1} & W_{k_2 k_2} \end{bmatrix} \succeq 0, \quad \forall (k_1, k_2) \in \mathcal{L}. \quad (7)$$

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

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

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

$$\mathcal{C}_{\text{SOCP}} \triangleq \{\mathbf{H} \in \mathbb{H}^{|\mathcal{N}|} \mid \mathbf{H}\{(k_1, k_2)\} \succeq 0, \forall (k_1, k_2) \in \mathcal{L}\}. \quad (9)$$

It is evident that the SOCP relaxation is dominated by the SDP relaxation, since constraint $\mathbf{W} \succeq 0$ ensures that every sub-matrix of \mathbf{W} , including the 2×2 principle sub-matrices, is positive semidefinite (Prussing, 1986). The SOCP relaxation is exact if and only if it possesses a solution $(\mathbf{p}^{\text{OPT}}, \mathbf{q}^{\text{OPT}}, \mathbf{W}^{\text{OPT}})$ such that: (i) $\text{rank}[\mathbf{W}^{\text{OPT}}\{(k_1, k_2)\}] = 1$ 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 grid (Lavaei et al., 2014; Sojoudi & Lavaei, 2012; 2014; Zhang & Tse, 2013). The papers (Bose, Gayme, Low, & Chandy, 2011; Sojoudi & Lavaei, 2012) prove that the SOCP relaxation is exact for radial networks if load over-satisfaction is allowed, i.e.

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

The papers (Gan, Li, Topcu, & Low, 2012; 2013; Li, Chen, & Low, 2012; Low, 2014a; Subhoni, Low, & Chandy, 2012) 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. 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 (Kocuk et al., 2016b), cutting planes (Hijazi, Coffrin, & Van Hentenryck, 2016; Kocuk et al., 2018; Madani, Atamturk et al., 2017), and convex envelope (Coffrin et al., 2016). The paper (Hijazi et al., 2016) proposed to tighten the SOCP relaxation by imposing the conic constraints on the 3×3 principle

sub-matrices of matrix variable \mathbf{W} as:

$$\mathbf{W}\{(k_1, k_2, k_3)\} = \begin{bmatrix} W_{k_1 k_1} & W_{k_1 k_2} & W_{k_1 k_3} \\ W_{k_2 k_1} & W_{k_2 k_2} & W_{k_2 k_3} \\ W_{k_3 k_1} & W_{k_3 k_2} & W_{k_3 k_3} \end{bmatrix} \succeq 0, \quad \forall (k_1, k_2, k_3) \in \bigcup_{k=1}^B \mathcal{B}_k \times \mathcal{B}_k \times \mathcal{B}_k. \quad (11)$$

where $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_B \subseteq \mathcal{N}$ are the overlapping subsets associated with an arbitrary tree decomposition of the network for which $|\mathcal{B}_k| \geq 3$, $\forall k = 1, \dots, B$. The intuition behind imposing constraint (11) is that a nonzero, Hermitian matrix $\mathbf{W} \in \mathbb{H}^{|\mathcal{N}|}$ is positive semidefinite if and only if all the principal minors of \mathbf{W} are non-negative. Based on this fact, constraint (11) is equivalent to constraint (7) together with a new constraint $\det\{\mathbf{W}\{(k_1, k_2, k_3)\}\} \geq 0$. The paper (Kocuk et al., 2018) proposed to strengthen SOCP relaxation by imposing constraints on both principal and non-principal 2×2 minors of the \mathbf{W} . These constraints are imposed based on the following characterization of rank-one positive semidefinite matrices.

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

It has been shown that constraints on non-principal 2×2 minors of \mathbf{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 (Coffrin et al., 2016) proposed a Quadratic Convex (QC) relaxation that imposes quadratic constraints in terms of the variables $|\nu_k|^2$, $|\nu_{k_1}| |\nu_{k_2}| \cos(\angle \nu_{k_1} - \angle \nu_{k_2})$, and $|\nu_{k_1}| |\nu_{k_2}| \sin(\angle \nu_{k_1} - \angle \nu_{k_2})$ to enforce the relation between \mathbf{W} and \mathbf{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 $\mathbf{W} = \mathbf{v}\mathbf{v}^*$ as:

$$\begin{aligned} W_{k,k} &= \langle |\nu_k|^2 \rangle^T, & \forall k \in \mathcal{N} \\ \text{real}\{W_{k_1 k_2}\} &= \langle \langle |\nu_{k_1}| |\nu_{k_2}| \rangle^M \langle \cos(\angle \nu_{k_1} - \angle \nu_{k_2}) \rangle^C \rangle^M & \forall (k_1, k_2) \in \mathcal{L} \\ \text{imag}\{W_{k_1 k_2}\} &= \langle \langle |\nu_{k_1}| |\nu_{k_2}| \rangle^M \langle \sin(\angle \nu_{k_1} - \angle \nu_{k_2}) \rangle^C \rangle^M & \forall (k_1, k_2) \in \mathcal{L} \end{aligned} \quad (12)$$

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

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

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

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). In general, the exactness of all such relaxations can be verified by checking the consistency equalities:

$$W_{k_1 k_2}^{\text{OPT}} = \nu_{k_1}^{\text{OPT}} (\nu_{k_2}^{\text{OPT}})^* \quad \forall (k_1, k_2) \in \mathcal{L}. \quad (15)$$

The paper (Zohrizadeh, Kheirandishfard, Quarm et al., 2018) 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 (Bienstock & Munoz, 2014) develops a linear relaxation of the rectangular formulation of the OPF problem in which the non-linearity and non-convexity come from the bi-linear equations in terms of the real and imaginary parts of the complex voltages. Let us define each ν_k at bus k as $\nu_k = e_k + i f_k$ where e_k and f_k are bounded variables corresponding to the real and imaginary parts of the ν_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:

$$u = \sum_{j=1}^T 2^{-j} y_j + \delta, \quad (16)$$

where $y_j \in \{0, 1\}$, $T \geq 1$, and $0 \leq \delta \leq 2^{-T}$. Given (16), the bi-linear term uv can be relaxed as

$$\sum_{j=1}^T 2^{-j} w_j \leq uv \leq \sum_{j=1}^T 2^{-j} w_j + 2^{-T} v, \quad (17)$$

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:

$$\begin{aligned} w_j &\leq \min\{v, y_j\}, \\ w_j &\geq \max\{v + y_j - 1, 0\}. \end{aligned} \quad (18)$$

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 (Madani et al., 2016; Madani, Sojoudi et al., 2015; Molzahn et al., 2015; Molzahn, Josz, Hiskens, & Panciatici, 2017; Natarajan et al., 2013; Zhang, Madani et al., 2018; Zohrizadeh, Kheirandishfard, Quarm et al., 2018). The paper (Madani, Sojoudi et al., 2015) proposed to penalize the total reactive power generation of the network by replacing objective function (2a) with:

$$h(\mathbf{p}) + \varepsilon_b \sum_{g \in \mathcal{G}} q_g, \quad (19)$$

where $\varepsilon_b > 0$ is the penalty coefficient. The paper (Madani et al., 2016) proposed to penalize the apparent power loss over the series impedance of the lines of the network using the following objective function:

$$h(\mathbf{p}) + \varepsilon_b \sum_{g \in \mathcal{G}} q_g + \varepsilon_l \sum_{(l, m) \in \mathcal{L}} \left| W_{ll} + W_{mm} - W_{lm} - W_{ml} \right| |y_{lm}^*|, \quad (20)$$

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

Table 2

References cited on convex relaxation of OPF.

Motivation for stronger relaxations of OPF	Lesieutre et al. (2011)
Polynomial optimization theory	Lasserre (2000, 2001), Parrilo (2000, 2003), Putinar (1993), Schmüdgen (1991), Marshall (2006, 2009), Stengle (1974)
Exploiting symmetry in polynomial optimization	D'Angelo and Putinar (2008), Waki et al. (2006), Riener et al. (2013)
Application of polynomial optimization to OPF	Molzahn and Hiskens (2014), Josz and Molzahn (2018), Josz et al. (2015), Molzahn and Hiskens (2015c), Ghaddar et al. (2016)

matrix \mathbf{W} , which indeed promotes low-rank solutions. Given \mathbf{W}^{OPT} of the penalized SDP relaxation, an approximate feasible solution \mathbf{v} of problem (1a)–(1g) is obtained by setting magnitude v_k equal to the square root of W_{kk}^{OPT} , for all $k \in \{1, \dots, |\mathcal{N}|\}$ and finding the phases of the entries of \mathbf{v} by minimizing the following optimization problem:

$$\underset{\mathbf{v} \in \mathbb{C}^{|\mathcal{N}|}}{\text{minimize}} \quad \sum_{(l,m) \in \mathcal{L}} |\angle W_{lm}^{\text{OPT}} - \angle v_l - \angle v_m| \quad (21a)$$

$$\text{subject to} \quad |\angle v_k| \leq \pi, \quad \forall k \in \mathcal{N}, \quad (21b)$$

$$\angle v_1 = 0. \quad (21c)$$

Note that if matrix \mathbf{W}^{OPT} has a dominant nonzero eigenvalue, problem (21a)–(21c) seeks to find vector \mathbf{v} such that its line angle differences are as close as possible to those suggested by \mathbf{W}^{OPT} . The papers (Zohrizadeh, Kheirandishfard, Nasir, & Madani, 2018; Zohrizadeh, Kheirandishfard, Quarm et al., 2018) 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 Kheirandishfard, Zohrizadeh, and Madani (2018), Madani, Kheirandishfard, Lavaei, and Atamturk (2018) that if the initial point is sufficiently close to the feasible set of the non-convex OPF problem (1a)–(1h) and satisfies a generalized linear independence constraint qualification (LICQ) condition, the penalized relaxations produce feasible solutions for the OPF problem. More precisely, penalization is guaranteed to result in a feasible point, if the initial distance from feasible set is smaller than $(2P)^{-1}S$, where S is the smallest singular value of the Jacobian matrix and P is a constant value regarded as pencil norm.

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 Lesieutre et al. (2011), in the LMBM3 3-bus system (named after its authors Lesieutre, Molzahn, Borden, and Demarco) of Fig. 2, the rank constraint is not always satisfied at optimality in the Shor relaxation of the optimal power flow as proposed in Lavaei and Low (2012). For instance, when the upper limit on the power flow on line 2–3 is equal to 48 megavolt amperes, the Shor relaxation provides a lower bound equal to 5819.02 dollar per hour. Using sum-of-squares of degree 2, one can construct a stronger convex relaxation whose optimal value is 5882.67 dollar per hour (Ghaddar et al., 2016; Josz et al., 2015; Molzahn & Hiskens, 2015c). 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 (Table 2).

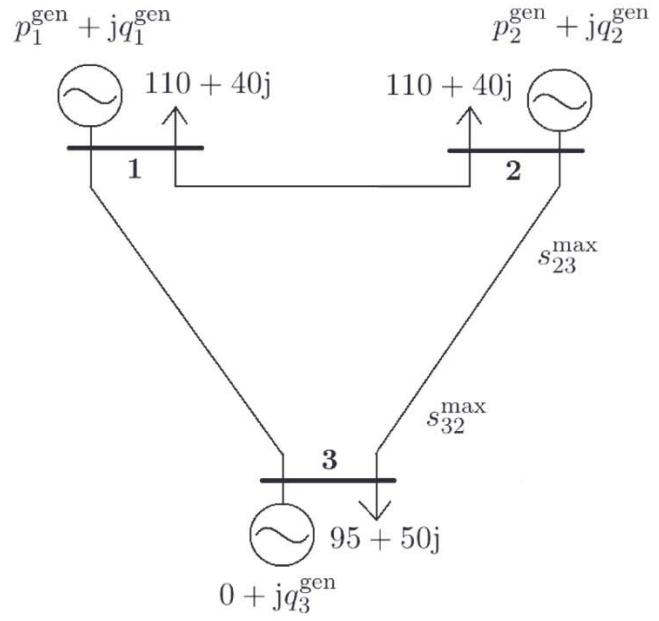


Fig. 2. LMBM3 3-bus system (Lesieutre et al., 2011).

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. The advantage of casting the optimal power flow as a polynomial optimization problem (as is performed next) is that it allows one to construct stronger convex relaxations. As a result, one can find global solutions to a larger class of benchmarks. Theoretically, all benchmarks instances could be solved with this approach, but some are too computationally expensive.

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 illustrate some of its main features and provide numerical examples. Note that all the constraints and objective of the optimal power flow are polynomial functions of degree 2 of the real and imaginary parts of the voltages. In other words, they can be written in the form

$$f(\mathbf{x}) \triangleq \sum_{|\alpha| \leq 2} f_\alpha \mathbf{x}^\alpha \quad (22)$$

where each monomial \mathbf{x}^α has coefficient f_α . A monomial is defined by

$$\mathbf{x}^\alpha := \text{real}\{\mathbf{v}_1\}^{\alpha_1} \dots \text{real}\{\mathbf{v}_{|\mathcal{N}|}\}^{\alpha_{|\mathcal{N}|}} \text{imag}\{\mathbf{v}_1\}^{\alpha_{|\mathcal{N}|+1}} \dots \text{imag}\{\mathbf{v}_{|\mathcal{N}|}\}^{\alpha_{2|\mathcal{N}|}}$$

where $\alpha_1, \dots, \alpha_{2|\mathcal{N}|}$ are integers such that $|\alpha| := \sum_{k=1}^{2|\mathcal{N}|} \alpha_k \leq 2$.

With these notations, one can define the Riesz functional

$$L_y(f) \triangleq \sum_{|\alpha| \leq 2} f_\alpha y_\alpha \quad (23)$$

where to each monomial \mathbf{x}^α we associate a lifted variable $y_\alpha \in \mathbb{R}$. The next step of the procedure is to define a vector

$$\begin{aligned} \mathbf{x}_d = & [1 \ \text{real}\{\nu_1\} \ \dots \text{real}\{\nu_{|\mathcal{N}|}\} \ \text{imag}\{\nu_1\} \ \dots \text{imag}\{\nu_{|\mathcal{N}|}\} \ \dots \\ & \text{real}\{\nu_1\}^2 \ \text{real}\{\nu_1\}\text{real}\{\nu_2\} \ \dots \text{real}\{\nu_1\}\text{imag}\{\nu_{|\mathcal{N}|}\} \ \dots \\ & \text{real}\{\nu_1\}^d \ \dots \ \text{imag}\{\nu_{|\mathcal{N}|}\}^d]^T \end{aligned} \quad (24)$$

containing monomials up to a certain degree d of the real and imaginary parts of the voltages. In order to define the moment relaxation of d , we define the moment matrix as $M_d\{y\} \triangleq L_y\{\mathbf{x}_d \mathbf{x}_d^T\}$ (see [Molzahn & Hiskens, 2015b](#) for example). Also, to each constraint $c(\nu) \geq 0$ of the optimal power flow of degree less than or equal to $2d'$, we define a localizing matrix defined as

$$\mathbf{M}_{d-d'}\{cy\} = L_{d-d'}\{c(\nu)\mathbf{x}_{d-d'}\mathbf{x}_{d-d'}^T\}. \quad (25)$$

With these definitions, we can define the moment relaxation of order d as follows where the variables are $y_\alpha \in \mathbb{R}$ for $|\alpha| \leq 2d$:

$$\text{minimize} \quad L_y(\mathbf{p}(\nu)) \quad (26a)$$

$$\text{subject to} \quad y_0 = 1, \quad (26b)$$

$$\mathbf{M}_d\{y\} \succeq 0, \quad (26c)$$

$$\mathbf{M}_{d-1}\{\mathbf{f}_k^{\max} - |\text{diag}\{\tilde{\mathbf{C}} \ \mathbf{v}\mathbf{v}^* \ \tilde{\mathbf{C}}^*\}\}_k\} \succeq 0, \quad k \in \mathcal{N}, \quad (26d)$$

$$\mathbf{M}_{d-1}\{|\mathbf{v}|_k^2 - (\mathbf{v}_k^{\min})^2\} \succeq 0, \quad k \in \mathcal{N}, \quad (26e)$$

$$\mathbf{M}_{d-1}\{(\mathbf{v}_k^{\max})^2 - |\mathbf{v}|_k^2\} \succeq 0, \quad k \in \mathcal{N}, \quad (26f)$$

$$\mathbf{M}_{d-1}\{\mathbf{p}_k(\nu) - \mathbf{p}_k^{\min}\} \succeq 0, \quad k \in \mathcal{N}, \quad (26g)$$

$$\mathbf{M}_{d-1}\{\mathbf{p}_k^{\max} - \mathbf{p}_k(\nu)\} \succeq 0, \quad k \in \mathcal{N}, \quad (26h)$$

$$\mathbf{M}_{d-1}\{\mathbf{q}_k(\nu) - \mathbf{q}_k^{\min}\} \succeq 0, \quad k \in \mathcal{N}, \quad (26i)$$

$$\mathbf{M}_{d-1}\{\mathbf{q}_k^{\max} - \mathbf{q}_k(\nu)\} \succeq 0, \quad k \in \mathcal{N}, \quad (26j)$$

$$\mathbf{M}_{d-1}\{\theta_{ij}^{\max} - \angle\text{diag}\{\tilde{\mathbf{C}} \ \mathbf{v}\mathbf{v}^* \ \tilde{\mathbf{C}}\}_{ij}\} \succeq 0, \quad (i, j) \in \mathcal{L}. \quad (26k)$$

Each non-linear optimal power flow constraint has been transformed into a conic constraint. At the first order of the hierarchy, it corresponds to a linear constraint; at high-orders it corresponds to a semidefinite constraint. In 2001, the [Lasserre \(2000, 2001\)](#) hierarchy (see also [Parrilo, 2000; Parrilo, 2003](#)) 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 ([Putinar, 1993](#)) proven in 1993. The assumption is Putinar's theorem is satisfied for the optimal power flow due to the constraints $\text{real}\{\nu_k\}^2 + \text{imag}\{\nu_k\}^2 \leq \mathbf{v}_k^{\max}, k \in \mathcal{N}$. Hence 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 ([Josz & Henrion, 2016](#)). 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.

We now focus on the result that justifies the global convergence of the Lasserre hierarchy, namely [Putinar's \(1993\)](#) Positivstellensatz. This result provided a crucial refinement of [Schmüdgen's \(1991\)](#) Positivstellensatz 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 can be seen in the theorem below.

Theorem 2 ([Putinar's, 1993](#) Positivstellensatz). *Assume that there exist $R > 0$ and sums-of-squares p_0, \dots, p_m such that*

$$R^2 - x_1^2 - \dots - x_n^2 = p_0 + \sum_{i=1}^m p_i g_i \quad (27)$$

If $f > 0$ on $\{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0 \dots g_m(\mathbf{x}) \geq 0\}$, then there exists sums-of-squares $\sigma_0, \dots, \sigma_m$ such that

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

Based on this theorem, it is simple to see that whenever λ is a strict lower bound of f on the feasible set $\{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0 \dots g_m(\mathbf{x}) \geq 0\}$, 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 [Josz \(2017\)](#):

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

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

$$x_1^2 + x_2^2 + 2x_1x_2 - 4x_1 - 4x_2 - (2 - 4\sqrt{2}) =$$

$$(\sqrt{2} - 1)(x_1 - x_2)^2 + \sqrt{2}(-\sqrt{2} + x_1 + x_2)^2$$

$$+ 2(\sqrt{2} - 1)(1 - x_1^2 - x_2^2)$$

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, the Lasserre hierarchy has finite convergence generically. This means that for a given arbitrary polynomial optimization problem, finite convergence will almost surely hold. It was [Nie \(2014\)](#) who proved this result, which had been observed in practice ever since the Lasserre hierarchy was introduced. He relied on theorems of [Marshall \(2006, 2009\)](#) 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 \(2000\)](#) 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 (2003) work panders to Stengle's (1974) Positivstellensatz, 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. In the case of optimal power flow, the upper bounds on voltage suffice to guarantee convergence of the Lasserre hierarchy and often a low order suffices to reach global optimality.

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, Kim, Kojima, and Muramatsu (2006) 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 (Ghaddar et al., 2016) (instead of 20 without exploiting sparsity Molzahn & Hiskens, 2014). However, by using the correlative sparsity graph discussed above, a lot of the sparsity is lost. The authors of Josz and Molzahn (2018) 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:

$$\begin{aligned} v_k i_k^* &= v_k \left(\sum_l i_{kl} \right)^* = v_k \left(\sum_l y_{kl} (v_k - v_l) \right)^* \\ &= \left(\sum_l y_{kl}^* \right) |v_k|^2 - \sum_l \bar{y}_{kl} v_k v_l^* \end{aligned} \quad (29)$$

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

$$2v_1v_1^* - (1+i)v_1v_2^* - (2-i)v_1v_3^* - (4+3i)v_1v_4^* = 1 - 3i \quad (30)$$

where the constant terms are chosen arbitrarily. Notice that the sparsity pattern associated to this constraint is exactly that the one described in Fig. 3. The reason why the monomial sparsity and correlative sparsity differ is due in part to the Kirchhoff's law, where 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 improved tractability. In addition, monomial sparsity allows one to apply different relaxation orders at each constraint.

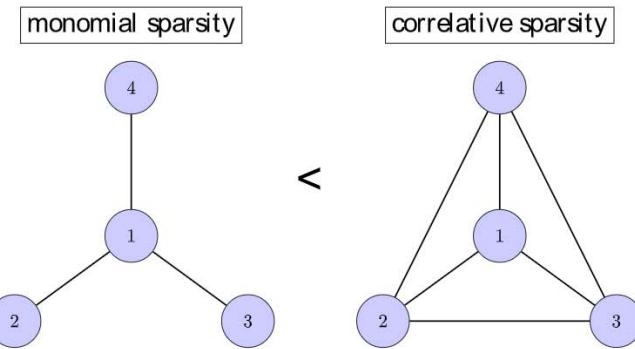


Fig. 3. Two different notions of sparsity.

3.3. Exploiting symmetry

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

$$\underset{\mathbf{v} \in \mathbb{C}^n}{\text{minimize}} \quad f(\mathbf{v}, \mathbf{v}^*) \triangleq \sum_{\alpha, \beta} f_{\alpha, \beta} \mathbf{v}^\alpha (\mathbf{v}^*)^\beta \quad (31a)$$

$$\text{subject to} \quad g_i(\mathbf{v}, \mathbf{v}^*) \triangleq \sum_{\alpha, \beta} g_{i, \alpha, \beta} \mathbf{v}^\alpha (\mathbf{v}^*)^\beta \geq 0, \quad i = 1, \dots, m. \quad (31b)$$

We again use the multi-index notation $\mathbf{v}^\alpha := v_1^{\alpha_1} \dots v_n^{\alpha_n}$, $\alpha \in \mathbb{N}^n$. The functions f, g_1, \dots, 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 $f_{\alpha, \beta}^* = f_{\beta, \alpha}$ and $g_{i, \alpha, \beta}^* = g_{i, \beta, \alpha}$. The feasible set is defined as $K := \{\mathbf{v} \in \mathbb{C}^n : g_i(\mathbf{v}, \mathbf{v}^*) \geq 0, i = 1, \dots, m\}$.

Example 2. In Bakhsh et al. (2013a, 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

$$\underset{\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}}{\text{minimize}} \quad 8|\mathbf{v}_1 - \mathbf{v}_2|^2 \quad (32a)$$

$$\text{subject to} \quad 0.9025 \leq |\mathbf{v}_1|^2 \leq 1.1025 \quad (32b)$$

$$0.9025 \leq |\mathbf{v}_2|^2 \leq 1.0568 \quad (32c)$$

$$(2 + 10i)\mathbf{v}_1\mathbf{v}_2^* + (2 - 10i)\mathbf{v}_2\mathbf{v}_1^* - 4|\mathbf{v}_2|^2 = 350 \quad (32d)$$

$$(-10 + 2i)\mathbf{v}_1\mathbf{v}_2^* + (-10 - 2i)\mathbf{v}_2\mathbf{v}_1^* + 20|\mathbf{v}_2|^2 = -350 \quad (32e)$$

Notice that if $(\mathbf{v}_1, \mathbf{v}_2)$ is a feasible point, then so is $(e^{i\theta}\mathbf{v}_1, e^{i\theta}\mathbf{v}_2)$ for all $\theta \in \mathbb{R}$ (this is due to the absence of any constraints on the angle of slack bus in our formulation, which is generally regarded as the global phase ambiguity). When converted to real numbers $\mathbf{v}_1 \triangleq x_1 + x_3i$ and $\mathbf{v}_2 \triangleq x_2 + x_4i$, it yields

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

$$\text{subject to} \quad 0.9025 \leq x_1^2 + x_3^2 \leq 1.1025 \quad (33b)$$

$$0.9025 \leq x_2^2 + x_4^2 \leq 1.0568 \quad (33c)$$

$$4x_1x_2 + 4x_3x_4 + 20x_1x_4 - 20x_3x_2 - 4x_2^2 + 4x_4^2 = 350 \quad (33d)$$

$$-20x_1x_2 - 20x_3x_4 + 4x_1x_4 - 4x_3x_2 + 20x_2^2 + 20x_4^2 = -350 \quad (33e)$$

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 (2008) Positivstellensatz, 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 megawatt 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 Riener, Theobald, An-dréan, and Lasserre (2013).

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 Majumdar, Ahmadi, and Tedrake (2014) and Ahmadi and Majumdar (2019), a polynomial $\sigma(x) = \sum_{|\alpha| \leq 2d} \sigma_\alpha x^\alpha$ is a scaled diagonally-dominant sum-of-squares (SDSOS) if it is of the form

$$\sigma(x) = \sum_k (p_k x^{\alpha_k} + q_k x^{\beta_k})^2 \quad \text{where} \quad \alpha_k, \beta_k \in \mathbb{N}^n, \quad p_k, q_k \in \mathbb{R}. \quad (34)$$

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

$$\sum_{\substack{|\alpha|, |\beta| \leq d \\ \alpha \neq \beta}} \mathbf{N}^{\alpha, \beta} \quad (35)$$

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

$$\begin{pmatrix} \mathbf{N}^{\alpha, \beta}(\alpha, \alpha) & \mathbf{N}^{\alpha, \beta}(\alpha, \beta) \\ \mathbf{N}^{\alpha, \beta}(\beta, \alpha) & \mathbf{N}^{\alpha, \beta}(\beta, \beta) \end{pmatrix} \succeq 0. \quad (36)$$

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 $\sigma_0, \dots, \sigma_m$ to be SDSOS, then in the moment problem of order d , we are relaxing each semidefinite constraint as follows

$$\begin{pmatrix} \mathbf{M}_d(y)(\alpha, \alpha) & \mathbf{M}_d(y)(\alpha, \beta) \\ \mathbf{M}_d(y)(\beta, \alpha) & \mathbf{M}_d(y)(\beta, \beta) \end{pmatrix} \succeq 0 \quad \forall |\alpha|, |\beta| \leq d, \quad \alpha \neq \beta, \quad (37)$$

$$\begin{pmatrix} \mathbf{M}_{d-k_i}(g_i y)(\alpha, \alpha) & \mathbf{M}_{d-k_i}(g_i y)(\alpha, \beta) \\ \mathbf{M}_{d-k_i}(g_i y)(\beta, \alpha) & \mathbf{M}_{d-k_i}(g_i y)(\beta, \beta) \end{pmatrix} \succeq 0 \quad \forall |\alpha|, |\beta| \leq d - k_i, \quad \alpha \neq \beta. \quad (38)$$

Naturally, these can also be viewed as second-order conic constraints. For their application to polynomial optimization, see Kuang, Ghaddar, Naoum-Sawaya, and Zuluaga (2018), QM-DSOSr,

QM-SDSOSr) and Kuang et al. (2017). In particular, they have been applied to the optimal power flow problem, enabling one to compute lower bound for medium-sized test cases (several hundreds of variables). In a similar flavor, in Molzahn and Hiskens (2015a), 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. Such are the conclusions in Molzahn and Hiskens (2015a), and there is no clear explanation for when it is beneficial to use second-order conic constraints. It is known that using the cheaper second-order conic constraints does not guarantee convergence, even in the case of a convex optimization problem with two variables (Josz, 2017; Kurpisz & de Wolff, 2019; Zhang, Josz, & Sojoudi, 2018). Specifically, one could use second-order cone programming in the moment relaxation of the optimal power flow of order d by replacing the constraints (26c)–(26k) with constraints that enforce all 2×2 submatrices to be positive semidefinite.

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 cone 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 alternating direction method of multipliers (ADMM). Next, we explain the reasoning behind interior-point methods which were originally designed for linear programming. Last, we discuss non-convex 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:

$$\begin{aligned} & \text{minimize} && \mathbf{C} \bullet \mathbf{X} \\ & \mathbf{X} \in \mathbb{S}^n \end{aligned} \quad (39a)$$

$$\text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{b} \quad (39b)$$

$$\mathbf{X} \succeq 0, \quad (39c)$$

where $\mathbf{C} \in \mathbb{S}^n$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathcal{A}(\mathbf{X}) = [\mathbf{A}_1 \bullet \mathbf{X} \quad \dots \quad \mathbf{A}_m \bullet \mathbf{X}]^\top$ is the linear matrix operator. The dual problem is given by:

$$\begin{aligned} & \text{maximize} && \mathbf{b}^\top \mathbf{y} \\ & \mathbf{y} \in \mathbb{R}^m, \mathbf{S} \in \mathbb{S}^n \end{aligned} \quad (40a)$$

$$\text{subject to} \quad \mathcal{A}^*(\mathbf{y}) + \mathbf{S} = \mathbf{C} \quad (40b)$$

$$\mathbf{S} \succeq 0, \quad (40c)$$

where $\mathcal{A}^*(\mathbf{y}) = \sum_{i=1}^m y_i \mathbf{A}_i$ 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 (Alizadeh, Haeberly, & Overton, 1997).

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

1. (Linear independence) The matrix $\mathbf{A} = [\text{vec}\{\mathbf{A}_1\} \dots \text{vec}\{\mathbf{A}_m\}]$ has full column-rank (i.e., $\mathbf{A}^\top \mathbf{A}$ is nonsingular);
2. (Slater's condition) There exist \mathbf{y} and $\mathbf{X}, \mathbf{S} > 0$ such that $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ and $\mathcal{A}^*(\mathbf{y}) + \mathbf{S} = \mathbf{C}$ 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 (Ye, Todd, & Mizuno, 1994), 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 $\mathbf{A}_i \bullet \mathbf{X} \geq b_i$, we can introduce $\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0}^\top & x \end{bmatrix}$ that extends the dimension of \mathbf{X} by 1, and introduce an extra equality equation that constrains the element $x = \mathbf{A}_i \bullet \mathbf{X} - b_i$. In addition, all of the complexity bounds discussed below can be generalized to conic programs posed on the Cartesian product of multiple semidefinite cones $\mathcal{K} = \mathbf{S}_+^{n_1} \times \mathbf{S}_+^{n_2} \times \dots \times \mathbf{S}_+^{n_\ell}$.

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 (Gabay & Mercier, 1976; Glowinski & Marroco, 1975). 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. (2011) 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 (2010) and O'Donoghue, Chu, Parikh, and Boyd (2016). Following the treatment of Wen et al. (2010), we start by defining the augmented Lagrangian function for the dual problem:

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

where $\mathbf{X} \in \mathbb{S}^n$ and $\mu > 0$ is the penalty coefficient. By adding the quadratic term to the Lagrangian function, the augmented Lagrangian $\mathcal{L}_\mu(\mathbf{X}, \mathbf{y}, \mathbf{S})$ is strongly convex in \mathbf{y} and \mathbf{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 \mathbf{X}^0 , the augmented Lagrangian method solves in each iteration

$$\begin{aligned} & \underset{\mathbf{y} \in \mathbb{R}^m, \mathbf{S} \in \mathbb{S}^n}{\text{minimize}} \quad \mathcal{L}_\mu(\mathbf{X}^k, \mathbf{y}, \mathbf{S}) \quad \text{subject to} \quad \mathbf{S} \succeq 0, \\ & \end{aligned} \quad (42)$$

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

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

Instead of jointly minimizing $\mathcal{L}_\mu(\mathbf{X}^k, \mathbf{y}, \mathbf{S})$ with respect to \mathbf{y} and \mathbf{S} , which can be very time consuming for large-scale problems, ADMM minimizes the augmented Lagrangian function with respect to \mathbf{y} and \mathbf{S} separately, and hence no longer solving problem (42) exactly (Wen et al., 2010):

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

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

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 $\mu > 0$ (Wen et al., 2010, Thm 2). Here, the \mathbf{y} -update has a closed-form solution based on first-order optimality of the unconstrained problem

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

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

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

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

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

where $\{\mathbf{v}_i \in \mathbb{C}^n\}_{i=1}^n$ and $\{\sigma_i \in \mathbb{R}\}_{i=1}^n$ denote the sets of eigenvectors and eigenvalues of \mathbf{V} , respectively. The above iteration is closely related to the regularization method (Malick, Povh, Rendl, & Wiegele, 2009) and the boundary point method (Povh, Rendl, & Wiegele, 2006), where \mathbf{X}^k is fixed until $\frac{1}{\mu}(\mathbf{S}^{k+1} - \mathbf{V}^{k+1})$ 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 Wen et al., 2010, Sec. 3.2; Boyd et al., 2011, Sec. 3.4 for implementation details).

As discussed in Section 2, due to the sparsity of network topologies, the capability of solving large-scale SDP relaxations of OPF can be enhanced by exploiting the sparsity pattern and parallelization (Fukuda, Kojima, Murota, & Nakata, 2001; Madani, Kalbat, & Lavaei, 2018; Sun, Andersen, & Vandenberghe, 2014). A fundamental difficulty of SDP is induced by the positive semidefinite constraint, $\mathbf{X} \succeq 0$, which densely couples all $O(n^2)$ elements of the matrix \mathbf{X} . The complexity comes from solving the linear systems known as the normal equation or the Schur complement equation, which is fully-dense. 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 \mathbf{X} with the submatrices $\mathbf{X}[\mathcal{B}_k]$ 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 \mathbf{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 Madani, Kalbat et al., 2018, Sec. IV-B for the implementation). The idea of tree decomposition of a sparse SDP into smaller sized SDPs is first proposed in Fukuda et al. (2001), which then solves the subproblems by interior point methods. Using a first-order splitting method, Sun et al. (2014) solves the decomposed SDP problem created by Fukuda et al. (2001), but the algorithm needs to solve an optimization subproblem at every iteration. The employment of ADMM to solve the reduced problem is proposed in Madani, Kalbat, and Lavaei (2015) and Madani, Kalbat et al. (2018). 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 (Josz, Fliscounakis, Maeght, & Panciatici, 2016) amounts to simple operations over 857 matrices of size 31 by 31 and 14,035 matrices of size 2 by 2).

As for the convergence rate of ADMM, it has been shown in He and Yuan (2012) 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 (Wen et al., 2010; Zheng, Fantuzzi, Papachristodoulou, Goulart, & Wynn, 2017). The sparse structure that arises from OPF can be further exploited in the ADMM updates. The y -update is dominated by the solution of $(\mathbf{A}^\top \mathbf{A})\mathbf{y} = \mathbf{r}$, which has the worst-case complexity of $O(n^6)$ time and $O(n^4)$ memory using standard Cholesky factorization for fully dense matrix \mathbf{A} . However, for large and sparse \mathbf{A} , efficiency can be substantially improved by using an incomplete factorization as the preconditioner within an iterative solution algorithm like conjugate gradients (Saad, 2003), and by reordering the columns of \mathbf{A} using a fill-minimizing ordering like minimum degree and nested dissection (George & Liu, 1981). In addition to SDP reduction via tree decomposition, the \mathbf{S} -update can be further improved when the optimal solution \mathbf{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 (Nesterov, 2007), augmented Lagrangian methods (Kočvara & Stingl, 2003), conjugate gradients (Toh & Kojima, 2002; Zhao, Sun, & Toh, 2010), 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 (1984) and their extension to SDPs was due to Nesterov and Nemirovskii (1994) and Alizadeh (1995). The development of primal-dual interior-point methods began with Kojima, Mizuno, and Yoshise (1989) and Kojima, Megiddo, Noma, and Yoshise (1991) and was further developed in early papers such as Vandenberghe and Boyd (1995) and Helmberg, Rendl, Vanderbei, and Wolkowicz (1996). It was eventually extended to SDP and SOCP in a unified way by Nesterov and Todd (1997, 1998). To solve SDPs, modern solvers, such as SDPT3 (Tütüncü, Toh, & Todd, 2003), SeDuMi (Sturm, 1999; 2002), and MOSEK (Andersen & Andersen, 2000), 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 (Vandenberghe, Boyd, & Wu, 1998), since

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

where the eigenvalues $\lambda_j(\mathbf{X})$ 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:

$$\begin{aligned} & \text{minimize} && \mathbf{C} \bullet \mathbf{X} - \mu \log \det \mathbf{X} \\ & \mathbf{X} \in \mathbb{S}^n \end{aligned} \quad (48a)$$

$$\text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{b}, \quad (48b)$$

and

$$\begin{aligned} & \text{maximize} && \mathbf{b}^\top \mathbf{y} + \mu \log \det \mathbf{S} \\ & \mathbf{y} \in \mathbb{R}^m, \mathbf{S} \in \mathbb{S}^n \end{aligned} \quad (49a)$$

subject to

$$\mathbf{A}^*(\mathbf{y}) + \mathbf{S} = \mathbf{C}, \quad (49b)$$

whose solutions are denoted as \mathbf{X}_μ and $(\mathbf{y}_\mu, \mathbf{S}_\mu)$, 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:

$$\text{Primal feasibility: } \mathcal{A}(\mathbf{X}_\mu) = \mathbf{b} \quad (50a)$$

$$\text{Dual feasibility: } \mathbf{A}^*(\mathbf{y}_\mu) + \mathbf{S} = \mathbf{C} \quad (50b)$$

$$\text{Complementarity slackness: } \mathbf{X}_\mu \mathbf{S}_\mu = \mu \mathbf{I} \quad (50c)$$

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

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

and

$$\mathbf{C} \bullet \mathbf{X}^* \leq \mathbf{C} \bullet \mathbf{X}_\mu \leq \mathbf{C} \bullet \mathbf{X}^* + n\mu,$$

$$\mathbf{b}^\top \mathbf{y}^* - n\mu \leq \mathbf{b}^\top \mathbf{y}_\mu \leq \mathbf{b}^\top \mathbf{y}^*,$$

where $\mathbf{X}^* \in \mathbb{S}^n$ and $(\mathbf{y}^*, \mathbf{S}^*)$ are the optimal solution of (39) and (40), respectively. By gradually decreasing μ towards 0, the solutions $\{\mathbf{X}_\mu, \mathbf{y}_\mu, \mathbf{S}_\mu\}$ 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

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

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 (Andersen et al., 2014; Fukuda et al., 2001; Kim, Kojima, Mevissen, & Yamashita, 2011; Madani et al., 2016; Madani, Sojoudi et al., 2015; Nakata, Fujisawa, Fukuda, Kojima, & Murota, 2003; Vandenberghe, Andersen et al., 2015; Zhang & Lavaei, 2017b).

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

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

where $\mathbf{H} = [\mathbf{A}^\top (\mathbf{K}^{\text{sc}} \otimes \mathbf{K}^{\text{sc}}) \mathbf{A}]$, $\mathbf{A} = [\text{vec}(\mathbf{A}_1) \quad \dots \quad \text{vec}(\mathbf{A}_m)]$, and \mathbf{K}^{sc} is the positive definite scaling matrix, which is the matrix geometric mean of the primal iterate \mathbf{X} and the inverse of the dual iterate \mathbf{S} . Despite that the data matrix \mathbf{A} are sparse, the Hessian matrix \mathbf{H} is often fully-dense because \mathbf{K}^{sc} 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 and

low-rank problem and dense and high-rank problem, which result in approximately the same amount of time and memory use.

To deal with this issue, one can employ the preconditioned conjugate gradient (PCG) algorithm, which requires a single matrix-vector product with the governing matrix \mathbf{H} and a single matrix-vector product with the inverse of the preconditioner $\tilde{\mathbf{H}}$ in each iteration (Barrett et al., 1994). The key to numerical performance is a preconditioner that has desirable characteristics, such as being similar to \mathbf{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 Zhang and Lavaei (2017a), where

$$\tilde{\mathbf{H}} = \tau^2 \mathbf{A}^\top \mathbf{A} + \mathbf{K} \mathbf{K}^\top, \quad (51)$$

where \mathbf{K} is a matrix with rank at most $n \text{rank}(\mathbf{X}^*)$, which is constructed to improve the conditioning of the preconditioner. In particular, it has been proved that the joint condition number $\kappa = \lambda_{\max}(\tilde{\mathbf{H}}^{-1} \mathbf{H}) / \lambda_{\min}(\tilde{\mathbf{H}}^{-1} \mathbf{H})$ is an absolute constant $O(1)$. This can be combined with the standard PCG convergence result to yield that PCG with the described $\tilde{\mathbf{H}}$ can solve the Hessian equation to L digits of accuracy 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 Fukuda et al. (2001), Nakata et al. (2003), and was later refined in Kim et al. (2011).

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

$$\text{subject to} \quad \sum_{k=1}^B \mathbf{A}_{i,k} \bullet \mathbf{W} \{ \mathcal{B}_k \} = b_i, \quad i \in \{1, \dots, m\} \quad (52b)$$

$$\mathbf{W} \{ \mathcal{B}_k \} \succeq 0, \quad k \in \{1, 2, \dots, B\} \quad (52c)$$

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 (Jabri, 2012; Madani et al., 2016; Madani, Sojoudi et al., 2015; Molzahn et al., 2013). 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 (Fukuda et al., 2001; Kim et al., 2011; Madani et al., 2016; Madani, Sojoudi et al., 2015; Nakata et al., 2003). 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. Here, ω is the treewidth of the network and n is one of the dimensions of \mathbf{X} . Further speed-ups were obtained by dropping some of the overlapping constraints, though the reformulation may no longer be exact (Andersen et al., 2014). Nevertheless, it has been shown in Zhang and Lavaei (2017b) 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 Zhang and Lavaei (2017b).

Definition 1 (Decoupled SDP). Given index sets $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_B\}$, we say that the linear constraints $\mathbf{A}_i \bullet \mathbf{W} = b_i$ is decoupled if there

exists $\mathcal{B}_k \in \mathcal{B}$ and some choice of $\mathbf{A}_{i,k} \in \mathbb{S}^{|\mathcal{B}_k|}$ such that

$$\mathbf{A}_i \bullet \mathbf{W} = \mathbf{A}_{i,k} \bullet \mathbf{W} \{ \mathcal{B}_k \} \quad \forall \mathbf{W} \in \mathbb{S}^n.$$

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 Zhang & Lavaei, 2017b, 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 (Zhang & Lavaei, 2017b). 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 10,000 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 $\mathbf{X}^\top = \mathbf{R} \mathbf{R}^\top$, where $\mathbf{R} \in \mathbb{R}^{n \times r}$ is a rank- r matrix (Burer & Monteiro, 2003; Journée, Bach, Absil, & Sepulchre, 2010). This reformulation is mainly motivated by the following theoretical result for low-rank SDP, which was proved concurrently in Barvinok (1995) and Pataki (1998).

Theorem 3 (Barvinok, 1995; Pataki, 1998). Consider (39) with m equality constraints. Then, there exists an optimal solution \mathbf{X}^* of (39) with rank \tilde{r} satisfying $\tilde{r}(\tilde{r} + 1)/2 \leq m$.

This result indicates that as long as r (i.e., the number of columns of the reformulated matrix variable \mathbf{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:

$$\underset{\mathbf{R} \in \mathbb{R}^{n \times r}}{\text{minimize}} \quad \mathbf{C} \bullet (\mathbf{R} \mathbf{R}^\top) \quad (53a)$$

$$\text{subject to} \quad \mathbf{A}(\mathbf{R} \mathbf{R}^\top) = \mathbf{b}, \quad (53b)$$

which replaces the positive semidefinite constraint in (39) with the quadratic equality constraint $\mathbf{X} = \mathbf{R} \mathbf{R}^\top$. An immediate advantage is the reduction of the number of variables from n^2 to $n \times r$, which can be substantial when \mathbf{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 Burer and Monteiro (2003), which optimizes over the augmented Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{R}, \mathbf{y}, \sigma) = & \mathbf{C} \bullet (\mathbf{R} \mathbf{R}^\top) - \sum_{i=1}^m y_i (\mathbf{A}_i \bullet (\mathbf{R} \mathbf{R}^\top) - b_i) \\ & + \frac{\sigma}{2} \sum_{i=1}^m (\mathbf{A}_i \bullet (\mathbf{R} \mathbf{R}^\top) - b_i)^2, \end{aligned} \quad (54)$$

where the variables $\mathbf{R} \in \mathbb{R}^{n \times r}$ and $\mathbf{y} \in \mathbb{R}^m$ 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 (\mathbf{y}^k, σ_k) given the current primal iterate \mathbf{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 (Burer & Monteiro, 2003) only provides guarantees of local convergence due the nonconvex nature. Recently, the theory gap started to close with a series of papers (Bandeira, Boumal, & Voroninski, 2016; Bhojanapalli, Boumal, Jain, & Netrapalli, 2018; Boumal, Voroninski, & Bandeira, 2016; Park, Kyriolidis, Bhojanapalli, Caramanis, & Sanghavi, 2016). For instance, the paper (Boumal et al., 2016) 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 (Bhojanapalli et al., 2018) 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 Mareček and Takáč (2017), which employs the well-established Augmented Lagrangian approach (Conn, Gould, & Toint, 1991) and a parallel coordinate descent with a closed-form step. The method compares favorably with SDP based methods (Madani et al., 2016) on small-scale instances, but experienced numerical difficulties on large instances, such as the Polish network with more than 2000 nodes. This motivates the combination of first- and second-order methods (Liu, Liddell, Mareček, & Takáč, 2017) 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 solutions of 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. However, several challenges remain. 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.

In this work, we have summarized the body of research on conic relaxations of the OPF problem. We cast OPF in a lifted rectangular form and surveyed the state-of-the-art:

- semidefinite programming relaxation and strengthening valid inequalities,
- reduced semidefinite programming based on graph-theoretic decompositions
- second-order cone programming relaxation
- quadratic convex relaxation
- linear programming relaxation.

Additionally, we covered heuristics and recovery algorithms that can be used towards projecting an infeasible solution of convex relaxations onto the non-convex feasible set of OPF. Complex sum-of-squares hierarchy is surveyed as another remedy for cases where basic convex relaxations fail to produce a globally optimal solution. Lastly, we covered scalable first- and second-order numerical algorithm for solving large instances of OPF relaxations.

As demonstrated in this paper, semidefinite programming and sum-of-squares based relaxations are effective tools for tackling challenging instances of OPF. Hence, one of the major barriers towards incorporating accurate OPF models into real-world power system problems is the lack of reliable mixed-integer solvers that are compatible with high-order semidefinite programming constraints. It should be emphasized that power flow nonlinearity is one among many sources of complexity in power optimization problems. Three of the other major research challenges involve the presence of integer parameters, uncertainties, and transient phenomena. Finer modeling of the power system should be considered to account for discrete aspects such as transformer ratios and the turning on or off of generating units (Bingane et al., 2018; Bingane et al., 2019). Additionally, dynamical aspects could be considered, as in transient stability analysis. Finally, data-driven approaches should be incorporated into OPF formulations to realize secure and resilient operation.

Acknowledgment

F. Zohrizadeh and R. Madani were funded by NSF award #1809454. J. Ming, C. Josz, J. Lavaei and S. Sojoudi were supported by grants from NSF, AFOSR, ONR and ARO.

References

Ahmadi, A., & 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.

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

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

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

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

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

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

Bai, X., & 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.

Bai, X., & 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.

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

Baker, K., Dall'Anese, E., & Summers, T. (2016). Distribution-agnostic stochastic optimal power flow for distribution grids. In *Proceedings of the 2016 North American power symposium (NAPS)* (pp. 1–6). IEEE.

Bakirtzis, A. G., Biskas, P. N., Zoumas, C. E., & Petridis, V. (2002). Optimal power flow by enhanced genetic algorithm. *IEEE Transactions on Power Systems*, 17(2), 229–236.

Bandeira, A. S., Boumal, N., & Voroninski, V. (2016). On the low-rank approach for semidefinite programs arising in synchronization and community detection. In *Proceedings of the conference on learning theory* (pp. 361–382).

Baran, M., & Wu, F. (1989). Optimal capacitor placement on radial distribution systems. *IEEE Transactions on Power Delivery*, 4(1), 725–734. doi:10.1109/61.19265.

Barati, M., & Kargarian, A. (2017). A global algorithm for AC optimal power flow based on successive linear conic optimization. In *Proceedings of the 2017 IEEE power & energy society general meeting* (pp. 1–5).

Barrett, R., Berry, M. W., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., ... Van der Vorst, H. (1994). *Templates for the solution of linear systems: building blocks for iterative methods*: 43. SIAM.

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

Ben-Tal, A., & Nemirovski, A. (2001). Lectures on modern convex optimization – Analysis, algorithms, and engineering applications. *MPS-SIAM series on optimization* 2. SIAM.

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

Bienstock, D., & Munoz, G. (2014). On linear relaxations of OPF problems. arXiv preprint arXiv:1411.1120.

Bienstock, D., & Verma, A. (2015). Strong NP-hardness of AC power flows feasibility. arXiv preprint arXiv:1512.07315.

Bingane, C., Anjos, M. F., & 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.

Bingane, C., Anjos, M. F., & Le Digabel, S. (2019). Tight-and-cheap conic relaxation for the optimal reactive power dispatch problem. *IEEE Transactions on Power Systems*, 34(6), 4684–4693.

Bloomberg New Energy Finance, J. (2016). New energy outlook 2018. <https://about.bnef.com/new-energy-outlook/>.

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

Bose, S., Gayme, D. F., Low, S., & Chandy, K. M. (2011). Optimal power flow over tree networks. In *Proceedings of the annual Allerton conference on communication, control, and computing* (pp. 1342–1348).

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

Bouktir, T., & Slimani, L. (2005). Optimal power flow of the algerian electrical network using an ant colony optimization method. *Leonardo Journal of Sciences*, 6, 43–57.

Boumal, N., Voroninski, V., & Bandeira, A. (2016). The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In *Proceedings of the advances in neural information processing systems* (pp. 2757–2765).

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.

Bukhsh, W., Grothey, A., McKinnon, K., & Trodden, P. (2013a). Local solutions of the optimal power flow problem. *IEEE Transactions on Power Systems*, 28, 4780–4788.

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

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

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

Cain, M. B., O'Neil, R. P., & Castillo, A. (2012). History of optimal power flow and formulations (OPF paper 1). *Technical Report*. US Federal Energy Regulatory Commission.

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

Capitanescu, F., Ramos, J. M., Panciatici, P., Kirschen, D., Marcolini, A. M., Platbrood, L., & 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.

Carpentier, M. (1962). Contribution à l'étude du dispatching économique. *Bulletin de la Société Chimique de France*, 8, 431–447.

Castillo, A., Laird, C., Silva-Monroy, C. A., Watson, J.-P., & O'Neill, R. P. (2016). The unit commitment problem with AC optimal power flow constraints. *IEEE Transactions on Power Systems*, 31(6), 4853–4866.

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

Castillo, A., & O'Neill, R. (2013a). Computational performance of solution techniques applied to the ACOPF (OPF paper 5). *Technical Report*. US FERC.

Castillo, A., & O'Neill, R. (2013b). Survey of approaches to solving the ACOPF (OPF paper 4). *Technical Report*. US Federal Energy Regulatory Commission.

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

Coffrin, C., Hijazi, H. L., & Van Hentenryck, P. (2015). DistFlow extensions for AC transmission systems. arXiv preprint arXiv:1506.04773.

Coffrin, C., Hijazi, H. L., & 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.

Coffrin, C., Hijazi, H. L., & 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.

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

Conn, A. R., Gould, N. I., & 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.

Cui, B., & 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.

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

Dall'Anese, E., Baker, K., & Summers, T. (2017). Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems*, 32(5), 3427–3438.

D'Angelo, J., & Putinar, M. (2008). Polynomial optimization on odd-dimensional spheres. In *Emerging applications of algebraic geometry*. New York: Springer.

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

Douglas, J., & Rachford, H. H. (1956). On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82(2), 421–439.

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

Eldridge, B., O'Neill, R., & Castillo, A. (2018). An improved method for the DCOPF with losses. *IEEE Transactions on Power Systems*, 33(4), 3779–3788.

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

Farivar, M., Clarke, C. R., Low, S. H., & Chandy, K. M. (2011). Inverter VAR control for distribution systems with renewables. In *Proceedings of the 2011 IEEE international conference on smart grid communications (SmartGridComm)* (pp. 457–462).

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

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

Fattah, S., Ashraphijuo, M., Lavaei, J., & Atamtürk, A. (2017). Conic relaxations of the unit commitment problem. *Energy*, 134, 1079–1095.

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

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

Gabay, D., & Mercier, B. (1976). A dual algorithm for the solution of non linear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1), 17–40.

Gan, L., Li, N., Topcu, U., & Low, S. (2012). On the exactness of convex relaxation for optimal power flow in tree networks. In *Proceedings of the IEEE conference on decision and control* (pp. 465–471).

Gan, L., Li, N., Topcu, U., & Low, S. H. (2013). Optimal power flow in tree networks. In *Proceedings of the IEEE conference on decision and control* (pp. 2313–2318).

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

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

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

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

Glowinski, R., & 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.

Gopalakrishnan, A., Raghunathan, A. U., Nikovski, D., & Biegler, L. T. (2011). Global optimization of optimal power flow using a branch & bound algorithm. In *Proceedings of the annual Allerton conference on communication, control, and computing*.

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

Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., & Summers, T. (2018a). Data-based distributionally robust stochastic optimal power flow, Part I: Methodologies. *IEEE Transactions on Power Systems*, 34(2), 1483–1492.

Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., & Summers, T. (2018b). Data-based distributionally robust stochastic optimal power flow, Part II: Case studies. *IEEE Transactions on Power Systems*, 34(2), 1493–1503.

Gutierrez-Martinez, V. J., Cañizares, C. A., Fuerte-Esquivel, C. R., Pizano-Martinez, A., & Gu, X. (2010). Neural-network security-boundary constrained optimal power flow. *IEEE Transactions on Power Systems*, 26(1), 63–72.

Halilbašić, L., Thams, F., Venzke, A., Chatzivassileiadis, S., & Pinson, P. (2018). Data-driven security-constrained AC-OPF for operations and markets. In *Proceedings of the 2018 power systems computation conference (PSCC)* (pp. 1–7). IEEE.

He, B., & 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.

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

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

Helberg, C., Rendl, F., Vanderbei, R. J., & Wolkowicz, H. (1996). An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6(2), 342–361.

Hijazi, H., Coffrin, C., & Van Hentenryck, P. (2016). Polynomial SDP cuts for optimal power flow. In *Proceedings of the power systems computation conference (PSCC)* (pp. 1–7).

Huang, S., Wu, Q., Wang, J., & 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.

Jabr, R. A. (2006a). Radial distribution load flow using conic programming. *IEEE Transactions on Power Systems*, 21, 1458–1459.

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

Jabr, R. A. (2012). Exploiting sparsity in SDP relaxations of the OPF problem. *IEEE Transactions on Power Systems*, 27(2), 1138–1139. doi:10.1109/TPWRS.2011.2170772.

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

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

Josz, C. (2017). Counterexample to global convergence of DSOS and SDSOS hierarchies. <https://arxiv.org/pdf/1707.02964.pdf>.

Josz, C., Fliscounakis, S., Maeght, J., & Panciatici, P. (2016). AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE. <https://arxiv.org/abs/1603.01533>.

Josz, C., & Henrion, D. (2016). Strong duality in Lasserres hierarchy for polynomial optimization. *Optimization Letters*, 10(1), 3–10.

Josz, C., Maeght, J., Panciatici, P., & 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. doi:10.1109/TPWRS.2014.2320819.

Josz, C., & Molzahn, D. K. (2018). Lasserre hierarchy for large scale polynomial optimization in real and complex variables. *SIAM Journal on Optimization*, 28(2), 1017–1048.

Journée, M., Bach, F., Absil, P.-A., & Sepulchre, R. (2010). Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5), 2327–2351.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. In *Proceedings of the annual ACM symposium on theory of computing* (pp. 302–311). ACM.

Kekatos, V., & Giannakis, G. B. (2013). Distributed robust power system state estimation. *IEEE Transactions on Power Systems*, 28(2), 1617–1626.

Kheirandishfard, M., Zohrizadeh, F., & Madani, R. (2018). Convex relaxation of bilinear matrix inequalities Part I: Theoretical results. In *Proceedings of the IEEE conference on decision and control*.

Kim, S., Kojima, M., Mevissen, M., & Yamashita, M. (2011). Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. *Mathematical Programming*, 129(1), 33–68.

Kocuk, B., Dey, S. S., & Sun, X. A. (2016a). Inexactness of SDP relaxation and valid inequalities for optimal power flow. *IEEE Transactions on Power Systems*, 31(1), 642–651.

Kocuk, B., Dey, S. S., & Sun, X. A. (2016b). Strong SOCP relaxations for the optimal power flow problem. *Operations Research*, 64(6), 1177–1196.

Kocuk, B., Dey, S. S., & 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.

Kocuk, B., Dey, S. S., & Sun, X. A. (2018). Matrix minor reformulation and SOCP-based spatial branch-and-cut method for the AC optimal power flow problem. *Mathematical Programming*, 10(4), 557–596.

Kočvara, M., & Stingl, M. (2003). Pennon: A code for convex nonlinear and semidefinite programming. *Optimization Methods and Software*, 18(3), 317–333.

Kojima, M., Megiddo, N., Noma, T., & Yoshise, A. (1991). *A unified approach to interior point algorithms for linear complementarity problems*: 538. Springer Science & Business Media.

Kojima, M., Mizuno, S., & Yoshise, A. (1989). A primal-dual interior point algorithm for linear programming. In *Progress in mathematical programming* (pp. 29–47). Springer.

Kuang, X., Ghaddar, B., Naoum-Sawaya, J., & Zuluaga, F. (2018). Alternative SDP and SOCP approximations for polynomial optimization. *EURO Journal on Computational Optimization*, 7(2), 1–23.

Kuang, X., Ghaddar, B., Naoum-Sawaya, J., & Zuluaga, L. F. (2017). Alternative LP and SOCP Hierarchies for ACOPF Problems. *IEEE TPS*, 32, 2828–2836.

Kurpisz, A., & de Wolff, T. (2019). New dependencies of hierarchies in polynomial optimization. In *Proceedings of the international symposium on symbolic and algebraic computation* (pp. 251–258).

LaCommare, K. H., & Eto, J. H. (2006). Cost of power interruptions to electricity consumers in the united states (US). *Energy*, 31(12), 1845–1855.

Lam, A. Y., Zhang, B., & David, N. T. (2012). Distributed algorithms for optimal power flow problem. In *Proceedings of the 2012 conference on decision and control* (pp. 430–437).

Lasserre, J. B. (2000). Optimisation globale et théorie des moments. *Comptes rendus de l'Academie des Sciences, Paris, Série I*, 331, 929–934.

Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11, 796–817.

Lasserre, J. B. (2009). *Moments, positive polynomials and their applications*: 1. World Scientific.

Lavaei, J., & Low, S. (2010a). Convexification of optimal power flow problem. In *Proceedings of the annual Allerton conference on communication, control, and computing*.

Lavaei, J., & Low, S. (2010b). Relationship between power loss and network topology in power systems. In *Proceedings of the IEEE conference on decision and control*.

Lavaei, J., & Low, S. (2012). Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1), 92–107.

Lavaei, J., Tse, D., & Zhang, B. (2014). Geometry of power flows and optimization in distribution networks. *IEEE Transactions on Power Systems*, 29(2), 572–583.

Lee, D., Nguyen, H. D., Dvijotham, K., & Turitsyn, K. (2018). Convex restriction of power flow feasibility set. *arXiv preprint arXiv:1803.00818*.

Lehmann, K., Grastien, A., & Van Hentenryck, P. (2016). AC-feasibility on tree networks is NP-hard. *IEEE Transactions on Power Systems*, 31(1), 798–801.

Lesieutre, B., Molzahn, D., Borden, A., & DeMarco, C. (2011). Examining the limits of the application of semidefinite programming to power flow problems. In *Proceedings of the annual Allerton conference on communication, control, and computing* (pp. 28–30).

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

Li, N., Chen, L., & Low, S. H. (2012). Exact convex relaxation of OPF for radial networks using branch flow model. In *Proceedings of the IEEE SmartGridComm* (pp. 7–12).

Li, Q., & 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.

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

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

Lipka, P., Oren, S. S., O'Neill, R. P., & Castillo, A. (2017). Running a more complete market with the SLP-IV-ACOPF. *IEEE Transactions on Power Systems*, 32(2), 1139–1148.

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

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

Louca, R., & Bitar, E. (2018). Robust AC optimal power flow. *IEEE Transactions on Power Systems*, 33(4), 1669–1681.

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

Lovasz, & Schrijver (1991). Cones of matrices and set-functions and 01 optimization. *SIAM Journal on Optimization*, 1, 166–190.

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.

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

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

Madani, R., Atamtürk, A., & Davoudi, A. (2017). A scalable semidefinite relaxation approach to grid scheduling. *arXiv preprint arXiv:1707.03541*.

Madani, R., Kalbat, A., & Lavaei, J. (2015). ADMM for sparse semidefinite programming with applications to optimal power flow problem. In *Proceedings of the IEEE conference on decision and control* (pp. 5932–5939).

Madani, R., Kalbat, A., & Lavaei, J. (2018). A low-complexity parallelizable numerical algorithm for sparse semidefinite programming. *IEEE Transactions on Control of Network Systems*, 5(4), 1898–1909. doi:10.1109/TCNS.2017.2774008.

Madani, R., Kheirandishfard, M., Lavaei, J., & Atamtürk, A. (2018). Penalized conic relaxations for quadratically-constrained quadratic programming. Preprint: https://lavaei.ieor.berkeley.edu/Penalized_conic_2018.pdf.

Madani, R., Lavaei, J., & Baldick, R. (2019). Convexification of power flow equations in the presence of noisy measurements. *IEEE Transactions on Automatic Control*, 64(8), 3101–3116.

Madani, R., Lavaei, J., Baldick, R., & Atamtürk, A. (2017). Power system state estimation and bad data detection by means of conic relaxation. In *Proceedings of the 50th Hawaii international conference on system sciences*.

Madani, R., Sojoudi, S., & Lavaei, J. (2015). Convex relaxation for optimal power flow problem: mesh networks. *IEEE Transactions on Power Systems*, 30(1), 199–211.

Majumdar, A., Ahmadi, A. A., & Tedrake, R. (2014). Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *Proceedings of the IEEE conference on decision and control* (pp. 394–401).

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

Mareček, J., & 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.

Marley, J. F., Molzahn, D. K., & 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.

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

Marshall, M. (2009). Representation of non-negative polynomials, degree bounds and applications to optimization. *The Canadian Journal of Mathematics*, 61, 205–221.

Mehta, D., Molzahn, D. K., & Turitsyn, K. (2016a). Recent advances in computational methods for the power flow equations. In *Proceedings of the 2016 American control conference (ACC)* (pp. 1753–1765).

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

Meith, R., & Dvorkin, Y. (2018). Data-driven distributionally robust optimal power flow for distribution systems. *IEEE Control Systems Letters*, 2(3), 363–368.

Misra, S., Molzahn, D. K., & Dvijotham, K. (2018). Optimal adaptive linearizations of the AC power flow equations. In *Proceedings of the 2018 power systems computation conference (PSCC)* (pp. 1–7).

Molzahn, D., & Hiskens, I. (2014). Moment-based relaxation of the optimal power flow problem. In *Proceedings of the 18th power systems computation conference (PSCC)*.

Molzahn, D., & Hiskens, I. (2015a). Mixed SDP/SOCP moment relaxations of the optimal power flow problem. In *Proceedings of the IEEE eindhoven powertech*.

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

Molzahn, D., Josz, C., Hiskens, I., & Panciatici, P. (2015). Solution of optimal power flow problems using moment relaxations augmented with objective function penalization. In *Proceedings of the 54th IEEE conference on decision and control (CDC)* (pp. 31–38). doi:10.1109/CDC.2015.7402083.

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

Molzahn, D. K., & 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.

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.

Molzahn, D. K., Holzer, J. T., Lesieutre, B. C., & 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.

Molzahn, D. K., Josz, C., Hiskens, I. A., & 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.

Momoh, J., Adapa, R., & El-Hawary, M. (1999). 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. doi:10.1109/59.744492.

Momoh, J., El-Hawary, M., & Adapa, R. (1999). 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. doi:10.1109/59.744495.

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

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

National Academies of SciencesEngineering, and Medicine and others (2016). *Analytic research foundations for the next-generation electric grid*. National Academies Press. doi:10.17226/21919.

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

Nesterov, Y., & Nemirovskii, A. (1994). *Interior-point polynomial algorithms in convex programming*: 13. SIAM.

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

Nesterov, Y. E., & Todd, M. J. (1998). Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8(2), 324–364.

Nick, M., Cherkouci, R., Le Boudec, J.-Y., & 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.

Nie, J. (2014). Optimality conditions and finite convergence of Lasserre's hierarchy. *Mathematical Programming*, 146, 97–121.

O'Donoghue, B., Chu, E., Parikh, N., & Boyd, S. (2016). Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3), 1042–1068.

Pandya, K., & Joshi, S. (2008). A survey of optimal power flow methods. *Journal of Theoretical and Applied Information Technology*, 4, 450–458.

Park, D., Kyriolidis, A., Bhojanapalli, S., Caramanis, C., & Sanghavi, S. (2016). Provable Burer-Monteiro factorization for a class of norm-constrained matrix problems. arXiv preprint arXiv:1606.01316.

Parrilo, P. (2000). *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. Cal. Inst. of Tech. Ph.D. thesis.

Parrilo, P. (2003). Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96, 293–320.

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.

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

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

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

Putinar, M. (1993). Positive polynomials on compact semi-algebraic sets. *The Indiana University Mathematics Journal*, 42, 969–984.

Riener, C., Theobald, T., Andrén, L. J., & Lasserre, J. B. (2013). Exploiting symmetries in SDP-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1), 122–141.

Saad, Y. (2003). *Iterative methods for sparse linear systems*: 82. SIAM.

Schmüdgen, K. (1991). The K-moment problem for semi-algebraic sets. *Mathematische Annalen*, 289, 203–206.

Shchetinin, D., De Rubira, T. T., & Hug-Glazmann, G. (2018). On the construction of linear approximations of line flow constraints for AC optimal power flow. *IEEE Transactions on Power Systems*, 34(2), 1182–1192.

Sherali, H., & Adams, W. (1990). A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3, 411–430.

Shor, N. (1987). Quadratic optimization problems. *Soviet Journal of Computer and System Sciences*, 25, 1–11.

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

Sojoudi, S., & Lavaei, J. (2011). Network topologies guaranteeing zero duality gap for optimal power flow problem. submitted for publication. preprint: https://lavaei.ier.berkeley.edu/Net_Top.pdf.

Sojoudi, S., & Lavaei, J. (2012). Physics of power networks makes hard optimization problems easy to solve. In *Proceedings of the IEEE power and energy society general meeting* (pp. 1–8).

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

Stengle, G. (1974). A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207, 87–97.

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

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.

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.

Subhonnmesh, B., Low, S. H., & Chandy, K. M. (2012). Equivalence of branch flow and bus injection models. In *Proceedings of the annual Allerton conference on communication, control, and computing* (pp. 1893–1899).

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

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

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

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

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

Toh, K.-C., Todd, M. J., & 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* (pp. 715–754). Springer.

Tütüncü, R. H., Toh, K.-C., & Todd, M. J. (2003). Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, 95(2), 189–217.

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

Vandenberghe, L., & Boyd, S. (1995). A primaldual potential reduction method for problems involving matrix inequalities. *Mathematical Programming*, 69(1–3), 205–236.

Vandenberghe, L., Boyd, S., & Wu, S.-P. (1998). Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2), 499–533.

Venzke, A., & Chatzivasileiadis, S. (2018). Convex relaxations of security constrained AC optimal power flow under uncertainty. In *Proceedings of the 2018 power systems computation conference (PSCC)* (pp. 1–7).

Venzke, A., & Chatzivasileiadis, S. (2019). Convex relaxations of probabilistic AC optimal power flow for interconnected AC and HVDC grids. *IEEE Transactions on Power Systems*, 34(4), 1–13.

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

Waki, H., Kim, S., Kojima, M., & Muramatsu, M. (2006). Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1), 218–242.

Wang, C., Cui, B., Wang, Z., & Gu, C. (2018). SDP-based optimal power flow with steady-state voltage stability constraints. *IEEE Transactions on Smart Grid*, 10(4), 1–9.

Wang, Z., Cui, B., & 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.

Watson, J.-P., Silva Monroy, C. A., Castillo, A., Laird, C., & O'Neill, R. (2015). Security-Constrained Unit Commitment with Linearized AC Optimal Power Flow. *Technical Report*, Albuquerque, NM (United States): Sandia National Lab. (SNL-NM).

Wei, W., Wang, J., Li, N., & 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.

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

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

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

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

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

Yu, S., Nguyen, H. D., & Turitsyn, K. S. (2015). Simple certificate of solvability of power flow equations for distribution systems. In *Proceedings of the 2015 IEEE power & energy society general meeting* (pp. 1–5).

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

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

Zhang, R., Josz, C., & Sojoudi, S. (2018). Conic optimization with applications to machine learning and energy systems. *Annual Reviews in Control*, 47(18), 323–340.

Zhang, R. Y., & Lavaei, J. (2017a). Modified interior-point method for large-and-sparse low-rank semidefinite programs. In *Proceedings of the IEEE conference on decision and control* (pp. 5640–5647).

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

Zhang, Y., Hong, M., Dall'Anese, E., Dhople, S., & Xu, Z. (2017). Distributed controllers seeking AC optimal power flow solutions using ADMM. *IEEE Transactions on Smart Grid*, 9(5), 4525–4537.

Zhang, Y., Madani, R., & Lavaei, J. (2018). Conic relaxations for power system state estimation with line measurements. *IEEE Transactions on Control of Network Systems*, 5(3), 1193–1205.

Zhang, Y., Shen, S., & Mathieu, J. L. (2015). Data-driven optimization approaches for optimal power flow with uncertain reserves from load control. In *Proceedings of the 2015 American control conference (ACC)* (pp. 3013–3018). IEEE.

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

Zheng, Y., Fantuzzi, G., Papachristodoulou, A., Goulart, P., & Wynn, A. (2017). Fast ADMM for semidefinite programs with chordal sparsity. In *Proceedings of the American control conference* (pp. 3335–3340).

Zhu, H., & 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.

Zohrizadeh, F., Kheirandishfard, M., Nasir, A., & Madani, R. (2018). Sequential relaxation of unit commitment with AC transmission constraints. In *Proceedings of the IEEE conference on decision and control* (pp. 2408–2413).

Zohrizadeh, F., Kheirandishfard, M., Quarm, E., & Madani, R. (2018). Penalized parabolic relaxation for optimal power flow problem. In *Proceedings of the IEEE conference on decision and control* (pp. 1616–1623).