2013 IREP Symposium-Bulk Power System Dynamics and Control -IX (IREP), August 25-30, 2013, Rethymnon, Greece

Convex Relaxation of Optimal Power Flow: A Tutorial

Steven H. Low
Engineering & Applied Science, Caltech
slow@caltech.edu

Abstract

This is a short survey of recent advances in the convex relaxation of the optimal power flow problem. Our focus is on understanding structural properties, especially the underlying convexity structure, of optimal power flow problems rather than different computational algorithms.

Introduction

The optimal power flow (OPF) problem is fundamental in power systems as it underlies many applications such as economic dispatch, unit commitment, state estimation, volt/var control, demand response, etc. OPF seeks to optimize a certain objective function, such as power loss, generation cost and/or user utilities, subject to Kirchhoff's laws, power balance as well as capacity, stability and security constraints on the voltages and power flows. There has been a great deal of research on OPF since Carpentier's first formulation in 1962 [1]. An early solution appears in [2] and recent surveys can be found in, e.g., [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14].

OPF is generally nonconvex and NP-hard, and a large number of optimization algorithms and relaxations have been proposed. A popular approximation is the DC OPF which is a linearization and therefore easy to solve, e.g. [15], [16], [17], [18]. An important observation was made in [19], [20] that AC OPF can be formulated as a quadratically constrained quadratic program and therefore can be approximated by a semidefinite program. While this approach is illustrated in [19], [20] on several IEEE test systems using an interior-point method, whether or when the semidefinite relaxation will turn out to be exact is not studied.

This extended abstract surveys main results on convex relaxations of OPF, formulated both using the bus injection model and the branch flow model. The bus injection model is the standard model for power flow analysis and optimization. It focuses on nodal voltages. The branch flow model, on the other hand, focuses on currents and powers on the branches. It has been used mainly for modeling distribution circuits which tend to be radial, but has received far less attention; see e.g. [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. We formulate OPF in each of these two models and summarize the main relaxation results in each model. Finally we show that these two models and their relaxations are equivalent.

The focus of this paper is on understanding structural properties, especially the underlying convexity structure, of OPF rather than numerical algorithms to compute a solution for which a huge literature exists.

Mathematical preliminaries

In this section we summarize basic concepts and notations in optimization, graph theory, and matrix completion that we will use in presenting recent advances in the convex relaxation of optimal power flow problems. More details can be found in, e.g., [32], [33], [34], [35], [36], [37], [38], [39].

Notations

Let $\mathbb{C}$ denote the set of complex numbers, $\mathbb{R}$ the set of real numbers, and $\mathbb{Z}$ the set of integers. For $a \in \mathbb{C}$, Re a and Im a denote the real and imaginary parts of a respectively. For any set $A \subseteq \mathbb{C}^n$ let conv A denote the convex hull of A. For $a \in \mathbb{R}$, $[a]^+ := \max\{a, 0\}$. For $a, b \in \mathbb{C}$, $a \le b$ means Re $a \le$ Re b and Im $a \le$ Im b. In general scalar or vector variables are in small letters, e.g., u, w, x, y, z. Most power system quantities however are in capital letters, e.g., $S_{ij}, P_{ij}, Q_{ij}, I_i, V_i$. A variable without a subscript denotes a vector with appropriate components, e.g., $s := (s_i, i = 1, \ldots, n)$, $S := (S_{ij}, (i,j) \in E)$. For a vector $a = (a_1, \ldots, a_k)$, a_{-i} denotes $(a_1, \ldots, a_{i-1}, a_{i+1}, a_k)$. For vectors x, y, $x \le y$ denotes inequality componentwise.

Matrices are in capital letters. The transpose of a matrix A is denoted by A^T and its Hermitian (complex conjugate) transpose by A^H. A matrix A is Hermitian if $A = A^H$. A is positive semidefinite (or psd), denoted by $A \succeq 0$, if A is Hermitian *and* $x^H A x \ge 0$ for all $x \in \mathbb{C}^n$; in particular if $A \succeq 0$ then by definition $A = A^H$. For matrices A, B, $A \succeq B$ means $A - B$ is psd. Let $\mathbb{S}^n$ be the set of all $n \times n$ Hermitian matrices and $\mathbb{S}^n_+$ the set of $n \times n$ psd matrices.

A graph $G = (N, E)$ consists of a set N of nodes and a set of edges $E \subseteq N \times N$. If G is undirected then $(j,k) \in E$ if and only if $(k,j) \in E$. If G is directed then $(j,k) \in E$ only if $(k,j) \notin E$; in this case we will use (j,k) and $j \rightarrow k$ interchangeably to denote an edge pointing from j to k. By "$j \sim k$" we mean an edge (j,k) if G is undirected and either $j \rightarrow k$ or $k \rightarrow j$ if G is directed.

978-1-4799-0199-9/13/ $31.00

Optimization

Quadratic constrained quadratic program (QCQP) is the following problem:

$$\min_{x\in\mathbb{C}^n} \quad x^H C_0 x \tag{1a}$$
$$\text{subj. to} \quad x^H C_m x \le b_m, \quad m = 1,\dots,M \tag{1b}$$

where, for $m = 0,\dots,M$, $C_m \in \mathbb{S}^n$ (so that $x^H C_m x$ are real) and $b_m \in \mathbb{R}$ are given. If C_m, $m = 0,\dots,M$, are positive semidefinite then (1) is a convex QCQP. Otherwise it is generally nonconvex.

Any psd rank-1 matrix X has a unique spectral decomposition $X = xx^H$. Using $x^H C_m x = \text{tr } C_m xx^H = \text{tr } C_m X$ we can rewrite a QCQP as the following equivalent problem where the optimization is over Hermitian matrices:

$$\min_{X\in\mathbb{S}^n} \quad \text{tr } C_0 X \tag{2a}$$
$$\text{subj. to} \quad \text{tr } C_m X \le b_m, \quad m = 1,\dots,M \tag{2b}$$
$$X \succeq 0, \quad \text{rank } X = 1 \tag{2c}$$

The objective function and the constraints (2b) are linear in X and $X \succeq 0$ is a convex constraint ($\mathbb{S}^n_+$ is a convex set). The rank constraint in (2c) is the only nonconvex constraint and the only source of computational difficulty.

Removing this constraint results in a semidefinite program (SDP):

$$\min_{X\in\mathbb{S}^n} \quad \text{tr } C_0 X \tag{3a}$$
$$\text{subj. to} \quad \text{tr } C_m X \le b_m, \quad m = 1,\dots,M \tag{3b}$$
$$X \succeq 0 \tag{3c}$$

SDP is a convex program and can be efficiently computed. We call (3) a *SDP relaxation* of QCQP (1) because the feasible set of (2) is a subset of the feasible set of SDP (3). A strategy for solving QCQP (1) is to solve SDP (3) for an optimal X^* and check its rank. If rank $X^* = 1$ then X^* is optimal for (2) as well and an optimal solution x^* of QCQP (1) can be recovered from X^* through spectral decomposition $X^* = x^*(x^*)^H$. If $X^* > 1$ then no feasible solution of QCQP can be directly obtained from X^* but the optimal objective value of SDP provides a lower bound to that of QCQP.

A special case of SDP is a second-order cone program (SOCP) in the following rotated form:

$$\min_{x\in\mathbb{C}^n} \quad c_0^H x \tag{4a}$$
$$\text{subj. to} \quad \|C_m x + b_m\|^2 \le (c_m^H x + d_m)(\hat{c}_m^H x + \hat{d}_m)$$
$$m = 1,\dots,M \tag{4b}$$

In this paper we will formulate optimal power flow problems as QCQPs and describe SDP and SOCP relaxations of OPF. The third relaxation we will describe is chordal relaxation based on the notion of chordal extension of a network graph. We now review some basic graph concepts, relate them to the solution matrices of semidefinite programs, and show that a chordal relaxation is indeed a semidefinite program.

Graph

Consider a graph $G = (N, E)$ with $N := (1,\dots,n)$. G can either be undirected or directed with an arbitrary orientation. Two nodes j and k are *adjacent* if $j \sim k \in E$. A *complete* graph is one where every pair of nodes is adjacent. A subgraph of G is a graph $F = (N', E')$ with $N' \subseteq N$ and $E' \subseteq E$. A *clique* of G is a complete subgraph of G. A *maximal clique* of G is a clique that is not a subgraph of another clique of G.

By a *path* connecting nodes j and k we mean either a set of *distinct* nodes $(j, n_1, \dots, n_i, k)$ such that $(j \sim n_1), (n_1 \sim n_2), \dots, (n_i \sim k)$ are edges in E or this set of edges, depending on the context. A *cycle* $(n_1,\dots,n_i)$ is a path such that $(n_1 \sim n_2),\dots,(n_i \sim n_1)$ are edges in E. By convention we exclude a pair of adjacent nodes (j,k) as a cycle. We will only consider connected graphs in which there is path between every pair of nodes.

A cycle in G that has no chord (an edge connecting two nodes that are non-adjacent in the cycle) is called a *minimal cycle*. G is *chordal* if all its minimal cycles are of length 3 (recall that an edge (j,k) is not considered a cycle). A *chordal extension* of G is a chordal graph on the same set of nodes as G that contains G as a subgraph. Every graph has a chordal extension; e.g. the complete graph on the same set of nodes is a trivial chordal extension.

Partial matrix and completion

Fix a graph $G = (N, E)$. For our purposes here we assume G is undirected so that $(j,k) \in E$ if and only if $(k,j) \in E$. A *G-partial matrix* (or simply a *partial matrix* if G is clear from the context) is a set of complex numbers:

$$X_G \;:=\; ([X_G]_{jj} \in \mathbb{C}, j \in N, \; [X_G]_{ij} \in \mathbb{C}, (i,j) \in E)$$

One can treat a partial matrix X_G as entries of an $n \times n$ matrix X whose entries X_{jk} are unspecified if $(j,k) \notin E$. Given a partial matrix X_G we call an $n \times n$ matrix X a *completion of* X_G if $X_{jj} = [X_G]_{jj}, j \in N$, and $X_{jk} = [X_G]_{jk}, (j,k) \in E$, i.e., X agrees with X_G on G.[1]

Consider any $n \times n$ matrix X. Given any $k \le n$ nodes $(n_1, n_2, \dots, n_k)$ let $X(n_1,\dots,n_k)$ denote the $k \times k$ principal submatrix of X defined by:

$$[X(n_1,\dots,n_k)]_{ij} \quad := \quad X_{ij}, \quad i,j \in \{n_1,\dots,n_k\}$$

Any maximal clique $q = (n_1, n_2, \dots, n_k)$ of G with k nodes defines a (fully specified) $k \times k$ principal submatrix denoted by $X(q) := X(n_1,\dots,n_k)$. In particular each edge $(i,j) \in E$

[1] We abuse the X_G notation: given G, X_G is a partial matrix defined on G, and given an $n \times n$ matrix X, X_G is a submatrix $(X_{jj}, j \in N, X_{jk}, (j,k) \in E)$ of X. The meaning should be clear from the context.

is a clique and defines a 2×2 principal submatrix $X(i,j)$, which we will use heavily in discussing optimal power flow problems.

We extend the notions of Hermitian, psd, rank-1, and trace to partial matrices as follows. We say that a partial matrix X_G is *Hermitian*, denoted by $X_G = X_G^H$, if $[X_G]_{kj} = ([X_G]_{jk})^H$. An $n \times n$ matrix X is psd if and only if all its principal submatrices (including X itself) is psd. We extend the notion of psd to partial matrices using this property, by saying that a partial matrix X_G is psd if all its "principal submatrices" that are fully specified are psd. Formally X_G is psd, denoted by $X_G \succeq 0$, if $X_G(q) \succeq 0$ for all maximal cliques q of G. Similarly we say that a partial matrix X_G is *rank-1*, denoted by rank $X_G = 1$, if $X_G(q)$ is rank-1 for all maximal cliques q of G. This means in particular that, since each edge $(j,k) \in E$ is a clique, if a partial matrix X_G is psd or rank-1 then it is automatically Hermitian. We say W_G is 2×2 *psd on* G, denoted by $W_G(j,k) \succeq 0$, if for all $(j,k) \in E$

$$\begin{aligned} W_G(j,j) \geq 0, \quad W_G(k,k) \geq 0 \\ W_G(j,j)\, W_G(k,k) \ \geq \ |W_G(j,k)|^2 \end{aligned} \tag{5}$$

We say W_G is 2×2 *psd rank-1 on* G, denoted by rank $W_G = 1$, if W_G is 2×2 psd and equality is attained in (5). Finally we say that an $n \times n$ matrix C is *defined on graph* G if $C_{jk} = 0$ if $(j,k) \notin E$. If C and X_G are defined on the same graph G then tr $CX_G = \sum_{j \in N} C_{jj}[X_G]_{jj} + \sum_{(j,k) \in E} C_{jk}[X_G]_{jk}$.

To simplify exposition suppose the matrices C_m in (3), $m = 0, \ldots, M$, are all defined on G, i.e., for all m, $[C_m]_{jk} = 0$ if $(j,k) \notin E$. Then for any $n \times n$ matrix X, tr $C_m X$ = tr $C_m X_G$. Conversely, given a partial matrix X_G that satisfies (3b), *any* completion X of X_G satisfies (3b). Even though both the objective function (3a) and the constraints (3b) depend only on a partial matrix X_G the constraint $X \succeq 0$ in (3c) depends also on entries not in X_G. Indeed the number of complex entries in X is n^2 while the number of complex variables in X_G is only $n + 2|E|$, which is much smaller than n^2 if G is large but sparse. Hence instead of solving for a full psd matrix X directly as in SDP (3) we would like to compute a partial matrix X_G that has a psd completion X that satisfies (3b)–(3c). If X is rank-1 then it also solves the problem (2) and hence yields a solution to the original QCQP (1) through spectral decomposition of X. Solving for such a partial matrix X_G is however difficult in general. In the next section we provide two characterizations of partial matrices that guarantee a psd rank-1 completion.

One of these characterizations is in terms of a chordal extension of G based on the following fundamental result in [39].

Theorem 1 ([39], Theorem 7): Fix a graph F. Every psd partial matrix X_F has a psd completion if and only if the graph F is chordal.

This result suggests a way to exploit the sparsity of graph G to solve SDP (3): solve for a partial matrix X_F defined on a chordal extension F of G instead of solving for the whole matrix $X \in \mathbb{S}^n_+$. Given a solution X_F^* that is psd we can compute a psd completion X^* guaranteed by Theorem 1 that solves SDP (3) using the algorithm in [39].

Given a chordal extension F we now formulate chordal relaxation as an SDP.

Chordal relaxation

Let $F = (N, E')$ be a chordal extension of G with $E' \supseteq E$. Let $q_1, \ldots, q_K$ be the set of maximal cliques of F and $X(q_k), k = 1, \ldots, K$, be the set of principal submatrices of X defined by these cliques. Consider the following problem where the optimization variable is the Hermitian partial matrix $W_F \in \mathbb{C}^{n+2|E'|}$ defined on the chordal extension F:

$$\min_{X_F = X_F^H} \quad \text{tr } C_0 X_G \tag{6a}$$

$$\text{subj. to} \quad \text{tr } C_m X_G \leq b_m, \quad m = 1, \ldots, M \tag{6b}$$

$$X_F(q_k) \succeq 0, \quad k = 1, \ldots, K \tag{6c}$$

We call this problem a *chordal relaxation* of QCQP (1). It is equivalent to SDP (3) in the sense that given any feasible solution X_F of (6), Theorem 1 guarantees a psd completion X that is feasible for (3); conversely given any feasible solution X of (3), its submatrix X_G is feasible for (6). Moreover these two problems have the same objective value, i.e., tr $C_0 X$ = tr $C_0 X_F$.

The first step in constructing the chordal relaxation (6) is to list all the maximal cliques q_k. Even though listing all maximal cliques of a general graph is NP-hard it can be done efficiently for a chordal graph. This is because a graph is chordal if and only if it has a perfect elimination ordering [40] and computing this ordering takes linear time in the number of nodes and edges [41]. Given a perfect elimination ordering all maximal cliques q_k can be enumerated and $X_F(q_k)$ constructed efficiently [35]. For optimal power flow problems the computation depends only on the topology of the power network, not on operational data, and therefore can be done offline.

Bus injection model

OPF formulation

Consider a power network modeled by a connected undirected graph $G(N,E)$ where each node in $N := \{1, 2, \ldots, n\}$ represents a bus and each edge in E represents a line. For each edge $(i,j) \in E$ let y_{ij} be its admittance. A bus $j \in N$ can have a generator, a load, both or neither. Traditionally the loads are specified and the generations are variables to be determined. Let s_j be the net complex power injection (generation minus load) at bus $j \in N$. Let V_j be the complex voltage at bus $j \in N$ and $|V_j|$ denote its magnitude. Bus 1 is the slack bus with a fixed magnitude $|V_1|$ (normalized to 1 p.u.). The *bus injection model* is defined by the following power flow equations that describe the Kirchhoff's law:

$$s_j \quad = \quad \sum_{k:(j,k)\in E} V_j(V_j^H - V_k^H)y_{jk}^H, \quad j \in N \tag{7}$$

The power injections satisfy

$$\underline{s}_j \le s_j \le \overline{s}_j, \quad j \in N \tag{8}$$

where $\underline{s}_j$ and $\overline{s}_j$ are given bounds on the net generation at bus j. If there is no upper bound on the load or on the generation at bus j then $\underline{s}_j = -\infty - \mathbf{i}\infty$ or $\overline{s}_j = \infty + \mathbf{i}\infty$ respectively. This is usually the case at the slack bus 1. We can eliminate the variables s_k from the OPF formulation by combining (7)–(8) into

$$\underline{s}_j \quad \le \sum_{k:(j,k)\in E} V_j(V_j^H - V_k^H)y_{jk}^H \quad \le \quad \overline{s}_j, \; j \in N \tag{9}$$

Then OPF in the bus injection model can be formulated in terms of just the voltage vector V. All voltage magnitudes are constrained:

$$\underline{v}_j \;\; \le \;\; |V_j|^2 \;\; \le \;\; \overline{v}_j, \quad j \in N \tag{10a}$$

where $\underline{v}_j$ and $\overline{v}_j$ are given lower and upper bounds on voltages. Usually $\underline{v}_1 = \overline{v}_1 = 1$. These constraints define the feasible set of the optimal power flow problem in the bus injection model:

$$\mathbb{V} := \{V \in \mathbb{C}^n \mid V \text{ satisfies } (9) - (10)\} \tag{11}$$

Let the cost function be $C(V)$. Typical costs include the cost of generating real power at each generator bus or line loss over the network. All these costs can be expressed as functions of V. Then the optimal power flow problem is:

OPF:

$$\min_V C(V) \quad \text{subject to} \quad V \in \mathbb{V} \tag{12}$$

Since (9) is quadratic, $\mathbb{V}$ is generally a nonconvex set. OPF is thus a nonconvex problem and NP-hard to solve in general.

Remark 1: OPF (12) as defined is a simplified version that ignores other important constraints such as line limits and security constraints (see e.g. [42]). Our model also ignores shunt elements. Some of these (e.g. shunt elements and line limits) can be incorporated without any change to the results in this paper.

OPF as QCQP

Before we describe convex relaxations of OPF we first show that, when $C(V) := V^H CV$ is quadratic in V, OPF (12) is indeed a QCQP by converting it into the standard form (1), following the derivation in [43].

To write (9) in the standard form (1a), define the $n \times n$ admittance matrix Y as

$$Y_{ij} \quad = \quad \begin{cases} \displaystyle\sum_{k:k\sim i} y_{ik}, & \text{if } i = j, \\ -y_{ij}, & \text{if } i \ne j \text{ and } i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

Y is symmetric but not necessarily Hermitian. Then

$$s_j \quad := \quad V_j I_j^H \;\; = \;\; (e_j^H V)(I^H e_j)$$

where e_j is the n-dimensional vector with 1 in the jth entry and 0 elsewhere. Hence, since $I = YV$, we have

$$\begin{aligned} s_j \;\; = \;\; \text{tr}\; \left(e_j^H V V^H Y^H e_j\right) \quad &= \quad \text{tr}\; \left(Y^H e_j e_j^H\right) V V^H \\ &= \quad V^H Y_j^H V \end{aligned}$$

where $Y_j^H := Y^H e_j e_j^H$. Y_j is an $n \times n$ matrix with its jth row equal to the jth row of the admittance matrix Y and all other rows equal to the zero vector. Y_j^H is in general not Hermitian so that $V^H Y_j^H V$ is in general a complex number and not in the standard form (1b). Its real and imaginary parts can be expressed in terms of the Hermitian and skew Hermitian components of Y_j^H defined as:

$$\Phi_j := \frac{1}{2}\left(Y_j^H + Y_j\right) \text{ and } \Psi_j := \frac{1}{2\mathbf{i}}\left(Y_j^H - Y_j\right) \tag{13}$$

Then

$$\text{Re } s_j = V^H \Phi_j V \quad \text{and} \quad \text{Im } s_j = V^H \Psi_j V$$

Let their upper and lower bounds be denoted by

$$\begin{aligned} \underline{p}_j := \text{Re } \underline{s}_j \quad &\text{and} \quad \overline{p}_j := \text{Re } \overline{s}_j \\ \underline{q}_j := \text{Re } \underline{s}_j \quad &\text{and} \quad \overline{q}_j := \text{Re } \overline{s}_j \end{aligned}$$

Let $J_j := e_j e_j^H$ denote the Hermitian matrix with a single 1 in the (j,j)th entry and 0 everywhere else. Suppose $C(V) := V^H CV$ for some Hermitian matrix C. Then OPF (12) can be written as a standard form QCQP:

$$\min_{x\in\mathbb{C}^n} \quad V^H CV \tag{14a}$$

$$\text{s.t.} \quad V^H \Phi_j V \le \overline{p}_j, \;\; V^H(-\Phi_j)V \le -\underline{p}_j \tag{14b}$$

$$V^H \Psi_j V \le \overline{q}_j, \;\; V^H(-\Psi_j)V \le -\underline{q}_j \tag{14c}$$

$$V^H J_j V \le -\overline{v}_j, \;\; V^H(-J_j)V \le -\underline{v}_j \tag{14d}$$

where $j \in N$ in (14).

Feasible sets

The cost function C in OPF is usually assumed to be convex. The difficulty of OPF (12) thus arises mainly from the nonconvex quadratic constraints. In this subsection we summarize the main results in [44], [45] that characterize the feasible set $\mathbb{V}$ of OPF in terms of partial matrices. These characterizations lead naturally to SDP, chordal, and SOCP relaxations of OPF.

Given a voltage vector $V \in \mathbb{V}$ we can define a partial matrix W_G by:

$$\begin{aligned} [W_G]_{jj} &:= |V_j|^2 \\ [W_G]_{jk} &:= V_j V_k^H \;\; =: \;\; [W_G]_{kj}^H \end{aligned}$$

Then the constraints (9)–(10) imply that the partial matrix W_G satisfies [2]

$$\underline{s}_j \leq \sum_{k:(j,k)\in E} \left([W_G]_{jj} - [W_G]_{jk}\right) y_{jk}^H \leq \overline{s}_j, j \in N \qquad (15a)$$
$$\underline{v}_j \leq [W_G]_{jj} \leq \overline{v}_j, \qquad j \in N \qquad (15b)$$

These constraints can also be written in a (partial) matrix form as:

$$\begin{array}{ccccc} \underline{p}_j & \leq & \text{tr } \Phi_j W_G & \leq & \overline{p}_j \\ \underline{q}_j & \leq & \text{tr } \Psi_j W_G & \leq & \overline{q}_j \\ \underline{v}_j & \leq & \text{tr } J_j W_G & \leq & \overline{v}_j \end{array}$$

The converse is not always true: given a partial matrix W_G that satisfies (15) it is not always possible to recover a voltage vector V in $\mathbb{V}$. It is clear that *any* completion W of such a W_G would also satisfy (15) since $y_{jk} = 0$ if $(j,k) \notin E$. We can recover a voltage vector $V \in \mathbb{V}$ if W happens to be psd rank-1 because in that case it can be uniquely decomposed into $W = VV^H$ with $V \in \mathbb{V}$.

We hence seek additional conditions on the partial matrix W_G that guarantee that it has a psd rank-1 completion W from which $V \in \mathbb{V}$ can be recovered. To this end we say that a partial matrix W_G satisfies the *cycle condition* if, for every cycle c in G,

$$\sum_{(i,j)\in c} \angle W_{ij} \;\; = \;\; 0 \mod 2\pi \qquad (16)$$

We recall that a partial matrix W_F is defined to be psd (or rank-1) if $W_F(q) \geq 0$ (or rank $W_F(q) = 1$) for all its submatrices $W_F(q)$ defined on all the maximal cliques q of F.

The following result characterizes when a definite matrix W is of rank 1 in terms of its restriction onto G or any chordal extension $c(G)$, i.e., in terms of its submatrices W_G and $W_{c(G)}$. It is proved in [44, Theorem 3], [45]. It implies that a partial matrix W_G has a psd rank-1 completion W if and only if W_G is 2×2 rank-1 on G and satisfies the cycle condition (16) if and only if it has a chordal extension $W_{c(G)}$ that is psd rank-1. It leads to three convex relaxations of OPF – SDP, chordal, and SOCP relaxations – as explained below.

Theorem 2: Fix a graph G on n nodes and any chordal extension $c(G)$ of G. Given an $n \times n$ positive or negative semidefinite matrix W the following are equivalent:

(1) rank $W = 1$.
(2) rank $W_{c(G)} = 1$.
(3) rank $W_G(i,j) = 1$ for all $(i,j) \in E$ and the partial matrix W_G satisfies the cycle condition (16).

[2] The constraint (15a) can also be written compactly in terms of the admittance matrix Y as [46]:

$$\underline{s} \;\leq\; \text{diag}\left(WY^H\right) \;\leq\; \overline{s}$$

Remark 2: As shown in [44] the set of completions of a partial matrix W_G that satisfies the condition in Theorem 2(3) can consist of a single positive semidefinite rank-1 matrix, or a single negative semidefinite rank-1 matrix, and infinitely many indefinite non-rank-1 matrices. For OPF we are interested only in a partial matrix W_G that is 2×2 psd ($W(i,j) \succeq 0$ for all $(i,j) \in E$), in addition to the condition in Theorem 2(3). Such a partial matrix W_G cannot have a negative definite completion. Hence it must have a *unique* psd rank-1 completion W (as well as indefinite non-rank-1 completions); see [44, Theorems 5 and 8] and discussions therein. We will construct this W explicitly below.

The rank-1 condition is a property of the whole matrix W. Theorem 2 characterizes this in terms of two different submatrices W_G and $W_{c(G)}$ of W, defined on graph G and its chordal extension $c(G)$ respectively. This is important because the submatrices are typically much smaller than W and can be much more efficiently computed for large sparse networks. The theorem thus allows us to solve simpler problems in terms of partial matrices, as we now explain.

Define the set of $n \times n$ Hermitian matrices:

$$\begin{aligned} \mathbb{W}_1 \;\; := \;\; \{W \in \mathbb{S}^n \mid \; & W_G \text{ satisfies (15)}, \\ & W \succeq 0, \;\text{rank } W = 1\} \end{aligned}$$

Fix any chordal extension $c(G)$ of G and define the set of Hermitian *partial* matrices $W_{c(G)}$:

$$\begin{aligned} \mathbb{W}_{c(G)} \;\; := \;\; \{W_{c(G)} \mid \; & W_G \text{ satisfies (15)}, \\ & W_{c(G)} \succeq 0, \;\text{rank } W_{c(G)} = 1\} \end{aligned}$$

where given a $W_{c(G)}$, W_G denotes the submatrix of $W_{c(G)}$ with off diagonal entries $[W_G]_{jk}$ defined only for $(j,k) \in E$. Finally define the set of Hermitian partial matrices W_G:

$$\begin{aligned} \mathbb{W}_2 \;\; := \;\; \{W_G \mid & W_G \text{ satisfies (15)}, (16), W_G(i,j) \succeq 0, \\ & \text{rank } W_G(i,j) = 1 \text{ for all } (i,j) \in E\} \end{aligned}$$

Remark 2 immediately implies

Corollary 3: Given a partial matrix $W_{c(G)} \in \mathbb{W}_{c(G)}$ or $W_G \in \mathbb{W}_2$ there is a unique psd rank-1 completion $W \in \mathbb{W}_1$.

We say two sets A and B are *equivalent*, denoted by $A \equiv B$, if there is a bijection between them. Even though $\mathbb{W}_1, \mathbb{W}_{c(G)}, \mathbb{W}_2$ are different (e.g. their matrices have different dimensions) Theorem 2 and Corollary 3 imply that they are all equivalent to the feasible set of OPF.

Corollary 4: $\mathbb{V} \equiv \mathbb{W}_1 \equiv \mathbb{W}_{c(G)} \equiv \mathbb{W}_2$.

Corollary 4 suggests three equivalent problems to OPF. We assume the cost function $C(V)$ in OPF depends on V only through the G-partial matrix W_G. For example if the cost is total real line loss in the network then $C(V) = \sum_j \text{Re } s_j \;=\; \sum_j \sum_{k:(j,k)\in E} \text{Re}\left([W_G]_{jj} - [W_G]_{jk}\right) y_{jk}^H$. If the cost is a weighted sum of real generation power then $C(V) = \sum_j \left(c_j \operatorname{Re} s_j + p_j^d\right)$ where p_j^d are the given real power demands at buses j; again $C(V)$ is a function of the G-partial matrix W_G. Then OPF (12) is equivalent to

$$\min_{W} C(W_G) \quad \text{subject to} \quad W \in \mathbb{W} \tag{17}$$

where $\mathbb{W}$ is any one of the sets $\mathbb{W}_1, \mathbb{W}_{c(G)}, \mathbb{W}_2$ and the variable W is of appropriate dimension. This makes use of the fact that these four problems have the same objective function and there is a bijection among their feasible sets. More specifically, given an optimal solution W^* in $\mathbb{W}_1$, since W is rank-1, it can be uniquely decomposed into $W^* = V^*(V^*)^H$. Then V^* is in $\mathbb{V}$ and an optimal solution of OPF (12). Alternatively given an optimal solution W_F^* in $\mathbb{W}_{c(G)}$ or $\mathbb{W}_2$, Corollary 3 guarantees that W_F^* has a psd rank-1 completion W^* in $\mathbb{W}_1$ from which an optimal $V^* \in \mathbb{V}$ can be recovered through spectral decomposition. For a partial matrix $W_G^* \in \mathbb{W}_2$ we will provide below an alternative, a much more direct, construction of $V^* \in \mathbb{V}$.

Convex relaxations

The difficulty with solving (17) is that the feasible sets $\mathbb{W}_1$, $\mathbb{W}_{c(G)}$, and $\mathbb{W}_2$ are still nonconvex due to the rank-1 constraints and the cycle condition (16). Their removal leads to the SDP, chordal, and SOCP relaxations of OPF respectively.

Relax $\mathbb{W}_1$ to a convex subset of $\mathbb{S}^n$:

$$\mathbb{W}_1^+ \;:=\; \{W \in \mathbb{S}^n \mid W_G \text{ satisfies (15)}, W \succeq 0\} \tag{18}$$

Relax $\mathbb{W}_{c(G)}$ to a convex set of Hermitian partial matrices:

$$\mathbb{W}_{c(G)}^+ \;:=\; \{W_{c(G)} \mid W_G \text{ satisfies (15)}, W_{c(G)} \succeq 0\} \tag{19}$$

Relax $\mathbb{W}_G$ to a convex set of Hermitian partial matrices by dropping both the 2×2 rank-1 condition and the cycle condition:

$$\begin{aligned}\mathbb{W}_2^+ \;:=\; \{W_G \mid\; & W_G \text{ satisfies (15)}, \\ & W_G(j,k) \succeq 0,\ (j,k) \in E\}\end{aligned}$$

Define the problems:

OPF-sdp:

$$\min_{W} C(W_G) \quad \text{subject to} \quad W \in \mathbb{W}_1^+ \tag{20}$$

OPF-ch:

$$\min_{W_{c(G)}} C(W_G) \quad \text{subject to} \quad W_{c(G)} \in \mathbb{W}_{c(G)}^+ \tag{21}$$

OPF-socp:

$$\min_{W_G} C(W_G) \quad \text{subject to} \quad W_G \in \mathbb{W}_2^+ \tag{22}$$

Since $\mathbb{W}_1 \subseteq \mathbb{W}_1^+$, $\mathbb{W}_{c(G)} \subseteq \mathbb{W}_{c(G)}^+$, $\mathbb{W}_2 \subseteq \mathbb{W}_2^+$, OPF-sdp, OPF-ch, OPF-socp provide lower bounds on the optimal value of OPF (12) in light of Corollary 4.

We make two comments on these semidefinite programs. First the condition $W_G(j,k) \succeq 0$ in the definition of $\mathbb{W}_2^+$ is equivalent to

$$[W_G]jj \geq 0,\ [W_G]_{kk} \geq 0,\ [W_G]_{jj}[W_G]_{kk} \geq |[W_G]_{jk}|^2$$

This is a second-order cone and hence OPF-socp is indeed an SOCP in the rotated form (4). Second OPF-ch is a convex chordal relaxation in the standard form (6). SOCP relaxation for OPF seems to be first observed in [47] for the bus injection model and in [48] for the branch flow model. Chordal relaxation for OPF is first proposed in [29].

For a mapping $f : A \rightarrow B$ let $f(A)$ denote the set $\{f(x) \mid x \in A\} \subseteq B$. For two sets A and B that are not necessarily in the same space we say that A is an *equivalent subset* of B, denoted by $A \sqsubseteq B$, if there is a mapping $f : A \rightarrow B$ such that $f(A) \subseteq B$ and f is a bijection from A to $f(A)$. Clearly $A \equiv B$ if and only if $A \sqsubseteq B$ and $B \sqsubseteq A$. The feasible set of OPF (12) is an equivalent subset of the feasible sets of the relaxations, as the following results from [44], [45] show.

Theorem 5 ([44], [45]): 1) If G is radial then $\mathbb{V} \sqsubseteq \mathbb{W}_1^+ \equiv \mathbb{W}_{c(G)}^+ \equiv \mathbb{W}_2^+$.
2) If G has cycles then $\mathbb{V} \sqsubseteq \mathbb{W}_1^+ \equiv \mathbb{W}_{c(G)}^+ \sqsubseteq \mathbb{W}_2^+$.

Let $C^*, C_{sdp}, C_{ch}, C_{socp}$ be the optimal values of OPF, OPF-sdp, OPF-ch, OPF-socp respectively. Theorem 5 and Corollary 4 directly imply

Corollary 6: 1) If G is radial then $C^* \geq C_{sdp} = C_{ch} = C_{socp}$.
2) If G has cycles then $C^* \geq C_{sdp} = C_{ch} \geq C_{socp}$.

We now comment on the computational aspect of these three relaxations. First the choice of the chordal extension $c(G)$ of G determines the number of variables in OPF-ch and hence the required computation effort, but it does not affect its optimal value. A good choice of $c(G)$ is nontrivial. In the worst case OPF-ch can require as much effort as OPF-sdp, but simulation results on IEEE test systems in [45] confirm that it can be much more efficiently solved than OPF-sdp when the network is large and sparse, as practical systems are. Indeed the numbers of lines in IEEE test systems (with 14, 30, 57, 118, 300 buses) are less than 1.6 times the numbers of buses, much less than the squares of them.

Second though all OPF-sdp, OPF-ch, and OPF-socp are convex and hence can be solved in polynomial time, SOCP in general requires a much smaller computational effort than SDP for large sparse networks. Indeed G is a subgraph of any chordal extension $c(G)$ which is a subgraph of the complete graph defined on N, and hence the number of complex variables (matrix entries) is the smallest in OPF-socp ($|W_G|$), the largest in OPF-sdp (n^2), with OPF-ch typically in between.

Finally, and most importantly, Corollary 6 suggests that, when G is a tree, we should *always* solve OPF-socp. When G has cycles then there is a tradeoff between computational effort and exactness in deciding between solving OPF-socp or OPF-sdp/OPF-ch. Between OPF-sdp and OPF-ch, OPF-ch seems much more preferable as they have the same accuracy (in terms of exactness) but OPF-ch is much faster to solve for large sparse networks, as discussed above.

Solution strategy

The general strategy to solving OPF (12) based on convex relaxation is illustrated in Figure 1. For OPF-sdp if the optimal

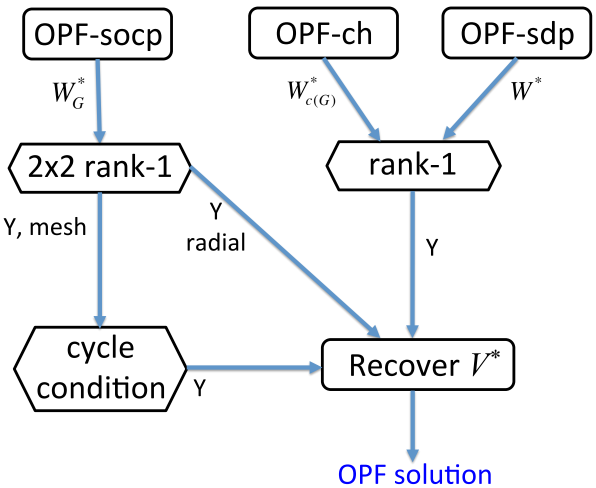

Fig. 1: Solving OPF (12) based on convex relaxations in the bus injection model.

solution W^* is rank-1 then an optimal solution V^* for OPF is recovered through the unique spectral decomposition $W^* = V^*(V^*)^H$. For OPF-ch if the optimal solution $W_{c(G)}$ is rank-1 then a psd rank-1 completion W^* can be computed from $W_{c(G)}$ according to Theorem 1 from which an optimal V^* can be recovered.

For OPF-socp if the optimal solution W_G is 2×2 rank-1 on G and satisfies the cycle condition (16) (a radial network automatically satisfies (16)) then a unique V^* can be constructed explicitly as follows. Let T be an arbitrary spanning tree of G. Let $\mathbb{P}_{ij}$ denote the unique path from node i to node j in T. For $i = 1, \ldots, n$, let

$$|V_i^*| \;:=\; \sqrt{[W_G]_{ii}} \tag{23}$$

$$\angle V_i^* \;:=\; \sum_{(j,k)\in\mathbb{P}_{1i}} \angle\,[W_G]_{jk} \tag{24}$$

Without loss of generality set $\angle V_1 = 0$ in this construction (otherwise, add $\angle V_1$ to (24)). Then it can be checked that V^* is an optimal solution of OPF (12).

The key for this solution strategy is that the convex relaxations are exact so that an V^* can be recovered. Formally if *every* optimal solution W^* of OPF-sdp is rank-1, and hence also effectively solves OPF, then we say that OPF-sdp is *exact*. Similarly if every optimal solution $W_{c(G)}^*$ of OPF-ch is rank-1 then we say OPF-ch is *exact*. If every optimal solution W_G^* of OPF-socp is 2×2 rank-1 and satisfies the cycle condition (16) then we say OPF-socp is *exact*.

In the following subsections we summarize results from [49], [50], [43], [51], [47], [52] that provide sufficient conditions for the relaxations to be exact, i.e., for the first inequality in Corollary 6 to attain equality. We will present these conditions for two special cases: AC radial networks and DC mesh networks. AC radial networks are important as most distribution systems are radial.

Exact relaxation: AC radial

As discussed previously we should always solve SOCP if the network graph G is a tree. We hence focus in this subsection on the exactness of OPF-socp (22).

Separating line Fix an undirected graph $G = (N, E)$. Fix Hermitian matrices $C_m \in \mathbb{S}^n$, $m = 0, \ldots, M$, defined on G, i.e., $[C_m]_{jk} = 0$ if $(j,k) \notin E$. Consider QCQP:

$$\min_{x\in\mathbb{C}^n} \quad x^H C_0 x \tag{25a}$$

$$\text{s.t.} \quad x^H C_m x \le b_m, \;\; m = 1, \ldots, M \tag{25b}$$

and its SOCP relaxation where the optimization variable ranges over Hermitian partial matrices W_G:

$$\min_{W_G = W_G^H} \quad \text{tr}\,(C_0 W) \tag{26a}$$

$$\text{s.t. tr}\; C_m W_G \le b_m, \;\; m = 1, \ldots, M \tag{26b}$$

$$W_G(i,j) \succeq 0, \;\; (i,j) \in E \tag{26c}$$

The following result is proved in [43]. We assume[3]

A1: G is a tree.
A2: For each link $(j,k) \in E$ there exists an α_{jk} such that $\angle\,[C_m]_{jk} \in [\alpha_{ij}, \alpha_{ij} + \pi]$ for all $m = 0, \ldots, M$.

Let C^*, C_{socp} denote the optimal values of QCQP (25) and SOCP (26) respectively.

Theorem 7 ([43]): Suppose A1–A2 holds. Then SOCP is exact, i.e., $C^* = C_{socp}$ and an optimal solution x^* of QCQP (25) can be obtained from every optimal solution W_G^* of SOCP (26).

We now apply Theorem 7 to our OPF problem (14). When the network is radial the condition A2 in the theorem implies a simple pattern on the power injection constraints (14b)–(14c). A2 depends only on the off-diagonal entries of C, Φ_j, Ψ_j (J_j is a digonal matrix). Let $y_{jk} = g_{jk} - \mathbf{i}b_{jk}$ with $g_{jk} > 0, b_{jk} > 0$. Then we have from (13)

$$[\Phi_k]_{ij} = \begin{cases} \frac{1}{2}Y_{ij} \;=\; -\frac{1}{2}(g_{ij} - \mathbf{i}b_{ij}) & \text{if } k = i \\ \frac{1}{2}Y_{ij}^* \;=\; -\frac{1}{2}(g_{ij} + \mathbf{i}b_{ij}) & \text{if } k = j \\ 0 & \text{if } k \notin \{i,j\} \end{cases}$$

$$[\Psi_k]_{ij} = \begin{cases} \frac{-1}{2\mathbf{i}}Y_{ij} \;=\; -\frac{1}{2}(b_{ij} + \mathbf{i}g_{ij}) & \text{if } k = i \\ \frac{1}{2\mathbf{i}}Y_{ij}^* \;=\; -\frac{1}{2}(b_{ij} - \mathbf{i}g_{ij}) & \text{if } k = j \\ 0 & \text{if } k \notin \{i,j\} \end{cases}$$

[3]All angles should be interpreted "mod 2π", i.e., projected onto $[0, 2\pi)$.

Hence for each line $(j,k) \in E$ the relevant angles for A2 are those of C_{jk} and

$$[\Phi_j]_{jk} = -\frac{1}{2}(g_{jk} - \mathbf{i}b_{jk})$$

$$[\Phi_k]_{jk} = -\frac{1}{2}(g_{jk} + \mathbf{i}b_{jk})$$

$$[\Psi_j]_{jk} = -\frac{1}{2}(b_{jk} + \mathbf{i}g_{jk})$$

$$[\Psi_k]_{jk} = -\frac{1}{2}(b_{jk} - \mathbf{i}g_{jk})$$

as well as the angles of $-[\Phi_j]_{jk}, -[\Phi_k]_{jk}$ and $-[\Psi_j]_{jk}, -[\Psi_k]_{jk}$. These quantities are shown in Figure 2 with their magnitude *normalized* to a common value and explained in the caption of the figure.

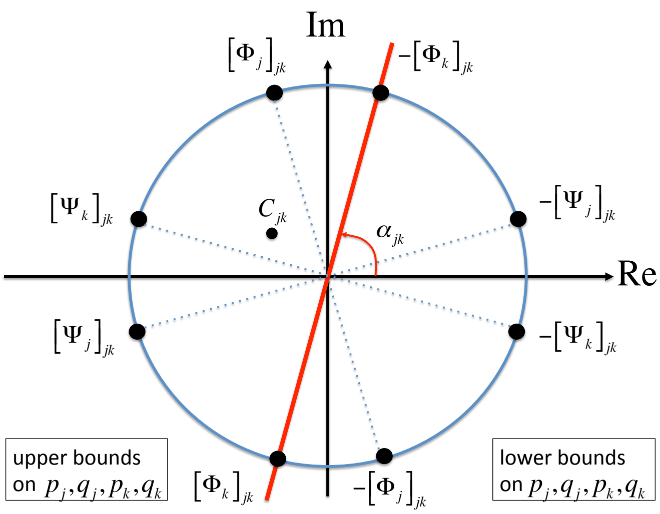

Fig. 2: Condition A2 on a line $(j,k) \in E$. The quantities $([\Phi_j]_{jk}, [\Phi_k]_{jk}, [\Psi_j]_{jk}, [\Psi_k]_{jk})$ on the left-half plane correspond to finite upper bounds on (p_j, p_k, q_j, q_k) in (14b)–(14c); $(-[\Phi_j]_{jk}, -[\Phi_k]_{jk}, -[\Psi_j]_{jk}, -[\Psi_k]_{jk})$ on the right-half plane correspond to finite lower bounds on (p_j, p_k, q_j, q_k). A2 is satisfied if there is a line through the origin, characterized by the angle α_{jk}, so that the quantities corresponding to *finite* upper or lower bounds on (p_j, q_j) and (p_k, q_k) lie on one side of the line, including on the line itself. The load over-satisfaction condition in [50], [47] and in [53, Theorem 7] corresponds to the Im-axis that excludes all quantities on the right-half plan (no lower bounds on power injections). The sufficient condition in [46, Theorem 2] corresponds to the red line in the figure that allows a finite lower bound on p_k.

The condition A2 applied to OPF (14) takes the following form (see Figure 2):

A2': For each link $(j,k) \in E$ there is a line in the complex plane through the origin such that $[C]_{jk}$ as well as those $\pm[\Phi_i]_{jk}$ and $\pm[\Psi_i]_{jk}$ corresponding to *finite* lower or upper bounds on (p_i, q_i), for $i = j,k$, are all on one side of the line, including on the line itself.

Corollary 8: Suppose G is a tree. If A2' holds then OPF-socp is exact, i.e., $C^* = C_{socp}$ and an optimal solution V^* of OPF (12) can be obtained from every optimal solution W_G^* of OPF-socp (22).

This is proved in [43] which also includes constraints on real branch power flows and line losses (see also the long version of [54] for a different proof). Corollary 8 includes several sufficient conditions for exact relaxation as special cases. For instance the load over-satisfaction condition in [50], [47] and in [53, Theorem 7] corresponds to the Im-axis that excludes all quantities on the right-half plan (no lower bounds on power injections). The sufficient condition in [46, Theorem 2] corresponds to the red line in Figure 2 that allows a finite lower bound on the real power injection at one end of a line (p_j or p_k but not both), and no finite lower bound on reactive power injections q_j. The approach in the independent works [50], [51], [46] proves that OPF-sdp (20) is exact, not by showing that an optimal W^* would be rank-1, but by showing a certain matrix A^* in the complementary slackness condition tr $A^*W^* = 0$ for SDP has rank $n-1$, as suggested in [49]. They make use of the fact that, if an $n \times n$ Hermitian matrix M is positive semidefinite and its underlying graph is a tree, then rank $M \geq n-1$; see e.g., [55], [56, Theorem 3.4] and [57, Corollary 3.9]. The complementary slackness condition tr $A^*W^* = 0$ then implies rank $W^* = 1$. The proof in [53] also makes use of the geometry of the power injection region, first explored in [51], to which we now turn.

Pareto front When the voltage magnitudes are fixed [46], [53], [58] provide a geometric insight on why convex relaxations are exact. For simplicity we will explain the intuition using the result in [53] for the OPF problem where $|V_i|$ are given for all $i \in N$ and reactive powers are ignored (the objective function and the constraints depend only on the real power injections $p_j, j \in N$). This result is extended to include reactive power in [58, Theorem 1] with fixed $|V_i|$ where an additional constraint is imposed on the lower bounds of reactive power injections to ensure these lower bounds are never tight. The case of variable $|V_i|$ without reactive power is considered in [53, Theorem 7] but the exact relaxation result there requires the load over-satisfaction condition and is therefore a special case of Corollary 8 with line limits.

Recall that $y_{jk} = g_{jk} - \mathbf{i}b_{jk}$ with $g_{jk} > 0, b_{jk} > 0$. Consider the following OPF problem:

$$\min_{p,P,V} \quad C(p) \tag{27a}$$
$$\text{s.t.} \quad \underline{p}_j \leq p_j \leq \overline{p}_j \tag{27b}$$
$$\underline{\theta}_{jk} \leq \theta_{jk} \leq \overline{\theta}_{jk} \tag{27c}$$
$$p_j = \sum_{k:k\sim j} P_{jk} \tag{27d}$$
$$P_{jk} = |V_j|^2 g_{jk} - |V_j||V_k| g_{jk} \cos\theta_{jk} \tag{27e}$$
$$+|V_j||V_k| b_{jk} \sin\theta_{jk} \tag{27f}$$
$$|V_j| = 1 \tag{27g}$$

where $j \in N$ in the above constraints, $\theta_{jk} := \theta_j - \theta_k$, and θ_j are the phase angles of V_j, i.e., $V_j = |V_j|\, e^{\mathbf{i}\theta_j}$.

We first comment on the constraints on angles θ_{jk} in (27). When the voltage magnitudes $|V_i|$ are fixed, constraints on real power flows, branch currents, line losses, and stability can all be represented in terms of θ_{jk}. Indeed a line flow constraint of the form $|P_{jk}| \leq \overline{P}_{jk}$ becomes a constraint on θ_{jk} using the expression for P_{jk} in (27). A constraint on current of the form $|I_{jk}| \leq \overline{I}_{jk}$ is also a constraint on θ_{jk} since $I_{jk} = y_{jk}(V_j - V_k)$ and $|V_j|, |V_k|$ are fixed. The line loss over $(j,k) \in E$ is equal to $P_{jk} + P_{kj}$ which is again a function of θ_{jk}. Stability typically requires $|\theta_{jk}|$ to stay within a small threshold. Therefore given constraints on branch power or current flows, losses, and stability, appropriate bounds $\underline{\theta}_{jk}, \overline{\theta}_{jk}$ can be determined in terms of these constraints. We assume this has been done and $\underline{\theta}_{jk}, \overline{\theta}_{jk}$ are given to ensure the satisfaction of these constraints.

We then express (27) as an optimization over only the bus injections $p := (p_j, j \in N)$ by using the equality constraints to eliminate the real branch power flows $P := (P_{jk}, (j,k) \in E)$, voltage magnitudes $|V| := (|V_j|, j \in N)$, and angles $\theta := (\theta_{jk}, (j,k) \in E)$. Define the injection region

$$\mathbb{P}_\theta \;:=\; \Bigg\{ p \in \mathbb{R}^n \mid p_j = \sum_{k:k\sim j} P_{jk}, \quad P_{jk} = g_{jk} - g_{jk}\cos\theta_{jk} + b_{jk}\sin\theta_{jk}, \quad \underline{\theta}_{jk} \leq \theta_{jk} \leq \overline{\theta}_{jk},\ (j,k) \in E \Bigg\} \tag{28}$$

Let $\mathbb{P}_p := \{p \in \mathbb{R}^n \,|\, \underline{p}_j \leq p_j \leq \overline{p}_j, j \in N\}$. Then (27) takes the form:

$$\min_p\ C(p) \quad \text{s.t.} \quad p \in \mathbb{P}_\theta \cap \mathbb{P}_p \tag{29}$$

This problem is hard because the set $\mathbb{P}_\theta$ defined in (28) is nonconvex.

For a set A let conv A denote the convex hull of A. A convex relaxation of (29) enlarges the nonconvex feasible set $\mathbb{P}_\theta \cap \mathbb{P}_p$ of (29) to a convex set:

$$\min_p\ C(p) \quad \text{s.t.} \quad p \in \text{conv}(\mathbb{P}_\theta) \cap \mathbb{P}_p \tag{30}$$

When $C(p)$ is linear in p this problem is SDP/SOCP [53]. For simplicity of exposition we will make this assumption even though the insight is about feasible sets and holds more generally than linear objective functions. Note that the feasible set of the relaxation (30) is generally a superset of the convex hull of the feasible set of (29) because conv and intersection do not commute, i.e.,

$$\text{conv}(\mathbb{P}_\theta) \cap \mathbb{P}_p \quad \supseteq \quad \text{conv}(\mathbb{P}_\theta \cap \mathbb{P}_p)$$

The key insight of [53], [58] shows that these two convex sets have the same Pareto front which turns out to coincide with the Pareto front of the nonconvex feasible set $\mathbb{P}_\theta \cap \mathbb{P}_p$, provided the θ_{jk} are suitably bounded.

More precisely, we say that a point $x \in A \subseteq \mathbb{R}^n$ is a *Pareto optimal point* in A if there does not exist another $x' \in A$ such that $x' \leq x$ with at least one strictly smaller component $x'_j < x_j$. The *Pareto front of* A, denoted by $\mathbb{O}(A)$, is the set of all Pareto optimal points in A. The significance of $\mathbb{O}(A)$ is that, for any increasing function, its minimizer, if exists, is necessarily in $\mathbb{O}(A)$ whether A is convex or not. If A is convex then x^* is a Pareto optimal point in $\mathbb{O}(A)$ if and only if there is a vector $c := (c_1, \ldots, c_n) \geq 0$ such that x^* is a minimizer of $c^T x$ over A.

The following result says that the sets of optimal points of (29) and its relaxation (30) are identical, implying that SDP/SOCP is exact when $C(p)$ is linear.

Theorem 9 ([53], [58]): Suppose G is a tree and for all $(j,k) \in E$

$$-\tan^{-1}\frac{b_{jk}}{g_{jk}} < \underline{\theta}_{jk} \leq \overline{\theta}_{jk} < \tan^{-1}\frac{b_{jk}}{g_{jk}} \tag{31}$$

If (29) is feasible then

(1) $\mathbb{O}(\mathbb{P}_\theta \cap \mathbb{P}_p) \;=\; \mathbb{O}(\,\text{conv}(\mathbb{P}_\theta) \cap \mathbb{P}_p\,)$.
(2) The SDP/SOCP relaxation is exact, i.e., every optimal solution p^* of (30) is also optimal for (29).

Theorem 9(1) implies that every optimal solution of (30) is feasible for (29) and hence SDP/SOCP relaxation is exact. The key observation that establishes Theorem 9(1) is $\mathbb{O}(\mathbb{P}_\theta) = \mathbb{O}(\text{conv}(\mathbb{P}_\theta))$.

Exact relaxation: DC mesh

In this subsection we consider purely resistive network, i.e., the impedance/admittance $z_{jk} = 1/y_{jk} = r_{jk}$, the power injections $s_j = p_j$, and the voltages V_j are all real.

Suppose the cost function of OPF depends only on the power injections $p := (p_j, j \in N)$ and is separable, i.e., $C(p) = \sum_j C_j(p_j)$.

Theorem 10 ([49], [59]): OPF-socp and OPF-sdp are exact if, for all $j \in N$, $\underline{p}_j = -\infty$ and $C_j(p_j)$ are strictly increasing.

The exactness of SDP relaxation for DC networks is first proved in [49] using a duality argument. The DC nature of the problem allows the application of Perron-Frobenius theorem to a irreducible matrix with nonpositive off-diagnoal elements. The exactness of SOCP is reported in [59] using [38, Theorem 3.1]. A similar result is also provided in Theorem 16 below for the branch flow model.

The following result is proved using a different technique.

Theorem 11 ([52]): Fix V_1. OPF-socp is exact if $\overline{v}_j = \infty$ for all $j \in N \setminus \{1\}$ and $C_1(p_1)$ is strictly increasing. Moreover if OPF-socp is exact then its solution is unique.

Branch flow model

OPF formulation

In the branch flow model we adopt a directed connected graph $\tilde{G} = (N, \tilde{E})$ to represent a power network where each node in $N := \{1, \ldots, n\}$ represents a bus and each edge in $\tilde{E}$ represents a line. Fix an arbitrary orientation for G and let $m := |\tilde{E}|$ be the number of directed edges in G. Denote an edge by (i, j) or $i \to j$ if it points from node i to node j. For each edge $(i, j) \in \tilde{E}$ let $z_{ij} := 1/y_{ij}$ be the complex impedance on the line, let I_{ij} be the complex current from buses i to j, and $S_{ij} = P_{ij} + \mathbf{i}Q_{ij}$ be the *sending-end* complex power from buses i to j. For each node $i \in N$ let V_i be the complex voltage at bus i. Let s_i be the net complex power injection (generation minus load) at bus i. We use s_i to denote both the complex number $p_i + \mathbf{i}q_i$ and the real pair (p_i, q_i) depending on the context.

The *branch flow model* is defined by the following set of power flow equations:

$$I_{jk} = y_{jk}(V_j - V_k), \quad j \to k \in \tilde{E} \tag{32a}$$
$$S_{jk} = V_j \; I_{jk}^H, \quad j \to k \in \tilde{E} \tag{32b}$$
$$s_j = \sum_{k:j\to k} S_{jk} - \sum_{i:i\to j} \left(S_{ij} - z_{ij}|I_{ij}|^2 \right), \; j \in N \tag{32c}$$

where (32a) describes the Ohm's law, (32b) defines branch power in terms of voltage and current, and (32c) imposes power balance at each bus where $z_{ij}|I_{ij}|^2$ represents the line loss so that $S_{ij} - z_{ij}|I_{ij}|^2$ is the receiving-end complex power at bus j from bus i. As in the bus injection model the power injections satisfy

$$\underline{s}_j \le s_j \le \overline{s}_j, \qquad j \in N \tag{33}$$

where $\underline{s}_j$ and $\overline{s}_j$ are given. We often assume $\underline{s}_1 = -\infty - \mathbf{i}\infty$ and $\overline{s}_1 = \infty + \mathbf{i}\infty$. We can eliminate the variables s_j by combining (32a) and (33) into

$$\underline{s}_j \;\le\; \sum_{k:j\to k} S_{jk} \;-\sum_{i:i\to j} \left(S_{ij} - z_{ij}|I_{ij}|^2 \right) \;\le\; \overline{s}_j, \qquad j \in N \tag{34}$$

All voltage magnitudes are constrained:

$$\underline{v}_j \le |V_j|^2 \le \overline{v}_j, \quad j \in N \tag{35}$$

where $\underline{V}_j$ and $\overline{V}_j$ are given. We often assume $\underline{v}_1 = \overline{v}_1 = 1$.

Denote the variables in the branch flow model by $\tilde{x} := (S, I, V) \in \mathbb{C}^{n+2m}$. These constraints define the feasible set of the OPF problem in the branch flow model:

$$\mathbb{X} := \{\tilde{x} \in \mathbb{C}^{n+2m} \mid \tilde{x} \text{ satisfies } (32a), (32b), (34), (35)\} \tag{36}$$

Let the cost function in the branch flow model be $C(\tilde{x})$. Then the optimal power flow problem in the branch flow model is:
OPF:

$$\min_{\tilde{x}} \;\; C(\tilde{x}) \quad \text{subject to} \quad \tilde{x} \in \mathbb{X} \tag{37}$$

Since (32) is quadratic, $\mathbb{X}$ is generally a nonconvex set. OPF is thus a nonconvex problem and NP-hard to solve in general.

SOCP relaxation

The cost function C is usually convex so the difficulty of OPF is due to the nonconvexity of the feasible set $\mathbb{X}$. Following [60] we now enlarge $\mathbb{X}$ into a second-order cone in two steps, leading to an SOCP relaxation of (37).

Angle relaxation. First we eliminate the phase angles from the complex voltages V and currents I. Formally this defines a mapping $h : \mathbb{C}^{n+2m} \to \mathbb{R}^{n+3m}$ that maps an $\tilde{x} = (S, I, V)$ to $h(\tilde{x}) = x := (S, \ell, v)$ with $\ell_{jk} = |I_{jk}|^2$ and $v_j = |V_j|^2$.[4] This set of new variables x satisfies, for $(j, k) \in \tilde{E}$ (eliminating the phase angles of the complex voltages V and currents I from (32a)–(32b)),

$$v_k = v_j - 2 \, \text{Re} \, (z_{jk}^H S_{jk}) + |z_{jk}|^2 \ell_{jk} \tag{38}$$
$$\ell_{jk} v_j = |S_{jk}|^2 \tag{39}$$

in addition to (32a). This is the model first proposed by Baran-Wu in [21], [22] for distribution systems. Define the set of solutions to the Baran-Wu model that also satisfy the OPF constraints as:

$$\mathbb{X}_2^{nc} := \{x \in \mathbb{R}^{n+3m} \mid x \text{ satisfies } (38), (39) \; (34), (35)\}$$

The mapping h relaxes a voltage V_j or current I_{jk} from a point in the complex plane into a circle with its magnitude as the radius of the circle. While (32) specifies $n + 2m$ (nonlinear) equations in $n + 2m$ complex variables, the Baran-Wu model (38), (39), (34), (35) specifies $2(n + m)$ equations in $n + 3m$ real variables. Since $m \ge n - 1$ ($\tilde{G}$ is connected), there are generally insufficient number of equations to determine uniquely all the variables x when $\tilde{G}$ contains cycles. When $\tilde{G}$ is a tree (which is the case in [21], [22]), v_1 is set to 1, and $s_1 := (p_1, q_1)$ are variables to be determined, then $m = n - 1$ and both the number of equations and the number of real variables become $4n - 1$. Indeed the results in [60] shows that $h : \mathbb{X} \to \mathbb{X}_2^{nc}$ is bijective when $\tilde{G}$ is a tree. Otherwise $h(\mathbb{X}) \subsetneq \mathbb{X}_2^{nc}$, i.e., h is not surjective on $\mathbb{X}_2^{nc}$. We now characterize the subset $h(\mathbb{X})$ of $\mathbb{X}_2^{nc}$ over which h is surjective and construct the inverse h^{-1} of h.

Given an $x := (S, \ell, v) \in \mathbb{R}^{n+3m}$ define $\beta := \beta(x) \in \mathbb{R}^m$ by

$$\beta_{ij}(x) := \angle \left(v_i - z_{ij}^H S_{ij} \right), \quad (i, j) \in \tilde{E} \tag{40}$$

Even though x does not include phase angles of V it turns out that x "implies" a phase angle difference across each line $(i, j) \in \tilde{E}$ given by $\beta_{ij}(x)$ [60, Theorem 2]. We are interested in the set of x such that $\beta_{ij}(x)$ can be expressed as $\theta_i - \theta_j$ for some voltage angles θ. To this end let B be the $m \times n$ (transposed) incidence matrix of $\tilde{G}$ defined as

$$B_{ei} = \begin{cases} 1 & \text{if edge } e \in \tilde{E} \text{ leaves node } i \\ -1 & \text{if edge } e \in \tilde{E} \text{ enters node } i \\ 0 & \text{otherwise} \end{cases}$$

[4] We abuse notation and use S both to denote a complex variable in $\mathbb{C}^m$ and as a shorthand for the real variables (P, Q) in $\mathbb{R}^{2m}$.

Consider the set of x such that

$$\exists\theta(x) \text{ that solves } \begin{cases} B\theta &= \beta(x) \mod 2\pi \\ \theta_1 &= 0 \end{cases} \quad (41)$$

i.e., $\beta(x)$ is in the range space of B (mod 2π).[5] A solution $\theta(x)$, if exists, is unique in $[-\pi,\pi)^n$. With such a $\theta(x)$ the inverse mapping $h^{-1} : \mathbb{R}^{n+3m} \to \mathbb{C}^{n+2m}$ is defined by $h^{-1}(S,\ell,v) = (S,I,V)$ where

$$V_j := \sqrt{v_j}\; e^{\mathbf{i}\theta_j(x)}, \quad j \in N \quad (42a)$$
$$I_{jk} := \sqrt{\ell_{jk}}\; e^{\mathbf{i}(\theta_j(x)-\angle S_{jk})}, \; (j,k)\in\tilde{E} \quad (42b)$$

Define the set:

$$\mathbb{X}_2 := \mathbb{X}_2^{nc} \cap \{x \in \mathbb{R}^{n+3m} \mid x \text{ satisfies } (41)\}$$

It is shown in [60] that, given any x, (41) either has no solution or has at most one solution $\theta(x)$ in $[-\pi,\pi)^n$. This means that h is a bijection from $\mathbb{X}$ to $\mathbb{X}_2$ with its inverse h^{-1} defined by (40)–(42). Hence $\mathbb{X} \equiv \mathbb{X}_2$; see Figure 3.

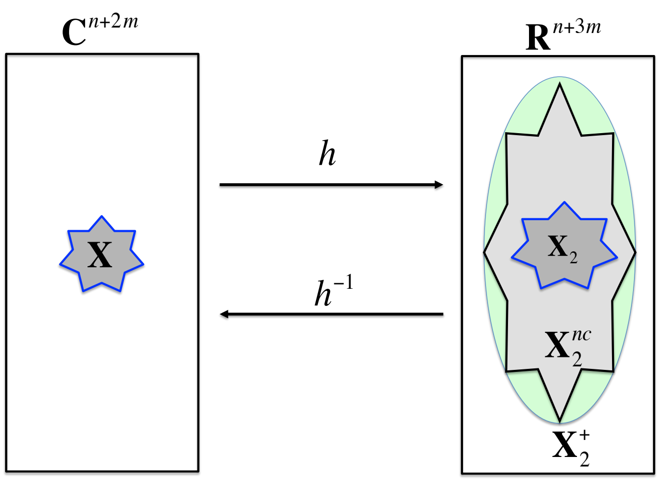

Fig. 3: Feasible set $\mathbb{X}$ of OPF in the branch flow model, the equivalent set $\mathbb{X}_2$ (defined by the bijection h) and its relaxations $\mathbb{X}_2^{nc}$, $\mathbb{X}_2^+$.

SOCP relaxation. The set $\mathbb{X}_2^{nc}$ is still nonconvex because of the quadratic equalities in (39). The second step of the relaxation relaxes them to inequalities:

$$\ell_{jk}v_j \geq |S_{jk}|^2, \qquad (j,k)\in\tilde{E} \quad (43)$$

Define the set:

$$\mathbb{X}_2^+ := \{x \in \mathbb{R}^{n+3m} \mid x \text{ satisfies } (38),(43)\; (34),(35)\}$$

Clearly $\mathbb{X}_2 \subseteq \mathbb{X}_2^{nc} \subseteq X_2^+$; see Figure 3.

These three sets define the following three optimization problems. We assume the cost function $c(\tilde{x})$ in OPF (37) depends on $\tilde{x} \in \mathbb{C}^{n+2m}$ only through $x \in \mathbb{R}^{n+3m}$. For example if the cost is total real line loss in the network then $c(\tilde{x}) = \sum_{(j,k)\in\tilde{E}} \text{Re}\, z_{jk}\ell_{jk}$. If the cost is a weighted sum of real generation power then $c(\tilde{x}) = \sum_j (c_j p_j + p_j^d)$ where p_j are the real parts of s_j in (32c) and p_j^d are the given real power demands at buses j; again $c(\tilde{x})$ depends only on x. Consider:

OPF:

$$\min_x \; c(x) \quad \text{subject to} \quad x \in \mathbb{X}_2$$

OPF-nc:

$$\min_x \; c(x) \quad \text{subject to} \quad x \in \mathbb{X}_2^{nc}$$

OPF-socp:

$$\min_x \; c(x) \quad \text{subject to} \quad x \in \mathbb{X}_2^{+}$$

Let C^* be the optimal cost of OPF (37) in the branch flow model. Let C_{opf}, C_{nc}, C_{socp} be the optimal costs of OPF, OPF-nc, OPF-socp respectively. The following result follows directly from [60, Theorems 2, 4].

Theorem 12 ([60]): For general networks

1) $\mathbb{X} \equiv \mathbb{X}_2 \subseteq \mathbb{X}_2^{nc} \subseteq \mathbb{X}_2^+$.
2) $C^* = C_{opf} \geq C_{nc} \geq C_{socp}$.

If $\tilde{G}$ is radial then

1) $\mathbb{X} \equiv \mathbb{X}_2 = \mathbb{X}_2^{nc} \subseteq \mathbb{X}_2^+$.
2) $C^* = C_{opf} = C_{nc} \geq C_{socp}$.

We say OPF-socp is *exact (with respect to OPF)* if every optimal solution x^* of OPF-socp attains equality in (43) and satisfies (41) so that an optimal solution to OPF can be recovered. An algorithm similar to that presented in the previous subsection on Solution Strategy can be used to compute an optimal OPF solution from x^*. Algebraically, for any optimal solution x^* of OPF-socp that attains equality in (43) and satisfies (41), there is a unique $\theta(x^*)$ in $[-\pi,\pi)^n$. The inverse $\tilde{x} := h^{-1}(x^*)$ is an optimal point for OPF.

We now turn to sufficient conditions under which OPF-socp is exact. We present the results first for AC radial networks and then for DC mesh networks. Finally we show that the case of AC mesh networks can be reduced to the case of AC radial networks using phase shifters.

Exact relaxation: AC radial

Recall that

$$s_j := (p_j, q_j) =: (p_j^g - p_j^c, q_j^g - q_j^c)$$

[5] The condition (41) on x has a familiar interpretation: the voltage angle differences implied by x sum to zero (mod 2π) around any cycle [60, Theorem 2(2)].

denote the net power injections at buses j. From (32c) they are given in terms of x by:

$$\begin{aligned} p_j^g - p_j^c &= \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} \left(P_{ij} - r_{ij}|I_{ij}|^2\right), \; j \in N \\ q_j^g - q_j^c &= \sum_{k:j\to k} Q_{jk} - \sum_{i:i\to j} \left(Q_{ij} - x_{ij}|I_{ij}|^2\right), \; j \in N \end{aligned}$$

Assume

A3: The cost function $C(x)$ is convex, strictly increasing in ℓ, nonincreasing in load (p^c, q^c), and independent of branch flows (P, Q).

To avoid triviality we assume that the problem OPF is feasible.

Theorem 13 ([48]): Suppose A3 holds. If $\overline{p}_j^c = \overline{q}_j^c = \infty$, $j \in N$, then OPF-socp is exact.

The next set of sufficient conditions, proved in [61], [62], allow finite upper bounds on loads (finite lower bounds on power injections), but remove upper limits on voltage magnitudes. To state them precisely, define $\mathbb{T}_j$ as the subtree rooted at bus j including bus j. Define the total net injections in a subtree as:

$$\begin{aligned} S_{jk}^{\text{lin}}(p,q) &:= P_{jk}^{\text{lin}}(p) + \mathbf{i}\, Q_{jk}^{\text{lin}}(q) \\ &:= \sum_{i\in\mathbb{T}_k} p_i \;+\; \mathbf{i} \sum_{i\in\mathbb{T}_k} q_i \end{aligned}$$

The quantity $-S_{jk}(p,q)$ represents the net load in the subtree $\mathbb{T}_j$ and is a lower bound on the real and reactive sending-end power S_{jk} on branch $(j,k) \in \tilde{E}$. It is a linear approximation in that S_{jk} equals $-S_{jk}(p,q)$ plus the real and reactive losses on the line $j \to k$ and those in $\mathbb{T}_k$. Let $\mathbb{P}_j$ be the unique path from bus 1 to bus j. Let $a_1^1 := 1, a_1^2 := 0, a_1^3 := 0, a_1^4 := 1$ and define for $i = 2, \ldots, N$:

$$\begin{aligned} a_i^1 &:= \prod_{(j,k)\in\mathbb{P}_i} \left(1 - \frac{2r_{jk}\left[P_{jk}^{\text{lin}}(\overline{p})\right]^+}{\underline{v}_k}\right) \\ a_i^2 &:= \sum_{(j,k)\in\mathbb{P}_i} \frac{2r_{jk}\left[Q_{jk}^{\text{lin}}(\overline{q})\right]^+}{\underline{v}_k} \\ a_i^3 &:= \sum_{(j,k)\in\mathbb{P}_i} \frac{2x_{jk}\left[P_{jk}^{\text{lin}}(\overline{p})\right]^+}{\underline{v}_k} \\ a_i^4 &:= \prod_{(j,k)\in\mathbb{P}_i} \left(1 - \frac{2x_{jk}\left[Q_{jk}^{\text{lin}}(\overline{q})\right]^+}{\underline{v}_k}\right) \end{aligned}$$

Consider the condition:

A4: $a_j^1 r_{jk} > a_j^2 x_{jk}$, $a_j^3 r_{jk} < a_j^4 x_{jk}$ for all $j \to k \in \tilde{E}$.

The following result is proved in [54].

Theorem 14 ([54]): Suppose A3–A4 hold. If $\overline{v}_j = \infty$, $j \in N \setminus \{1\}$, then OPF-socp is exact.

We make three remarks. First a variety of different sufficient conditions for the exactness of OPF-socp for radial networks have been proved in [61]. The condition A4 in Theorem 14 includes them as special cases. The set of sufficient conditions in [61] has the following simple interpretation: OPF-socp is exact provided either there are no reverse power flows in the network, or if the r/x ratios on all lines are equal, or if the r/x ratios increase in the downstream direction from the substation to the leaves then there are no reverse real power flows, or if the r/x ratios decrease in the downstream direction then there are no reverse reactive power flows. Second the condition A4 seems very mild and is satisfied by several test systems by a large margin [61], [54]. Finally the condition $\overline{v}_j = \infty$ in Theorem 14 can be replaced by a finite upper limit $\overline{v}_j$ but imposing $\overline{v}_j$ also on a linear function of the power injections (p,q) that represents an approximation to the voltage magnitudes. This suggests solving a modified OPF problem with this additional constraint on (p,q) [54]. It seems from several test systems that this modification is small, but guarantees that all the constraints on injections (p,q) as well as on voltage magnitudes v are met and that OPF-socp is exact with respect to the modified OPF problem.

The sufficient condition for the exactness of OPF-socp, proved in [62], [63], replaces A4 with a set of conditions on the r/x ratios in the network. Define the cumulative resistance and reactance from node i to node k as:

$$\begin{aligned} R_{ik} &:= \sum_{(j_1,j_2)\in\mathbb{P}_{ik}} r_{j_1 j_2} \\ X_{ik} &:= \sum_{(j_1,j_2)\in\mathbb{P}_{ik}} x_{j_1 j_2} \end{aligned}$$

Consider the condition:

A5: For each edge $(k,l) \in \tilde{E}$, if $\frac{r_{kl}}{x_{kl}} \geq \frac{R_{0k}}{X_{0k}}$ then for all edges $(i,j) \in \mathbb{P}_k$

$$\underline{v}_i \;>\; 2P_{ij}^{\text{lin}}(\overline{p})\left(\frac{r_{kl}}{x_{kl}} X_{0k} - R_{ik}\right) + 2Q_{ij}^{\text{lin}}(\overline{q}) X_{0i}$$

otherwise for all edges $(i,j) \in \mathbb{P}_k$

$$\underline{v}_i \;>\; 2P_{ij}^{\text{lin}}(\overline{p}) R_{0i} + 2Q_{ij}^{\text{lin}}(\overline{q})\left(\frac{x_{kl}}{r_{kl}} R_{0k} - X_{ik}\right)$$

Theorem 15: [[62], [63]] Suppose A3, A5 hold. If $\overline{v}_j = \infty$, $j \in N \setminus \{1\}$, then OPF-socp is exact.

While Theorem 14 is proved by considering the (primal) problem OPF-socp, Theorem 15 is proved by studying its Lagrangian dual.

Exact relaxation: DC mesh

In this subsection we consider purely resistive network, i.e., the impedance/admittance $z_{jk} = 1/y_{jk} = r_{jk}$, the power injections $s_j = p_j$, and the voltages V_j are all real. The constraint (35) should be replaced by

$$0 \ < \ \sqrt{\underline{v}_j} \ \leq \ V_j \ \leq \sqrt{\overline{v}_j}$$

In particular $V_j > 0$ for all $j \in N$.

We can therefore view DC voltages as complex quantities with zero phase angles $V_j = \sqrt{v_j} e^{\mathbf{i}0}$. We have from (32a)–(32b)

$$V_j^2 - r_{jk} P_{jk} \ = \ V_j V_k \ > \ 0$$

This implies that $\beta_{jk}(x)$ defined in (40) are zero for any $x \in \mathbb{X}_2^{nc}$. Hence (41) is always satisfied with $\theta(x) = 0$ and OPF-socp is exact if every optimal solution x^* in $\mathbb{X}_2^+$ attains equality in (43). Theorem 1 of [60] then implies the following result.

Theorem 16: Suppose A3 holds. If $\overline{p}_j^c = \overline{q}_j^c = \infty$, $j \in N$, then OPF-socp is exact.

Exact relaxation: AC mesh

For general AC networks a solution $x \in \mathbb{X}_2^{nc}$ of OPF-nc may not satisfy (41) and therefore cannot be mapped to a solution of OPF. It is proved in [60] however that if there are phase shifters in the network then any solution $x \in \mathbb{X}_2^{nc}$ becomes implementable, as we now explain. The use of phase shifters to convexify AC mesh networks has also been observed in [47] for the bus injection model.

A phase shifter can be a traditional transformer or a FACTS (Flexible AC Transmission Systems) device. We consider an idealized phase shifter that only shifts the phase angles of the sending-end voltage and current across a line, and has no impedance nor limits on the shifted angles. As before let V_i denote the sending-end voltage. Define I_{ij} to be the *sending-end* current leaving node i towards node j. Let k be the point between the phase shifter ϕ_{ij} and line impedance z_{ij}. Let V_k and I_k be the voltage at k and the current from k to j respectively. Then the effect of an idealized phase shifter, parametrized by ϕ_{ji}, is summarized by the following modeling assumption:

$$V_k \ = \ V_i \ e^{\mathbf{i}\phi_{ij}} \quad \text{and} \quad I_k \ = \ I_{ij} \ e^{\mathbf{i}\phi_{ij}}$$

The power transferred from nodes i to j is still (defined to be) $S_{ij} := V_i I_{ij}^*$ which, as expected, is equal to the power $V_k I_k^*$ from nodes k to j since the phase shifter is assumed to be lossless. Applying Ohm's law across z_{ij}, we define the *branch flow model with phase shifters* as the following set of equations:

$$I_{ij} \ = \ y_{ij} \left(V_i - V_j \ e^{-\mathbf{i}\phi_{ij}} \right) \tag{47a}$$
$$S_{ij} \ = \ V_i I_{ij}^* \tag{47b}$$
$$s_j \ = \sum_{k: j \to k} S_{jk} - \sum_{i: i \to j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) \tag{47c}$$

Without phase shifters ($\phi_{ij} = 0$), (47) reduces to the branch flow model (32).

The inclusion of phase shifters modifies the network and enlargers the solution set of the (new) branch flow equations. Formally, let

$$\overline{\mathbb{X}} \ := \ \left\{ \tilde{x} \in \mathbb{C}^{n+2m} \mid \tilde{x} \text{ solves (47a) for some } \phi, \right. \\ \left. \text{(47b), (34), (35)} \right\}$$

Unless otherwise specified, all angles should be interpreted as being modulo 2π and in $(-\pi, \pi]$. Hence we are primarily interested in $\phi \in (-\pi, \pi]^m$. For any spanning tree T of G, let "$\phi \in T^\perp$" be the shorthand for "$\phi_{ij} = 0$ for all $(i, j) \in T$", i.e., ϕ involves only phase shifters in branches not in the spanning tree T. Define

$$\overline{\mathbb{X}}_T := \left\{ \tilde{x} \in \mathbb{C}^{n+2m} | \tilde{x} \text{ solves (47a) for some } \phi \in T^\perp, \right. \\ \left. \text{(47b), (34), (35)} \right\}$$

Since (47) reduces to the branch flow model when $\phi = 0$, $\mathbb{X} \subseteq \overline{\mathbb{X}}_T \subseteq \overline{\mathbb{X}}$.

Recall the problem OPF (37) and OPF-nc. Define optimization problem where there is a phase shifter on every line in the network:

OPF-ps:

$$\min_{\tilde{x}, \phi} \ C(x) \quad \text{subject to} \quad \tilde{x} \in \overline{\mathbb{X}}$$

and the problem where, given any spanning tree T, there are phase shifters only outside T:

OPF-ps(T):

$$\min_{\tilde{x}, \phi} \ C(x) \quad \text{subject to} \quad \tilde{x} \in \overline{\mathbb{X}}_T, \ \phi \in T^\perp$$

Let the optimal values of OPF, OPF-nc, OPF-ps, and OPF-ps(T) be C^*, C_{nc}, C_{ps}, and C_T respectively. The following result is from [60]. It implies that if an optimal solution x^* of OPF-socp attains equality in (43) then x^* can be implemented by an appropriate choice of phase shifter angles ϕ. Such an x^* solves the problem OPF-nc. Moreover this benefit can be attained with phase shifters only outside an arbitrary spanning tree T of G.

Theorem 17 ([60]): Given any spanning tree T of G:

1) $\mathbb{X} = \overline{\mathbb{X}} = \overline{\mathbb{X}}_T$.
2) $C^* \geq C_{nc} = C_{ps} = C_T$.

Equivalence of bus injection and branch flow models

We now establish the equivalence between the bus injection model and the branch flow model and their relaxations.

Theorem 18 ([44], [45]): $\mathbb{W}_2 \equiv \mathbb{X}_2$ and $\mathbb{W}_2^+ \equiv \mathbb{X}_2^+$.

Corollary 4 establishes a bijection between $\mathbb{W}_2$ and the feasible set $\mathbb{V}$ of OPF in the bus injection model. Theorem 12 implies a bijection between $\mathbb{X}_2$ and the feasible set $\mathbb{X}$ of OPF in the branch flow model. Theorem 18 hence implies that there is a bijection between the feasible sets $\mathbb{V}$ and $\mathbb{X}$ of OPF in the bus injection model and the branch flow model respectively. It is in this sense that these two models are equivalent.

The bijection between these two models allows many results to be formulated and proved in either model.

References

[1] J. Carpentier. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 3(8):431–447, 1962. In French.

[2] H.W. Dommel and W.F. Tinney. Optimal power flow solutions. *Power Apparatus and Systems, IEEE Transactions on*, PAS-87(10):1866–1876, Oct. 1968.

[3] J. A. Momoh. *Electric Power System Applications of Optimization*. Power Engineering. Markel Dekker Inc.: New York, USA, 2001.

[4] M. Huneault and F. D. Galiana. A survey of the optimal power flow literature. *IEEE Trans. on Power Systems*, 6(2):762–770, 1991.

[5] J. A. Momoh, M. E. El-Hawary, and R. Adapa. A review of selected optimal power flow literature to 1993. Part I: Nonlinear and quadratic programming approaches. *IEEE Trans. on Power Systems*, 14(1):96–104, 1999.

[6] J. A. Momoh, M. E. El-Hawary, and R. Adapa. A review of selected optimal power flow literature to 1993. Part II: Newton, linear programming and interior point methods. *IEEE Trans. on Power Systems*, 14(1):105 – 111, 1999.

[7] K. S. Pandya and S. K. Joshi. A survey of optimal power flow methods. *J. of Theoretical and Applied Information Technology*, 4(5):450–458, 2008.

[8] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey, I: formulations and deterministic methods. *Energy Systems*, 3:221–258, September 2012.

[9] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey, II: nondeterministic and hybrid methods. *Energy Systems*, 3:259–289, September 2013.

[10] Mary B. Cain, Richard P. O'Neill, and Anya Castillo. History of optimal power flow and formulations (OPF Paper 1). Technical report, US FERC, December 2012.

[11] Richard P. O'Neill, Anya Castillo, and Mary B. Cain. The IV formulation and linear approximations of the AC optimal power flow problem (OPF Paper 2). Technical report, US FERC, December 2012.

[12] Richard P. O'Neill, Anya Castillo, and Mary B. Cain. The computational testing of AC optimal power flow using the current voltage formulations (OPF Paper 3). Technical report, US FERC, December 2012.

[13] Anya Castillo and Richard P. O'Neill. Survey of approaches to solving the ACOPF (OPF Paper 4). Technical report, US FERC, March 2013.

[14] Anya Castillo and Richard P. O'Neill. Computational performance of solution techniques applied to the ACOPF (OPF Paper 5). Technical report, US FERC, March 2013.

[15] B Stott and O. Alsaç. Fast decoupled load flow. *IEEE Trans. on Power Apparatus and Systems*, PAS-93(3):859–869, 1974.

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

[17] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans. Usefulness of DC power flow for active power flow analysis. In *Proc. of IEEE PES General Meeting*, pages 2457–2462. IEEE, 2005.

[18] B. Stott, J. Jardim, and O. Alsaç. DC Power Flow Revisited. *IEEE Trans. on Power Systems*, 24(3):1290–1300, Aug 2009.

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

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

[21] M. E. Baran and F. F Wu. Optimal Capacitor Placement on radial distribution systems. *IEEE Trans. Power Delivery*, 4(1):725–734, 1989.

[22] M. E Baran and F. F Wu. Optimal Sizing of Capacitors Placed on A Radial Distribution System. *IEEE Trans. Power Delivery*, 4(1):735–743, 1989.

[23] H-D. Chiang and M. E. Baran. On the existence and uniqueness of load flow solution for radial distribution power networks. *IEEE Trans. Circuits and Systems*, 37(3):410–416, March 1990.

[24] Hsiao-Dong Chiang. A decoupled load flow method for distribution power networks: algorithms, analysis and convergence study. *International Journal Electrical Power Energy Systems*, 13(3):130–138, June 1991.

[25] R. Cespedes. New method for the analysis of distribution networks. *IEEE Trans. Power Del.*, 5(1):391–396, January 1990.

[26] A. G. Expósito and E. R. Ramos. Reliable load flow technique for radial distribution networks. *IEEE Trans. Power Syst.*, 14(13):1063–1069, August 1999.

[27] R.A. Jabr. Radial Distribution Load Flow Using Conic Programming. *IEEE Trans. on Power Systems*, 21(3):1458–1459, Aug 2006.

[28] R. A. Jabr. A Conic Quadratic Format for the Load Flow Equations of Meshed Networks. *IEEE Trans. on Power Systems*, 22(4):2285–2286, Nov 2007.

[29] R. A. Jabr. Exploiting sparsity in sdp relaxations of the opf problem. *Power Systems, IEEE Transactions on*, 27(2):1138–1139, 2012.

[30] Joshua Adam Taylor. *Conic Optimization of Electric Power Systems*. PhD thesis, MIT, June 2011.

[31] Joshua A. Taylor and Franz S. Hover. Convex models of distribution system reconfiguration. *IEEE Trans. Power Systems*, 2012.

[32] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[33] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of semidefinite programming: theory, algorithms, and applications*, volume 27. Springer Netherlands, 2000.

[34] Miguel Soma Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.

[35] Mituhiro Fukuda, Masakazu Kojima, Kazuo Murota, and Kazuhide Nakata. Exploiting sparsity in semidefinite programming via matrix completion i: General framework. *SIAM Journal on Optimization*, 11:647–674, 1999.

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

[37] S. Zhang. Quadratic maximization and semidefinite relaxation. *Mathematical Programming*, 87(3):453–465, 2000.

[38] S. Kim and M. Kojima. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26(2):143–154, 2003.

[39] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz. Positive definite completions of partial Hermitian matrices. *Linear Algebra and its Applications*, 58:109–124, 1984.

[40] D. R. Fulkerson and O. A. Gross. Incidence matrices and interval graphs. *Pacific Journal of Mathematics*, 15(3):835–855, 1965.

[41] Donald J. Rose, Robert Endre Tarjan, and George S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM Journal on Computing*, 5(2):266–283, 1976.

[42] Deqiang Gan, Robert J. Thomas, and Ray D. Zimmerman. Stability-constrained optimal power flow. *IEEE Trans. Power Systems*, 15(2):535–540, May 2000.

[43] S. Bose, D. Gayme, S. H. Low, and K. M. Chandy. Quadratically constrained quadratic programs on acyclic graphs with application to power flow. arXiv:1203.5599v1, March 2012.

[44] Subhonmesh Bose, Steven H. Low, and Mani Chandy. Equivalence of branch flow and bus injection models. In *50th Annual Allerton Conference on Communication, Control, and Computing*, October 2012.

[45] Subhonmesh Bose, Steven H. Low, Mani Chandy, T. Teeraratkul, and Babak Hassibi. Equivalent relaxations of optimal power flow. Submitted for publication, 2013.

[46] Baosen Zhang and David Tse. Geometry of the injection region of power networks. *To appear in IEEE Trans. Power Systems*, 28(2):788–797, 2013.

[47] S. Sojoudi and J. Lavaei. Physics of power networks makes hard optimization problems easy to solve. In *IEEE Power & Energy Society (PES) General Meeting*, July 2012.

[48] Masoud Farivar, Christopher R. Clarke, Steven H. Low, and K. Mani Chandy. Inverter var control for distribution systems with renewables. In *Proceedings of IEEE SmartGridComm Conference*, October 2011.

[49] J. Lavaei and S. H. Low. Zero duality gap in optimal power flow problem. *IEEE Trans. on Power Systems*, 27(1):92–107, February 2012.

[50] S. Bose, D. Gayme, S. H. Low, and K. M. Chandy. Optimal power flow over tree networks. In *Proc. Allerton Conf. on Comm., Ctrl. and Computing*, October 2011.

[51] B. Zhang and D. Tse. Geometry of feasible injection region of power networks. In *Proc. Allerton Conf. on Comm., Ctrl. and Computing*, October 2011.

[52] Lingwen Gan and Steven H. Low. Optimal power flow in DC networks. Submitted for publication, March 2013.

[53] Javad Lavaei, David Tse, and Baosen Zhang. Geometry of power flows and optimization in distribution networks. *arXiv*, November 2012.

[54] Lingwen Gan, Na Li, Ufuk Topcu, and Steven H. Low. Optimal power flow in distribution networks. In *Submitted to 52nd IEEE Conference on Decision and Control*, December 2013. in arXiv:12084076.

[55] Y. Colin de Verdi'ere. Multiplicities of eigenvalues and tree-width graphs. *Journal of Combinatorial Theory*, 74:121146, 1998. Series B.

[56] Hein van der Holst. Graphs whose positive semidefinite matrices have nullity at most two. *Linear Algebra and its Applications*, 375:1–11, 2003.

[57] C.R. Johnson, A. Leal Duarte, C.M. Saiago, B.D. Sutton, and A.J. Witt. On the relative position of multiple eigenvalues in the spectrum of an Hermitian matrix with a given graph. *Linear Algebra and its Applications*, 363:147–159, 2003.

[58] Albert Y.S. Lam, Baosen Zhang, Alejandro Domínguez-García, and David Tse. Optimal distributed voltage regulation in power distribution networks. *arXiv*, April 2012.

[59] Javad Lavaei, Anders Rantzer, and Steven H. Low. Power flow optimization using positive quadratic programming. In *Proceedings of IFAC World Congress*, 2011.

[60] Masoud Farivar and Steven H. Low. Branch flow model: relaxations and convexification (parts I, II). *IEEE Trans. on Power Systems*, 2013.

[61] Lingwen Gan, Na Li, Ufuk Topcu, and Steven H. Low. On the exactness of convex relaxation for optimal power flow in tree networks. In *51st IEEE Conference on Decision and Control*, December 2012.

[62] Na Li, Lijun Chen, and Steven Low. Exact convex relaxation of opf for radial networks using branch flow models. In *IEEE International Conference on Smart Grid Communications*, November 2012.

[63] Na Li. *Distributed Optimization in Power Networks and General Multi-agent Systems*. PhD thesis, Caltech, May 2013.