Branch Flow Model: Relaxations and Convexification—Part I

Masoud Farivar and Steven H. Low, Fellow, IEEE

Abstract—We propose a branch flow model for the analysis and optimization of mesh as well as radial networks. The model leads to a new approach to solving optimal power flow (OPF) that consists of two relaxation steps. The first step eliminates the voltage and current angles and the second step approximates the resulting problem by a conic program that can be solved efficiently. For radial networks, we prove that both relaxation steps are always exact, provided there are no upper bounds on loads. For mesh networks, the conic relaxation is always exact but the angle relaxation may not be exact, and we provide a simple way to determine if a relaxed solution is globally optimal. We propose convexification of mesh networks using phase shifters so that OPF for the convexified network can always be solved efficiently for an optimal solution. We prove that convexification requires phase shifters only outside a spanning tree of the network and their placement depends only on network topology, not on power flows, generation, loads, or operating constraints. Part I introduces our branch flow model, explains the two relaxation steps, and proves the conditions for exact relaxation. Part II describes convexification of mesh networks, and presents simulation results.

Index Terms—Convex relaxation, load flow control, optimal power flow, phase control, power system management.

I. INTRODUCTION

A. Motivation

The bus injection model is the standard model for power flow analysis and optimization. It focuses on nodal variables such as voltages, current and power injections and does not directly deal with power flows on individual branches. Instead of nodal variables, the branch flow model 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. In this paper, we advocate the use of branch flow model for both radial and mesh networks, and demonstrate how it can be used for optimizing the design and operation of power systems.

One of the motivations for our work is the optimal power flow (OPF) problem. 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 contingency constraints on the voltages and power flows. There has been a great deal of research on OPF since Carpenter’s first formulation in 1962 [2]; surveys can be found in, e.g., [3]–[7]. 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 power flow problem, which is a linearization and therefore easy to solve, e.g., [8]–[11]. An important observation was made in [12] and [13] that the full 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 [12] and [13] 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. Instead of solving the OPF problem directly, [14] proposes to solve its convex Lagrangian dual problem and gives a sufficient condition that must be satisfied by a dual solution for an optimal OPF solution to be recoverable. This result is extended in [15] to include other variables and constraints and in [16] to exploit network sparsity. In [17] and [18], it is proved that the sufficient condition of [14] always holds for a radial (tree) network, provided the bounds on the power flows satisfy a simple pattern. See also [19] for a generalization. These results confirm that radial networks are computationally much simpler. This is important as most distribution systems are radial.

The limitation of semidefinite relaxation for OPF is studied in [20] using mesh networks with 3, 5, and 7 buses: as a line-flow constraint is tightened, the duality gap becomes nonzero and the solutions produced by the semidefinite relaxation becomes physically meaningless. Indeed, examples of nonconvexity have long been discussed in the literature, e.g., [21]–[23]. See, e.g., [24] for branch-and-bound algorithms for solving OPF when convex relaxation fails.

The papers above are all based on the bus injection model. In this paper, we introduce a branch flow model on which OPF and its relaxations can also be defined. Our model is motivated by a model first proposed by Baran and Wu in [25] and [26] for the optimal placement and sizing of switched capacitors in distribution circuits for Volt/VAR control. One of the insights we highlight here is that the Baran-Wu model of [25] and [26] can be treated as a particular relaxation of our branch flow model where the phase angles of the voltages and currents are ignored. By recasting their model as a set of linear and quadratic equality constraints, [27] and [28] observe that relaxing the quadratic equality constraints to inequality constraints yields a second-order cone program (SOCP). It proves that the SOCP relaxation is exact for radial networks, when there are no upper bounds on the loads. This result is extended here to mesh networks with...
line limits, and convex, as opposed to linear, objective functions (Theorem 1). See also [29] and [30] for various convex relaxations of approximations of the Baran-Wu model for radial networks.

Other branch flow models have also been studied, e.g., in [31]–[33], all for radial networks. Indeed [31] studies a similar model to that in [25] and [26], using receiving-end branch powers as variables instead of sending-end branch powers as in [25] and [26]. Both [32] and [33] eliminate voltage angles by defining real and imaginary parts of $V_i V_j^*$ as new variables and defining bus power injections in terms of these new variables. This results in a system of linear quadratic equations in power injections and the new variables. While [32] develops a Newton-Raphson algorithm to solve the bus power injections, [33] solves for the branch flows through an SOCP relaxation for radial networks, though no proof of optimality is provided.

This set of papers [25]–[33] all exploit the fact that power flows can be specified by a simple set of linear and quadratic equalities if voltage angles can be eliminated. Phase angles can be relaxed only for radial networks and generally not for mesh networks, as [34] points out for their branch flow model, because cycles in a mesh network impose nonconvex constraints on the optimization variables (similar to the angle recovery condition in our model; see Theorem 2 below). For mesh networks, [34] proposes a sequence of SOCP where the nonconvex constraints are replaced by their linear approximations and demonstrates the effectiveness of this approach using seven network examples. In this paper we extend the Baran-Wu model from radial to mesh networks and use it to develop a solution strategy for OPF.

B. Summary

Our purpose is to develop a formal theory of branch flow model for the analysis and optimization of mesh as well as radial networks. As an illustration, we formulate OPF within this alternative model, propose relaxations, characterize when a relaxed solution is exact, prove that our relaxations are always exact for radial networks when there are no upper bounds on loads but may not be exact for mesh networks, and show how to use phase shifters to convexify a mesh network so that a relaxed solution is always optimal for the convexified network.

Specifically we formulate in Section II the OPF problem using branch flow equations involving complex bus voltages and complex branch current and power flows. In Section III we describe our solution approach that consists of two relaxation steps (see Fig. 1):

- **Angle relaxation**: relax OPF by eliminating voltage and current angles from the branch flow equations. This yields the (extended) Baran-Wu model and a relaxed problem OPF-ar which is still nonconvex due to a quadratic equality constraint.

- **Conic relaxation**: relax OPF-ar by changing the quadratic equality into an inequality constraint. This yields a convex problem OPF-cr (which is an SOCP when the objective function is linear).

In Section IV we prove that the conic relaxation OPF-cr is always exact even for mesh networks, provided there are no upper bounds on real and reactive loads, i.e., any optimal solution of OPF-cr is also optimal for OPF-ar. Given an optimal solution of OPF-ar, whether we can derive an optimal solution of the original OPF depends on whether we can recover the voltage and current angles from the given OPF-ar solution. In Section V we characterize the exact condition (the angle recovery condition) under which this is possible, and present two angle recovery algorithms. The angle recovery condition has a simple interpretation: any solution of OPF-ar implies an angle difference across a line, and the condition says that the implied angle differences sum to zero $\pmod{2\pi}$ around each cycle. For a radial network, this condition holds trivially and hence solving the conic relaxation OPF-cr always produces an optimal solution for OPF. For a mesh network, the angle recovery condition corresponds to the requirement that the implied phase angle differences sum to zero around every loop. The given OPF-ar solution may not satisfy this condition, but our characterization can be used to check if it yields an optimal solution for OPF. These results suggest an algorithm for solving OPF as summarized in Fig. 2.

If a relaxed solution for a mesh network does not satisfy the angle recovery condition, then it is infeasible for OPF. In Part II of this paper, we propose a simple way to convexify a mesh network using phase shifters so that any relaxed solution of OPF-ar can be mapped to an optimal solution of OPF for the convexified network, with an optimal cost that is lower than or equal to that of the original network.
C. Extensions: Radial Networks and Equivalence

In [35] and [36], we prove a variety of sufficient conditions under which the conic relaxation proposed here is exact for radial networks. The main difference from Theorem 1 below is that [35] and [36] allow upper bounds on the loads but relax upper bounds on voltage magnitudes. Unlike the proof for Theorem 1 here, those in [35] and [36] exploit the duality theory.

The bus injection model and the branch flow model are defined by different sets of equations in terms of their own variables. Each model is self-contained: one can formulate and analyze power flow problems within each model, using only nodal variables or only branch variables. Both models (i.e., the sets of equations in their respective variables), however, are descriptions of the Kirchhoff’s laws. In [37] we prove formally the equivalence of these models, in the sense that given a power flow solution in one model, one can derive a corresponding power flow solution in the other model. Although the semidefinite relaxation in the bus injection model is very different from the convex relaxation proposed here, [37] also establishes the precise relationship between the various relaxations in these two models.

This is useful because some results are easier to formulate and prove in one model than in the other. For instance, it is hard to see how the upper bounds on voltage magnitudes and the technical conditions on the line impedances in [35] and [36] for exactness in the branch flow model affect the rank of the semidefinite matrix variable in the bus injection model, although [37] clarifies conditions that guarantee their equivalence.

II. BRANCH FLOW MODEL

Let \( \mathbb{R} \) denote the set of real numbers, \( \mathbb{C} \) complex numbers, and \( \mathbb{N} \) integers. A variable without a subscript denotes a vector with appropriate components, e.g., \( s := (s_i, i = 1, \ldots, n) \). For a vector \( a = (a_1, a_2, \ldots, a_k) \), \( a = (a_1, a_2, \ldots, a_k) \). For a scalar, vector, or matrix \( A \), \( A^T \) denotes its transpose and \( A^* \) its complex conjugate transpose.

Given a directed graph \( G = (N, E) \), denote a link in \( E \) by \((i, j)\) or \( i \rightarrow j \) if it points from node \( i \) to node \( j \). We will use \( i, j \), \((i, j)\), or \( i \rightarrow j \) interchangeably to refer to a link in \( E \). We write \( i \sim j \) if \( i \) and \( j \) are connected, i.e., if either \((i, j) \in E \) or \((j, i) \in E \) (but not both). We write \( \theta = 0 \) (\( \text{mod } 2\pi \)) if \( \theta = 2\pi k \), and \( \phi = \text{mod } 2\pi \) \( \phi = \text{mod } 2\pi \) \( \phi = \text{mod } 2\pi \), for some integer \( k \). For a \( d \)-dimensional vector \( \alpha \), \( P^T(\alpha) \) denotes its projection onto \((-\pi, \pi)^d \) by taking modulo \( 2\pi \) componentwise.

A. Branch Flow Model

Let \( G = (N, K) \) be a connected graph representing a power network, where each node in \( N \) represents a bus and each link in \( E \) represents a line (condition A1). We index the nodes by \( i = 0, 1, \ldots, n \). The power network is called radial if its graph \( G \) is a tree. For a distribution network, which is typically radial, the root of the tree (node 0) represents the substation bus. For a (generally meshed) transmission network, node 0 represents the slack bus.

We regard \( G \) as a directed graph and adopt the following orientation for convenience (only). Pick any spanning tree \( T := (N, E_T) \) of \( G \) rooted at node 0, i.e., \( T \) is connected and \( E_T \subseteq E \) has \( n \) links. All links in \( E_T \) point away from the root. For any link in \( E \setminus E_T \) that is not in the spanning tree \( T \), pick an arbitrary direction. Denote a link by \((i, j)\) or \( i \rightarrow j \) if it points from node \( i \) to node \( j \). Henceforth we will assume without loss of generality that \( G \) and \( T \) are directed graphs as described above.

For each \((i, j) \in E \), let \( x_{ij} = r_{ij} + jx_{ij} \) be the complex impedance on the line, and \( y_{ij} := 1/2z_{ij} := g_{ij} - jb_{ij} \) be the corresponding admittance. For each node \( i \in N \), let \( z_i = r_i + jx_i \) be the shunt impedance from \( i \) to ground, and \( y_i := 1/z_i := g_i - jb_i \).

For each \((i, j) \in E \), let \( I_{ij} \) be the complex current from buses \( i \) to \( j \) and \( S_{ij} = P_{ij} + jQ_{ij} \) be the sending-end complex power from buses \( i \) to \( j \). For each node \( i \in N \), let \( V_i \) be the complex voltage on bus \( i \). Let \( s_i \) be the net complex power injection, which is generation minus load on bus \( i \). We use \( s_i \) to denote both the complex number \( p_i + jq_i \) and the pair \((p_i, q_i)\) depending on the context.

As customary, we assume that the complex voltage \( V_0 \) is given and the complex net generation \( s_0 \) is a variable. For power flow analysis, we assume other power injections \( s := (s_i, i = 1, \ldots, n) \) are given. For optimal power flow, VAR control, or demand response, \( s \) are control variables as well.

Given \( z := (z_{ij}, (i, j) \in E, z_i, i \in N) \), \( V_0 \) and bus power injections \( s \), the variables \((S, I, V, s_0) := (S_{ij}, I_{ij}, (i, j) \in E, V_i, i = 1, \ldots, n, s_0) \) satisfy the Ohm’s law:

\[
V_i - V_j = z_{ij}I_{ij}, \quad \forall (i, j) \in E \quad (1)
\]

the definition of branch power flow:

\[
S_{ij} = V_iI_{ij}^*, \quad \forall (i, j) \in E \quad (2)
\]

and power balance at each bus: for all \( j \in N \),

\[
\sum_{k \rightarrow j} S_{kj} - \sum_{i \rightarrow j} (S_{ij} - z_{ij}|I_{ij}|^2) + y_j^2|V_j|^2 = s_j. \quad (3)
\]

We will refer to (1)-(3) as the branch flow model/equations. Recall that the cardinality \(|N| = n + 1\) and let \( E := m \). The branch flow equations (1)-(3) specify \( 2m + n + 1 \) nonlinear equations in \( 2m + n + 1 \) complex variables \((S, I, V, s_0)\), when other bus power injections \( s \) are specified.

We will call a solution of (1)-(3) a branch flow solution with respect to a given \( s \), and denote it by \( x[s] := (S, I, V, s_0) \). Let \( \mathbf{X}(s) \subseteq \mathbb{C}^{2m+n+1} \) be the set of all branch flow solutions with respect to a given \( s \):

\[
\mathbf{X}[s] := \{ x := (S, I, V, s_0) \, \mid \, x \text{ solves (1)-(3) given } s \} \quad (4)
\]

and let \( \mathbf{X} \) be the set of all branch flow solutions:

\[
\mathbf{X} := \bigcup_{s \in \mathbb{C}^n} \mathbf{X}(s). \quad (5)
\]

For simplicity of exposition, we will often abuse notation and use \( \mathbf{X} \) to denote either the set defined in (4) or that in (5), de-
pending on the context. For instance, $\mathbf{X}$ is used to denote the set in (4) for a fixed $s$ in Section V for power flow analysis, and to denote the set in (5) in Section IV for optimal power flow where $s$ itself is also an optimization variable. Similarly for other variables such as $x$ for $x(s)$.

### B. Optimal Power Flow

Consider the optimal power flow problem where, in addition to $(S, I, V_n, s_0)$, $s$ is also an optimization variable. Let $p_i := p_i^2 - p_i^4$ and $q_i := q_i^2 - q_i^4$ where $p_i^2$ and $q_i^2$ ($p_i^4$ and $q_i^4$) are the real and reactive power generation (consumption) at node $i$. For instance, [25] and [26] formulate a Volt/VAR control problem for a distribution circuit where $q_i^2$ represent the placement and sizing of shunt capacitors. In addition to (1)–(3), we impose the following constraints on power generation: for $i \in N$:

$$ p_i^2 \leq \bar{p}_i^2, \quad q_i^2 \leq \bar{q}_i^2. \tag{6} $$

In particular, any of $\bar{p}_i^2, \bar{q}_i^2$ can be a fixed constant by specifying that $p_i^2 = \bar{p}_i^2$ and/or $q_i^2 = \bar{q}_i^2$. For instance, in the inverter-based VAR control problem of [27] and [28], $p_i^2$ are the fixed (solar) power outputs and the reactive power $q_i^2$ are the control variables. For power consumption, we require, for $i \in N$:

$$ p_i^2 \leq \bar{p}_i^2, \quad q_i^2 \leq \bar{q}_i^2. \tag{7} $$

The voltage magnitudes must be maintained in tight ranges: for $i = 1, \ldots, n$:

$$ \underline{v}_i \leq |V_i|^2 \leq \bar{v}_i. \tag{8} $$

Finally, we impose flow limits in terms of branch currents: for all $(i, j) \in E$:

$$ |I_{ij}| \leq \bar{I}_{ij}. \tag{9} $$

We allow any objective function that is convex and does not depend on the angles $\mathcal{L}_V, \mathcal{L}_I$ of voltages and currents. For instance, suppose we aim to minimize real power losses $r_{ij} I_{ij}^2$ [38], [39], minimize real power generation costs $c_p p_i^2$, and maximize energy savings through conservation voltage reduction (CVR). Then the objective function takes the form (see [27] and [28]):

$$ \sum_{(i,j) \in E} r_{ij} |I_{ij}|^2 + \sum_{i \in N} c_p p_i^2 + \sum_{i \in N} \alpha_i |V_i|^2 \tag{10} $$

for some given constants $r_{ij}, \alpha_i \geq 0$.

To simplify notation, let $\ell_{ij} := |I_{ij}|^2$ and $v_i := |V_i|^2$. Let $s^2 := (s_i^2, i = 1, \ldots, n) = (p_i^2, q_i^2, i = 1, \ldots, n)$ be the power generations, and $s^c := (s_i^c, i = 1, \ldots, n) = (p_i^c, q_i^c, i = 1, \ldots, n)$ the power consumptions. Let $s$ denote either $s^2 - s^c$ or $(s^2, s^c)$ depending on the context. Given a branch flow solution $x := x(s) := (S, I, V, s_0)$ with respect to a given $s$, let $\hat{y} := \hat{y}(x) := (S, \bar{I}, \bar{V}, s_0)$ denote the projection of $x$ that have phase angles $\mathcal{L}_V, \mathcal{L}_I$ eliminated. This defines a projection function $\hat{h}$ such that $\hat{y} = \hat{h}(x)$, to which we will return in Section III. Then our objective function is $f(\hat{h}(x), s)$. We assume $f(\hat{y}, s)$ is convex (condition A2); in addition, we assume $f$ is strictly increasing in $\ell_{ij}, (i, j) \in E$, nonincreasing in load $s^2$, and independent of $S$ (condition A3). Let

$$ S := \{(S, v, s_0, x) | (v, s_0, x) \text{ satisfies (6) – (9)} \}. $$

All quantities are optimization variables, except $V_0$ which is given.

The optimal power flow problem is

$$ \text{OPF:} \begin{align*} \min_{x, s} & \quad f(\hat{h}(x), s) \tag{11} \\
\text{subject to} & \quad x \in \mathbf{X}, \quad (S, v, s_0, s) \in S \tag{12} \end{align*} $$

where $\mathbf{X}$ is defined in (5).

The feasible set is specified by the nonlinear branch flow equations and hence OPF (11)–(12) is in general nonconvex and hard to solve. The goal of this paper is to propose an efficient way to solve OPF by exploiting the structure of the branch flow model.

### C. Notations and Assumptions

The main variables and assumptions are summarized in Table I and below for ease of reference:

- **A1)** The network graph $G$ is connected.
- **A2)** The cost function $f(\hat{y}, s)$ for optimal power flow is convex.
- **A3)** The cost function $f(\hat{y}, s)$ is strictly increasing in $\ell$, nonincreasing in load $s^2$, and independent of $S$.
- **A4)** The optimal power flow problem OPF (11)–(12) is feasible.

These assumptions are standard and realistic. For instance, the objective function in (10) satisfies conditions A2–A3. A3 is a property of the objective function and not a property of power flow solutions; it holds if the cost function is strictly increasing in line loss.

### III. RELAXATIONS AND SOLUTION STRATEGY

#### A. Relaxed Branch Flow Model

Substituting (2) into (1) yields $V_j = V_i - z_{ij} S_{ij}^*/V_j^*$. Taking the magnitude squared, we have $v_j = v_i + |z_{ij}|^2 \ell_{ij} - (z_{ij} S_{ij}^*/z_{ij}^2 S_{ij})$. Using (3) and (2) and in terms of real variables, we therefore have

$$ p_j = \sum_{k:j \leftarrow k} P_{jk} - \sum_{i:j \rightarrow i} (P_{ij} - r_{ij} \ell_{ij}) + g_j v_j, \quad \forall j \tag{13} $$

$$ q_j = \sum_{k:j \leftarrow k} Q_{jk} - \sum_{i:j \rightarrow i} (Q_{ij} - x_{ij} \ell_{ij}) + b_j v_j, \quad \forall j \tag{14} $$

$$ v_j = v_i - 2(r_{ij} P_{ij} + x_{ij} Q_{ij}) + (r_{ij}^2 + x_{ij}^2) \ell_{ij} \quad \forall (i, j) \in E \tag{15} $$

$$ \ell_{ij} := \frac{p_{ij}^2 + q_{ij}^2}{v_i}, \quad \forall (i, j) \in E. \tag{16} $$

We will refer to (13)–(16) as the relaxed (branch flow) model/equations and a solution a relaxed (branch flow) solution. These
TABLE I
NOTATIONS

| G, T | (directed) network graph G and a spanning tree T of G |
| B, B_T | reduced (and transposed) incidence matrix of G and the submatrix corresponding to T |
| V_i, v_i | complex voltage on bus i with v_i := |V_i|^2 |
| s_i = p_i + jq_i | net complex load power on bus i |
| p_i = p_i^r - p_i^i | net real power equals generation minus load; |
| q_i = q_i^r - q_i^i | net reactive power equals generation minus load |
| I_{ij}, \ell_{ij} | complex current from buses i to j with \ell_{ij} := |I_{ij}|^2 |
| S_{ij} = P_{ij} + jQ_{ij} | complex power from buses i to j (sending-end) |

X \subseteq \mathbb{C}^{2m+n+1} the set of all branch flow solutions that satisfy (1)–(3) either for some \(s\), or for a given \(s\) (sometimes denoted more accurately by \(X(s)\));

\(\hat{Y}\) set of all relaxed branch flow solutions that satisfy (13)–(16) either for a given \(s\) or for some \(s\);

\(\hat{Y}\) set of all relaxed branch flow solutions that satisfy (13)–(15) and (22) either for a given \(s\) or for some \(s\);

\(x = (S, I, V, s_0) \in X\) vector of power flow variables and its projection \(\hat{g}\);

\(\tilde{x}_{ij}, \tilde{y}_i|\) impedance on line \((i, j)\) and shunt admittance from bus i to ground

\(f = f(\hat{h}(x), s)\) objective function of OPF

\(\hat{h}(\mathbf{X})\) the set of relaxed solutions. The inverse projection \(\hat{h}\) is defined in Section V.

In contrast to the original branch flow equations (1)–(3), the relaxed (13)–(16) specifies \(2(m+n+1)\) equations in \(3m+n+2\) real variables \((P, Q, \ell, v, p_0, q_0)\), given \(s\). For a radial network, i.e., \(G\) is a tree, \(m = |E| = |N| - 1 = n\). Hence the relaxed system (13)–(16) specifies \(4n+2\) equations in \(4n+2\) real variables. It is shown in [40] that there are generally multiple solutions, but for practical networks where \(|V_0| \approx 1\) and \(\ell_{ij}, x_{ij}\) are small p.u., the solution of (13)–(16) is unique. Exploiting structural properties of the Jacobian matrix, efficient algorithms have also been proposed in [41] to solve the relaxed branch flow equations.

For a connected mesh network, \(m = |E| > |N| - 1 = n\), in which case there are more variables than equations for the relaxed model (13)–(16), and therefore the solution is generally nonunique. Moreover, some of these solutions may be spurious, i.e., they do not correspond to a solution of the original branch flow equations (1)–(3).

Indeed, one may consider \((S, \ell, v, s_0)\) as a projection of \((S, I, V, s_0)\) where each variable \(I_{ij}\) or \(V_i\) is relaxed from a point in the complex plane to a circle with a radius equal to the distance of the point from the origin. It is therefore not surprising that a relaxed solution of (13)–(16) may not correspond to any solution of (1)–(3). The key is whether, given a relaxed solution, we can recover the angles \(\angle V_i, \angle I_{ij}\) correctly from it. It is then remarkable that, when \(G\) is a tree, indeed the solutions of (13)–(16) coincide with those of (1)–(3). Moreover, for general networks, (13)–(16) together with the angle recovery condition in Theorem 2 below are indeed equivalent to (1)–(3), as explained in Remark 5 of Section V.

To understand the relationship between the branch flow model and the relaxed model and formulate our relaxations precisely, we need some notations. Fix an \(s\).

Given a vector \((S, I, V, s_0) \in \mathbb{C}^{2m+n+1}\), define its projection \(\hat{h} : \mathbb{C}^{2m+n+1} \rightarrow \mathbb{R}^{2m+n+2}\) by \(\hat{h}(S, I, V, s_0) = (P, Q, \ell, v, p_0, q_0)\) where

\[P_{ij} = \text{Re} \ S_{ij}, \quad Q_{ij} = \text{Im} \ S_{ij}, \quad \ell_{ij} = |I_{ij}|^2 \]

\[p_i = \text{Re} \ s_i, \quad q_i = \text{Im} \ s_i, \quad v_i = |V_i|^2.\]

Let \(\hat{Y} \subseteq \mathbb{C}^{2m+n+1}\) denote the set of all \(y := (S, I, V, s_0)\) whose projections are the relaxed solutions:

\[\hat{Y} := \{ y := (S, I, V, s_0) : \hat{h}(y) \text{ solves } (13) - (16) \}.\]

Define the projection \(\hat{Y} := \hat{h}(Y)\) onto the space \(\mathbb{R}^{2m+n+1}\) as

\[\hat{Y} := \{ \hat{y} := (S, \ell, v, s_0) : \hat{y} \text{ solves } (13) - (16) \}.\]

Clearly

\[X \subseteq Y \text{ and } \hat{h}(X) \subseteq \hat{h}(Y) =: \hat{Y}.\]

Their relationship is illustrated in Fig. 3.

\(\text{B. Two Relaxations}\)

Consider the OPF with angles relaxed:

\[
\min_{x, s} f(\hat{h}(x), s) \]

subject to \(x \in Y, \quad (S, v, s_0, s) \in S\).

Clearly, this problem provides a lower bound to the original OPF problem since \(Y \supseteq X\). Since neither \(\hat{h}(x)\) nor the constraints in

\[\text{as mentioned earlier, the set defined in (19) is strictly speaking } Y(s) \text{ with respect to a fixed } s. \text{ To simplify exposition, we abuse notation and use } Y \text{ to denote both } Y(s) \text{ and } \bigcup_{s \in S} Y(s), \text{ depending on the context. The same applies to } Y \text{ and } \hat{Y}, \text{ etc.}\]
\( \mathbf{Y} \) involves angles \( \angle V_i, \angle I_{ij} \), this problem is equivalent to the following

**OPF-ar:**

\[
\begin{align*}
\min_{\hat{y}, s} & \quad f(\hat{y}, s) \\
\text{subject to} & \quad \hat{y} \in \hat{Y}, \quad (S, v, s_0, s) \in \mathbf{S}.
\end{align*}
\]

(20) (21)

The feasible set of OPF-ar is still nonconvex due to the quadratic equalities in (16). Relax them to inequalities:

\[
\ell_{ij} \geq \frac{P^2_{ij} + Q^2_{ij}}{v_i}, \quad (i,j) \in E.
\]

(22)

Define the convex second-order cone (see Theorem 1 below) \( \mathbf{Y} \subseteq \mathbb{R}^{2m+n+1} \) that contains \( \hat{Y} \) as

\[
\mathbf{Y} := \{ \hat{y} : (S, \ell, v, s_0) | \hat{y} \text{ solves (13) } - (15) \text{ and (22)} \}.
\]

Consider the following conic relaxation of OPF-ar:

**OPF-cr:**

\[
\begin{align*}
\min_{\hat{y}, s} & \quad f(\hat{y}, s) \\
\text{subject to} & \quad \hat{y} \in \mathbf{Y}, \quad (S, v, s_0, s) \in \mathbf{S}.
\end{align*}
\]

(23) (24)

Clearly OPF-cr provides a lower bound to OPF-ar since \( \mathbf{Y} \supseteq \hat{Y} \).

**C. Solution Strategy**

In the rest of this paper, we will prove the following:

1) OPF-cr is convex. Moreover, if there are no upper bounds on loads, then the conic relaxation is exact so that any optimal solution \((\hat{y}_{cr}, s_{cr})\) of OPF-cr is also optimal for OPF-ar for mesh as well as radial networks (Section IV, Theorem 1). OPF-cr is an SOCP when the objective function is linear.

2) Given a solution \((\hat{y}_{ar}, s_{ar})\) of OPF-ar, if the network is radial, then we can always recover the phase angles \( \angle V_i, \angle I_{ij} \) uniquely to obtain an optimal solution \((x, s)\) of the original OPF through an inverse projection (Section V, Theorems 2 and 4).

3) For a mesh network, an inverse projection may not exist to map the given \((\hat{y}_{ar}, s_{ar})\) to a feasible solution of OPF-ar. Our characterization can be used to determine if \((\hat{y}_{ar}, s_{ar})\) is globally optimal.

These results motivate the algorithm in Fig. 2.

In Part II of this paper, we show that a mesh network can be convexified so that \((\hat{y}_{ar}, s_{ar})\) can always be mapped to an optimal solution of OPF for the convexified network. Moreover, convexification requires phase shifters only on lines outside an arbitrary spanning tree of the network graph.

**IV. EXACT CONIC RELAXATION**

Our first key result says that OPF-cr is exact and an SOCP when the objective function is linear.

**Theorem 1:** Suppose \( \ell_{kj} - \ell_{ij} = \infty, i \in N \). Then OPF-cr is convex. Moreover, it is exact, i.e., any optimal solution of OPF-cr is also optimal for OPF-ar.

**Proof:** The feasible set is convex since the nonlinear inequalities in \( \mathbf{Y} \) can be written as the following second order cone constraint:

\[
\begin{align*}
\begin{bmatrix} 2P_{ij} \\ 2Q_{ij} \\ \ell_{ij} - v_i \end{bmatrix}_2 & \leq (\ell_{ij} + v_i).
\end{align*}
\]

Since the objective function is convex, OPF-cr is a conic optimization.\(^4\) To prove that the relaxation is exact, it suffices to show that any optimal solution of OPF-cr attains equality in (22).

Assume for the sake of contradiction that \((\hat{y}_{ar}, s_{ar}) := (S, \ell, v, s_0, s_{cr}, \tilde{s}_{cr}, \tilde{s}_{cr})\) is optimal for OPF-cr, but a link \((i, j) \in E\) has strict inequality, i.e., \( v_i, \ell_{ij} > 2P_{ij}^2 + 2Q_{ij}^2 \). For some \( \varepsilon > 0 \) to be determined below, consider another point \((\hat{y}, \hat{s}) = (S, \hat{\ell}, \hat{v}, \hat{s}_0, \hat{s}_{cr}, \hat{s}_{cr})\) defined by

\[
\begin{align*}
\hat{v} - v_i, & \quad \tilde{s} - \tilde{s}_{cr} = -\varepsilon \\
\hat{\ell}_{ij} - [\ell_{ij}], & \quad \hat{s}_{ij} - [s_{ij}], \quad \tilde{s}_{ij} - [s_{ij}] \\
\hat{s}_{ij} - [s_{ij}], & \quad \tilde{s}_{ij} - [s_{ij}]
\end{align*}
\]

where a negative index means excluding the indexed element from a vector. Since \( \hat{\ell}_{ij} - [\ell_{ij}] = \hat{v} - v_i \) has a strictly smaller objective value than \((\hat{y}_{ar}, s_{ar})\) because of assumption A3. If \((\hat{y}, \hat{s})\) is a feasible point, then it contradicts the optimality of \((\hat{y}_{ar}, s_{ar})\).

It suffices then to check that there exists an \( \varepsilon > 0 \) such that \((\hat{y}, \hat{s})\) satisfies (6)–(9), (13)–(15) and (22), and hence is indeed a feasible point. Since \((\hat{y}_{ar}, s_{ar})\) is feasible, (6)–(9) hold for \((\hat{y}, \hat{s})\) too. Similarly, \((\hat{y}, \hat{s})\) satisfies (13)–(14) at all nodes \( k \neq i, j \) and (15), (22) over all links \( (k, l) \neq (i, j) \). We now show that \((\hat{y}, \hat{s})\) satisfies (13)–(14) also at nodes \( i, j \), and (15), (22) over \((i, j)\).

Proving (13)–(14) is equivalent to proving (3). At node \( i \), we have

\[
\begin{align*}
\hat{s}_i & = s_i^k - s_i^l + [s_{ij}^k] - [s_{ij}^l] - s_{ij}k^l / 2 \\
& = \sum_{i \neq j, j^l} (s_{ij}^k - s_{ij}^l) + [s_{ij}^k] - [s_{ij}^l] - s_{ij}k^l / 2
\end{align*}
\]

\(^4\)The case of linear objective without line limits is proved in [27] for radial networks. This result is extended here to mesh networks with line limits and convex objective functions.
At node $j$, we have
\[
\hat{s}_j = \frac{\hat{y}_j^2 - \hat{y}_j^2}{2} = [s_{\Psi,j}^2 - s_{\Psi,j}^2]_j - z_i j \epsilon/2
\]
\[
= \sum_{j \to k} [S_{\Psi,j}^2 - \sum_{k \to j} (S_{\Psi,k}^2)].\]
\[
+ y_j^2 z_j - z_i j \epsilon/2
\]
\[
= \sum_{j \to k} [S_{\Psi,j}^2 - \sum_{k \to j} (S_{\Psi,k}^2)].\]
\[
- ((S_{\Psi,j}^2 - z_j \epsilon/2) - z_{i} j \epsilon/2)
\]
\[
- \sum_{j \to k} [S_{\Psi,j}^2 - \sum_{k \to j} (S_{\Psi,k}^2)].\]
\[
Hence (13)–(14) hold at nodes $i, j$.

For (15) across link $(i, j)$:
\[
\hat{v}_j = [v_j^2 - 2(r_{ij}^2 + x_{ij}^2)Q^2_{\Psi,i}]
\]
\[
+ (r_{ij}^2 + x_{ij}^2) [Q_{\Psi,i}^2]
\]
\[
= \hat{v}_i - 2(r_{ij}^2 + x_{ij}^2) [Q_{\Psi,i}^2] + (r_{ij}^2 + x_{ij}^2) \hat{e}_{ij}
\]

For (22) across link $(i, j)$, we have
\[
\hat{v}_i \hat{e}_{ij} = \hat{P}_{ij}^2 + \hat{Q}_{ij}^2
\]
\[
- \hat{v}_i [v_i^2 - 2(r_{ij}^2 + x_{ij}^2)] + (r_{ij}^2 + x_{ij}^2)
\]
\[
- \hat{e}_i - 2 [r_{ij}^2 + x_{ij}^2] [Q_{\Psi,i}^2] + (r_{ij}^2 + x_{ij}^2) \hat{e}_{ij}
\]
\[
Since [v_j^2 - 2(r_{ij}^2 + x_{ij}^2)] + (r_{ij}^2 + x_{ij}^2) > 0, we can choose an $\epsilon > 0$ sufficiently small such that $\epsilon_{ij} \geq (\hat{P}_{ij}^2 + \hat{Q}_{ij}^2)/\hat{v}_i$.

This completes the proof.

Remark 1: Assumption A3 is used in the proof here to contradict the optimality of $(\hat{y}_j, s_j)$. Instead of A3, if $f(\hat{y}_j, s_j)$ is non-decreasing in $\epsilon$, the same argument shows that, given an optimal $(\hat{y}_j, s_j)$ with a strict inequality $\hat{y}_j \neq S_{\Psi,j}^2$, one can choose $\epsilon > 0$ to obtain another optimal point $(\hat{y}_j, s_j)$ that attains equality and has a cost $f(\hat{y}_j, s_j) > f(\hat{y}_j, s_j)$. Without A3, there is always an optimal solution of OPF-cr that is also optimal for OPF-ar, even though it is possible that the convex relaxation OPF-cr may also have other optimal points with strict inequality that are infeasible for OPF-ar.

Remark 2: The condition in Theorem 1 is equivalent to the “over-satisfaction of load” condition in [14] and [17]. It is needed because we have increased the loads $s_{\Psi,j}$ on buses $i$ and $j$ to obtain the alternative feasible solution $(\hat{y}_j, s_j)$ as we show in the simulations in [42], it is sufficient but not necessary. See also [35] and [36] for exact conic relaxation of OPF-cr for radial networks where this condition is replaced by other assumptions.

V. ANGLE RELAXATION

Theorem 1 justifies solving the convex problem OPF-cr for an optimal solution of OPF-ar. Given a solution $(\hat{y}_j, s_j)$ of OPF-ar, when and how can we recover a solution $(x, s)$ of the original OPF (11)–(12)? It depends on whether we can recover a solution $x$ to the branch flow equations (1)–(3) from $\hat{y}_j$, given any $s$.

Hence, for the rest of Section V, we fix an $\epsilon$. We abuse notation in this section and write $x, y, \Theta, V, Y, \hat{y}$ instead of $x(s), \hat{y}(s), \Theta(s), V(s), Y(s), \hat{y}(s)$, respectively.

A. Angle Recovery Condition

Fix a relaxed solution $\hat{y} := (S, \ell, v, s_0) \in \hat{Y}$. Define the $(n+1) \times m$ incidence matrix $C$ of $G$ by
\[
C_{ie} = \begin{cases} 
1 & \text{if link } e \text{ leaves node } i \\
-1 & \text{if link } e \text{ enters node } i \\
0 & \text{otherwise}.
\end{cases}
\]

The first row of $C$ corresponds to node 0 where $V_0 = |V_0|e^{i\theta_0}$ is given. In this paper we will only work with the $m \times n$ reduced incidence matrix $B$ obtained from $C$ by removing the first row (corresponding to $V_0$) and taking the transpose, i.e., for $e \in E, i = 1, \ldots, n$. Then $B$ can be partitioned into
\[
H = \begin{bmatrix} B_T & B_L \end{bmatrix}
\]
where the $n \times n$ submatrix $B_T$ corresponds to links in $T$ and the $(m-n) \times n$ submatrix $B_L$ corresponds to links in $T^\perp := G \backslash T$.

Let $\beta := \beta(\hat{y}) \in (-\pi, \pi)^m$ be defined by
\[
\beta_{ij} := \angle(v_i - s_{i,j}^2, s_{i,j}), \quad (i, j) \in E.
\]

Informally, $\beta_{ij}$ is the phase angle difference across link $(i, j)$ that is implied by the relaxed solution $\hat{y}$. Write $\beta$ as
\[
\beta = \begin{bmatrix} \beta_T \\ \beta_L \end{bmatrix}
\]
where $\beta_T$ is $n \times 1$ and $\beta_L$ is $(m-n) \times 1$.

Recall the projection mapping $\hat{h} : \mathbb{C}^{2m+n+1} \to \mathbb{R}^{3m+n+2}$ defined in (17)–(18). For each $\theta := (\theta_1, \ldots, \theta_n, \ldots, \theta_m) \in (-\pi, \pi)^n$, define the inverse projection $h_\theta : \mathbb{R}^{3m+n+2} \to \mathbb{C}^{2m+n+1}$ by $h_\theta(P, Q, \Theta, V, s_0) = (S_{ij}, I_{ij}, s_{0})$ where
\[
S_{ij} := P_{ij} + iQ_{ij}
\]
\[
I_{ij} := \sqrt{\epsilon_{ij}} e^{-i \beta_{ij}}
\]
\[
s_0 := p_0 + i q_0.
\]

These mappings are illustrated in Fig. 3.

By definition of $h(\mathbf{x})$ and $\hat{y}$, a branch flow solution in $\mathbf{x}$ can be recovered from a given relaxed solution $\hat{y}$ if $\hat{y}$ is in $\hat{h}(\mathbf{x})$ and cannot if $\hat{y}$ is in $\mathbf{y} \setminus \hat{h}(\mathbf{x})$. In other words, $\hat{h}(\mathbf{x})$ consists of exactly those points $\hat{y} \in \hat{Y}$ for which there exist $\theta$ such that their inverse projections $h_\theta(\hat{y})$ are in $\mathbf{x}$. Our next key result characterizes the exact condition under which such an inverse projection exists, and provides an explicit expression for recovering the phase angles $\angle V_i, \angle I_{ij}$ from the given $\hat{y}$.
A cycle $c$ in $G$ is an ordered list $c = (i_1, \ldots, i_k)$ of nodes in $N$ such that $(i_1 \sim i_2), \ldots, (i_k \sim i_1)$ are all links in $E$. We will use "$(i, j) \in c$" to denote a link $i \sim j$ in the cycle $c$. Each link $i \sim j$ may be in the same orientation $(i, j) \in E$ or in the opposite orientation $(j, i) \in E$. Let $\beta$ be the extension of $\beta$ from directed links to undirected links: if $(i, j) \in E$ then $\beta_{ij} := \beta_{ji}$ and $\beta_{ji} := -\beta_{ij}$. For any $d$-dimensional vector $\alpha$, let $P(\alpha)$ denote its projection onto $(-\pi, \pi]^d$ by taking modulo $2\pi$ componentwise.

**Theorem 2:** Let $T$ be any spanning tree of $G$. Consider a relaxed solution $\hat{y} \in \hat{Y}$ and the corresponding $\beta = \hat{\beta}(\hat{y})$ defined in (27):

1) There exists a unique $\theta_\ast \in (-\pi, \pi]^n$ such that $h_{\theta_\ast}(\hat{y})$ is a branch flow solution in $\mathbf{X}$ if and only if
   \[ B_{\perp}B_T^{-1}\beta_T = \beta_{\perp} \pmod{2\pi}. \]  

2) The angle recovery condition (33) holds if and only if for every cycle $c$ in $G$
   \[ \sum_{(i,j) \in c} \beta_{ij} = 0 \pmod{2\pi}. \]  

3) If (33) holds, then $\theta_\ast = P(\hat{H}_{\perp}^{-1}\beta_T)$.

**Remark 3:** Given a relaxed solution $\hat{y}$, Theorem 2 prescribes a way to check if a branch flow solution can be recovered from it, and if so, the required computation. The angle recovery condition (33) depends only on the network topology through the reduced incidence matrix $B$. The choice of spanning tree $T$ corresponds to choosing $n$ linearly independent rows of $B$ to form $H_T$ and does not affect the conclusion of the theorem.

**Remark 4:** When it holds, the angle recovery condition (34) has a familiar interpretation (due to Lemma 3 below): the voltage angle differences (implied by $\hat{y}$) sum to zero $\pmod{2\pi}$ around any cycle.

**Remark 5:** A direct consequence of Theorem 2 is that the relaxed branch flow model (13)–(16) together with the angle recovery condition (33) is equivalent to the original branch flow model (1)–(3). That is, $x$ satisfies (1)–(3) if and only if $\hat{y} = \hat{h}(x)$ satisfies (13)–(16) and (33). The challenge in computing a branch flow solution $x$ is that (33) is nonconvex.

The proof of Theorem 2 relies on the following important lemma that gives a necessary and sufficient condition for an inverse projection $h_\theta(\hat{y})$ defined by (29)–(32) to be a branch flow solution in $\mathbf{X}$. Fix any $\hat{y} := (S, \ell, v, s_0)$ in $\hat{Y}$ and the corresponding $\beta := \beta(\hat{y})$ defined in (27). Consider the equation

\[ B\theta = \beta + 2\pi k \]  

where $k \in \mathbb{N}^m$ is an integer vector. Since $G$ is connected, $m \geq n$ and $\text{rank}(B) = n$. Hence, given any $k$, there is at most one $\theta$ that solves (35). Obviously, given any $\theta$, there is exactly one $k$ that solves (35); we denote it by $k(\theta)$ when we want to emphasize the dependence on $\theta$. Given any solution $[\theta, k]$ with $\theta \in (-\pi, \pi]^n$, define its equivalence class by\(^5\)

\[ \sigma(\theta, k) := \{ \theta + 2\pi\alpha, k + B\alpha \} \quad \alpha \in \mathbb{N}^n. \]

\(^5\)Using the connectedness of $G$ and the definition of $B$, one can argue that $\alpha$ must be an integer vector for $k + B\alpha$ to be integral.

We say $\sigma(\theta, k)$ is a solution of (35) if every vector in $\sigma(\theta, k)$ is a solution of (35), and $\sigma(\theta, k)$ is the unique solution of (35) if it is the only equivalence class of solutions.

**Lemma 3:** Given any $\hat{y} := (S, \ell, v, s_0)$ in $\hat{Y}$ and the corresponding $\beta := \beta(\hat{y})$ defined in (27):

1) $h_\theta(\hat{y})$ is a branch flow solution in $\mathbf{X}$ if and only if $(\theta, k(\theta))$ solves (35).

2) there is at most one $\sigma(\theta, k)$, $\theta \in (-\pi, \pi]^n$, that is the unique solution of (35), when it exists.

**Proof:** Suppose $(\theta, k)$ is a solution of (35) for some $k = k(\theta)$. We need to show that (13)–(16) together with (29)–(32) and (35) imply (1)–(3). Now (13) and (14) are equivalent to (3). Moreover (16) and (29)–(31) imply (2). To prove (1), substitute (2) into (35) to get

\[ \theta_i - \theta_j = \angle(V_i - z_{ij}I_{ij}) + 2\pi k_{ij}. \]

Hence

\[ \angle V_i = \theta_i - \angle(V_i - z_{ij}I_{ij}) = 2\pi k_{ij}. \]

From (15) and (2), we have

\[ V_i = V_i - z_{ij}I_{ij} \]

This and (36) imply $V_i = V_i - z_{ij}I_{ij}$ which is (1).

Conversely, suppose $h_\theta(\hat{y}) \in \mathbf{X}$. From (1) and (2), we have $V_i = V_i - z_{ij}I_{ij}$. Then $b_i - b_j = \beta_{ij} + 2\pi k_{ij}$ for some integer $k_{ij} = k_{ij}(\theta)$. Hence $(\theta, k)$ solves (35).

The discussion preceding the lemma shows that, given any $k \in \mathbb{N}^m$, there is at most one $\theta$ that satisfies (35). If no such $\theta$ exists for any $k \in \mathbb{N}^m$, then (35) has no solution $(\theta, k)$. If (35) has a solution $(\theta, k)$, then clearly $(\theta + 2\pi\alpha, k + B\alpha)$ are also solutions for all $\alpha \in \mathbb{N}^n$. Hence we can assume without loss of generality that $\theta \in (-\pi, \pi]^n$. We claim that $\sigma(\theta, k)$ is the unique solution of (35). Otherwise, there is a $(\theta, k) \not\in \sigma(\theta, k)$ with $B\theta = \beta + 2\pi k$. Then $B(\theta-\theta) = 2\pi k$ for some $\alpha$. Since $k \in \mathbb{N}^m$, $\alpha$ is an integer vector; moreover $\theta$ is unique given $k$. This means $(\theta, k) \in \sigma(\theta, k)$, a contradiction.

**Proof of Theorem 2:** Since $m \geq n$ and $\text{rank}(B) = n$, we can always find $n$ linearly independent rows of $B$ to form a basis. The choice of this basis corresponds to choosing a spanning tree of $G$, which always exists since $G$ is connected [44, Ch. 5]. Assume without loss of generality that the first $n$ rows are such a basis so that $\theta$ and $\beta$ are partitioned as in (26) and (28), respectively. Then Lemma 3 implies that $h_\theta(\hat{y}) \in \mathbf{X}$ with $\theta_\ast \in (-\pi, \pi]^n$ if and only if $(\theta_\ast, k_\ast(\theta_\ast))$ is the unique solution of

\[ B_T^{-1}\theta = \beta_T + 2\pi k_T \]

Since $T$ is a spanning tree, the $n \times n$ submatrix $B_T$ is invertible. Moreover (37) has a unique solution if and only if $B_T^{-1}(B_T + 2\pi k_T) = B_T^{-1} + 2\pi k_T$, i.e., $B_T B_T^{-1} \beta_T = B_T^{-1} + 2\pi k_T$, where $k_T := \perp - B_T^{-1} B_T k_T$. Then (38) below implies that $k_T$ is an integer vector. This proves the first assertion.
For the second assertion, recall that the spanning tree $T$ defines the orientation of all links in $T$ to be directed away from the root node 0. Let $T'(i \rightarrow j)$ denote the unique path from node $i$ to node $j$ in $T$; in particular, $T'(0 \rightarrow j)$ consists of links all with the same orientation as the path and $T'(j \rightarrow 0)$ of links all with the opposite orientation. Then it can be verified directly that

$$[B_T^{-1}]_{e,i} := \begin{cases} -1 & \text{if link } e \text{ is in } T(0 \rightarrow i) \\ 0 & \text{otherwise.} \end{cases} \quad (38)$$

Hence $B_T^{-1} \beta_T$ represents the (negative of the) sum of angle differences on the path $T(0 \rightarrow i)$ for each node $i \in T$:

$$[H_T^{-1} \beta_T]_i = - \sum_{e \in T(0 \rightarrow i)} [\beta_T]_e = - \sum_{v \in T(0 \rightarrow i)} [\beta_T]_{v,e}.$$

Hence $B_T^{-1} \beta_T$ is the sum of voltage angle differences from node $i$ to node $j$ along the unique path in $T$, for every link $(i, j) \in E \setminus E_T$ not in the tree $T$. To see this, we have, for each link $e := (i, j) \in E \setminus E_T$:

$$[B_T^{-1} \beta_T]_e = - \sum_{e' \in T(0 \rightarrow j)} [\beta_T]_{e'} - \sum_{e'' \in T(0 \rightarrow i)} [\beta_T]_{e''}.$$

Since

$$\sum_{e' \in T(0 \rightarrow j)} [\beta_T]_{e'} = - \sum_{e'' \in T(0 \rightarrow i)} [\beta_T]_{e''}$$

the angle recovery condition (33) is equivalent to

$$\sum_{e' \in T(0 \rightarrow j)} [\beta_T]_{e'} + [\beta_T]_{ij} + \sum_{e'' \in T(0 \rightarrow i)} [\beta_T]_{e''} = \sum_{e' \in T(0 \rightarrow i)} [\beta_T]_{e'} = 0 \pmod{2\pi}$$

where $c(i, j)$ denotes the unique basis cycle (with respect to $T$) associated with each link $(i, j)$ not in $T$ [44, Ch. 5]. Hence (33) is equivalent to (34) on all basis cycles, and therefore it is equivalent to (34) on all cycles.

Suppose (33) holds and let $(\theta_*, k_*)$ be the unique solution of (37) with $\theta_* \in (-\pi, \pi]^n$. We are left to show that $\theta_* = \mathcal{P}(B_T^{-1} \beta_T)$. By (37) we have $\theta_* = 2\pi B_T^{-1} [k_*]_I = \beta_T$. Consider $\alpha := -H_T^{-1} [k_*]_T$ which is in $\mathbb{R}^n$ due to (38). Then $[\theta_* + 2\pi \alpha, k_* + B_T \alpha] \in \sigma(\theta_*, k_*)$ and hence is also a solution of (37) by Lemma 3. Moreover $[\theta_* + 2\pi \alpha]_I = B_T^{-1} [k_*]_I \neq 0$. This means that $\theta_*$ is given by $\mathcal{P}(B_T^{-1} \beta_T)$ since $\theta_* \in (-\pi, \pi]^n$.

B. Angle Recovery Algorithms

Theorem 2 suggests a centralized method to compute a branch flow solution from a relaxed solution.

**Algorithm 1: centralized angle recovery.** Given a relaxed solution $\hat{\theta}_* \in \hat{Y}$:

1) Choose any $n$ basis rows of $B$ and form $B_T, B_L$.
2) Compute $\beta$ from $\hat{\theta}$ and check if $B_B^{-1} \beta_B - \beta_L = 0 \pmod{2\pi}$.
3) If not, then $\hat{\theta} \notin \hat{h}(\mathcal{X})$; stop.
4) Otherwise, compute $\theta_* = \mathcal{P}(B_T^{-1} \beta_T)$.
5) Compute $h_{\theta_*}(\hat{\theta}_*) \in \mathcal{X}$ through (29)–(32).

Theorem 2 guarantees that $h_{\theta_*}(\hat{\theta}_*)$, if exists, is the unique branch flow solution of (1)–(3) whose projection is $\hat{y}$.

The relations (2) and (35) motivate an alternative procedure to compute the angles $\mathcal{L}_{I_{jk}}, \mathcal{L}_{V_k}$, and a branch flow solution. This procedure is more amenable to a distributed implementation.

**Algorithm 2: distributed angle recovery.** Given a relaxed solution $\hat{\theta}_* \in \hat{Y}$:

1) Choose any spanning tree $T$ of $G$ rooted at node 0.
2) For $j = 0, 1, \ldots, n$ (i.e., as $j$ ranges over the tree $T$, starting from the root and in the order of breadth-first search), for all children $k$ with $j \rightarrow k$, set

$$\mathcal{L}_{I_{jk}} := \mathcal{L}_{V_j} - \mathcal{L}_{S_{jk}} \quad (39)$$

$$\mathcal{L}_{V_k} := \mathcal{L}_{V_j} - (v_j - z_{jk}^* S_{jk}). \quad (40)$$

3) For each link $(j, k) \in E \setminus E_T$ not in the spanning tree, node $j$ is an additional parent of $k$ in addition to $k$’s parent in the spanning tree from which $\mathcal{L}_{V_k}$ has already been computed in Step 2.

a) Compute current angle $\mathcal{L}_{I_{jk}}$ using (39).

b) Compute a new voltage angle $\theta_{jk}^*$ using the new parent $j$ and (40). If $\theta_{jk}^* \mathcal{L}_{V_k} \neq 0 \pmod{2\pi}$, then angle recovery has failed; stop.

If the angle recovery procedure succeeds in Step 3, then $\hat{\theta}_*$ together with these angles $\mathcal{L}_{V_k}, \mathcal{L}_{I_{jk}}$ are indeed a branch flow solution. Otherwise, a link $(j, k)$ not in the tree $T$ has been identified where condition (34) is violated over the unique basis cycle (with respect to $T$) associated with link $(j, k)$.

C. Radial Networks

Recall that all relaxed solutions in $\hat{Y} \setminus \hat{h}(\mathcal{X})$ are spurious. Our next key result shows that, for radial network, $\hat{h}(\mathcal{X}) = \hat{Y}$ and hence angle relaxation is always exact in the sense that there is always a unique inverse projection that maps any relaxed solution $\hat{\theta}_*$ to a branch flow solution in $\mathcal{X}$ (even though $\mathcal{X} \neq \hat{Y}$).

**Theorem 4:** Suppose $G = T$ is a tree. Then

1) $\mathcal{h}(\mathcal{X}) = \hat{Y}$.

2) Given any $\hat{\theta}_* : = \mathcal{P}(B_T^{-1} \beta_T)$ always exists and is the unique vector in $(-\pi, \pi]^n$ such that $h_{\theta_*}(\hat{\theta}_*) \in \mathcal{X}$.

**Proof:** When $G = T$ is a tree, $m = n$ and hence $B = B_T$ and $\beta = \beta_T$. Moreover $H$ is $n \times n$ and of full rank. Therefore $\theta_* = \mathcal{P}(B_T^{-1} \beta_T) \in \mathcal{X}$ always exists and, by Theorem 2, $h_{\theta_*}(\hat{\theta}_*)$ is the unique branch flow solution in $\mathcal{X}$ whose projection is $\hat{y}$. Since this holds for any arbitrary $\hat{\theta}_* \in \hat{Y}$, $\hat{Y} = \hat{h}(\mathcal{X})$.

A direct consequence of Theorem 1 and Theorem 4 is that, for a radial network, OPF is equivalent to the convex problem OPF-cr in the sense that we can obtain an optimal solution of one problem from that of the other.

**Corollary 5:** Suppose $G$ is a tree. Given any optimal solution $(\hat{y}_*, s_*)$ of OPF-cr, there exists a unique $\theta_* \in (-\pi, \pi]^n$ such that $h_{\theta_*}(\hat{y}_*, s_*)$ is optimal for OPF.
VI. CONCLUSION

We have presented a branch flow model for the analysis and optimization of mesh as well as radial networks. We have proposed a solution strategy for OPF that consists of two steps:

1) Compute a relaxed solution of OPF-ar by solving its second-order conic relaxation OPF-cr.
2) Recover from a relaxed solution an optimal solution of the original OPF using an angle recovery algorithm, if possible.

We have proved that this strategy guarantees a globally optimal solution for radial networks, provided there are no upper bounds on loads. For mesh networks the angle recovery condition may not hold but can be used to check if a given relaxed solution is globally optimal.

The branch flow model is an alternative to the bus injection model. It has the advantage that its variables correspond directly to physical quantities, such as branch power and current flows, and therefore are often more intuitive than a semidefinite matrix in the bus injection model. For instance, Theorem 2 implies that the number of power flow solutions depends only on the magnitude of voltages and currents, not on their phase angles.

ACKNOWLEDGMENT

The authors would like to thank S. Bose, K. M. Chandy, and L. Gan of Caltech; C. Clarke, M. Montoya, and R. Sherick of the Southern California Edison (SCE); and B. Lesieutre of Wisconsin for helpful discussions.

REFERENCES


Masoud Farivar received the B.S. degree with a double major in electrical engineering and computer science from Sharif University of Technology, Tehran, Iran, in 2009. He is pursuing the Ph.D. degree in electrical engineering at California Institute of Technology, Pasadena, CA, USA.

He is a fellow of the Resnick Sustainability Institute at California Institute of Technology. Since June of 2010, he has also been working in the Advanced Technology division of Southern California Edison on a number of smart grid projects.

Steven H. Low (F’08) received the B.S. degree from Cornell University, Ithaca, NY, USA, and the Ph.D. from the University of California, Berkeley, CA, USA, both in electrical engineering.

He is a Professor of the Computing and Mathematical Sciences and Electrical Engineering Departments at California Institute of Technology, Pasadena, CA, USA, and an adjunct professor of both Swinburne University, Australia, and Shanghai Jiao Tong University, China. Before that, he was with AT&T Bell Laboratories, Murray Hill, NJ, USA, and the University of Melbourne, Australia.

Prof. Low was a co-recipient of IEEE best paper awards, the R&D 100 Award, and an Okawa Foundation Research Grant. He was a member of the Networking and Information Technology Technical Advisory Group for the US President’s Council of Advisors on Science and Technology (PCAST) in 2006. He is a Senior Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the mentor for the annual JSAC series on Smart Grid Communications, a Senior Editor of the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, and a Steering Committee Member of the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING. He also served on the editorial boards of IEEE/ACM Transactions on Networking, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, ACM Computing Surveys, Computer Networks Journal, and NOW Foundations and Trends in Networking.
Abstract—We propose a branch flow model for the analysis and optimization of mesh as well as radial networks. The model leads to a new approach to solving optimal power flow (OPF) that consists of two relaxation steps. The first step eliminates the voltage and current angles and the second step approximates the resulting problem by a conic program that can be solved efficiently. For radial networks, we prove that both relaxation steps are always exact, provided there are no upper bounds on loads. For mesh networks, the conic relaxation is always exact but the angle relaxation may not be exact, and we provide a simple way to determine if a relaxed solution is globally optimal. We propose convexification of mesh networks using phase shifters so that OPF for the convexified network can always be solved efficiently for an optimal solution. We prove that convexification requires phase shifters only outside a spanning tree of the network and their placement depends only on network topology, not on power flows, generation, loads, or operating constraints. Part I introduces our branch flow model, explains the two relaxation steps, and proves the conditions for exact relaxation. Part II describes convexification of mesh networks, and presents simulation results.

Index Terms—Convex relaxation, load flow control, optimal power flow, phase control, power system management.

I. INTRODUCTION

In Part I of this two-part paper [2], we introduce a branch flow model that focuses on branch variables instead of nodal variables. We formulate optimal power flow (OPF) within the branch flow model and propose two relaxation steps. The first step eliminates phase angles of voltages and currents. We call the resulting problem OPF-ar which is still nonconvex. The second step relaxes the feasible set of OPF-ar to a second-order cone. We call the resulting problem OPF-cr which is convex, indeed a second-order cone program (SOCP) when the objective function is linear. We prove that the conic relaxation OPF-cr is always exact even for mesh networks, provided there are no upper bounds on real and reactive loads, i.e., any optimal solution of OPF-cr is also optimal for OPF-ar. Given an optimal solution of OPF-ar, whether we can derive an optimal solution to the original OPF depends on whether we can recover the voltage and current angles correctly from the given OPF-ar solution. We characterize the exact condition (the angle recovery condition) under which this is possible, and present two angle recovery algorithms. It turns out that the angle recovery condition has a simple interpretation: any solution of OPF-ar implies a phase angle difference across a line, and the angle recovery condition says that the implied phase angle differences sum to zero (mod $2\pi$) around each cycle. For a radial network, this condition holds trivially and hence solving the conic relaxation OPF-cr always produces an optimal solution for the original OPF. For a mesh network, the angle recovery condition may not hold, and our characterization can be used to check if a relaxed solution yields an optimal solution for OPF.

In this paper, we prove that, by placing phase shifters on some of the branches, any relaxed solution of OPF-ar can be mapped to an optimal solution of OPF for the convexified network, with an optimal cost that is no higher than that of the original network. Phase shifters thus convert an NP-hard problem into a simpler problem. Our result implies that when the angle recovery condition holds for a relaxed branch flow solution, not only is the solution optimal for the OPF without phase shifters, but the addition of phase shifters cannot further reduce the cost. On the other hand, when the angle recovery condition is violated, then the convexified network may have a strictly lower optimal cost. Moreover, this benefit can be attained by placing phase shifters only outside an arbitrary spanning tree of the network graph.

There are in general many ways to choose phase shifter angles to convexify a network, depending on the number and location of the phase shifters. While placing phase shifters on each link outside a spanning tree requires the minimum number of phase shifters to guarantee exact relaxation, this strategy might require relatively large angles at some of these phase shifters. On the other extreme, one can choose to minimize (the Euclidean norm of) the phase shifter angles by deploying phase shifters on every link in the network. We prove that this minimization problem is NP-hard. Simulations suggest, however, that a simple heuristic works quite well in practice.

These results lead to an algorithm for solving OPF when there are phase shifters in mesh networks, as summarized in Fig. 1. Since power networks in practice are very sparse, the number of lines not in a spanning tree can be relatively small compared to the number of buses squared, as demonstrated in simulations in Section V using the IEEE test systems with 14, 30, 57, 118, and 300 buses, as well as a 39-bus model of a New England power system and two models of a Polish power system with more than 2000 buses. Moreover, the placement of these phase shifters depends only on network topology, but not on power...
flows, generations, loads, or operating constraints. Therefore only one-time deployment cost is required to achieve subsequent simplicity in network operation. Even when phase shifters are not installed in the network, the optimal solution of a convex relaxation is useful in providing a lower bound on the true optimal objective value. This lower bound serves as a benchmark for other heuristic solutions of OPF.

The paper is organized as follows. In Section II, we extend the branch flow model of [2] to include phase shifters. In Section III, we describe methods to compute phase shifter angles to map any relaxed solution to a branch flow solution. In Section IV, we explain how to use phase shifters to simplify OPF. In Section V, we present our simulation results.

II. BRANCH FLOW MODEL WITH PHASE SHIFTERS

We adopt the same notations and assumptions A1–A4 of [2].

A. Review: Model Without Phase Shifters

For ease of reference, we reproduce the branch flow model of [2] here:

\[ I_{ij} = y_{ij}(V_i - V_j) \]  
\[ S_{ij} = V_I I_{ij} \]  
\[ s_j = \sum_{k:j\rightarrow k} S_{jk} - \sum_{i:i\rightarrow j} (S_{ij} - z_{ij}|I_{ij}|^2) + y_{ij}^2 V_j^2. \]

Recall the set \( X(s) \) of branch flow solutions given \( s \) defined in [2]:

\[ X(s) := \{ x := (S, I, V, s_0) \in \mathbb{R}^{n \times n} \mid x \text{ solves (1) - (3) given } s \} \]

and the set \( X \) of all branch flow solutions:

\[ X := \bigcup_{s \in \Sigma^n} X(s). \]

To simplify notation, we often use \( X \) to denote the set defined either in (4) or in (5), depending on the context. In this section we study power flow solutions and hence we fix an \( s \). All quantities, such as \( x, \hat{y}, X, \bar{X}, \bar{X}_T \), are with respect to the given \( s \), even though that is not explicit in the notation. In the next section, \( s \) is also an optimization variable and the sets \( X, \hat{X}, \bar{X}, \bar{X}_T \) are for any \( s \).

Given a relaxed solution \( \hat{y} \), define \( \beta := \beta(\hat{y}) \) by

\[ \beta_{ij} := \alpha\{v_i - z_{ij}|I_{ij}|^2\}, \quad (i, j) \in E. \]  

(6)

It is proved in [2, Theorem 2] that a given \( \hat{y} \) can be mapped to a branch flow solution in \( X \) if and only if there exists a \( (\theta, k) \) that solves

\[ B\theta = \beta + 2\pi k \]  

(7)

for some integer vector \( k \in \mathbb{N}^n \). Moreover if (7) has a solution, then it has a countably infinite set of solutions \( (\theta, k) \), but they are relatively unique, i.e., given \( k \), the solution \( \theta \) is unique, and given \( \theta \), the solution \( k \) is unique. Hence (7) has a unique solution \( (\theta_*, k_*) \) with \( \theta_* \in (-\pi, \pi]^n \) if and only if

\[ B \perp B_T^{-1} \beta_T = \beta_\perp \pmod{2\pi} \]  

(8)

which is equivalent to the requirement that the (implied) voltage angle differences sum to zero around any cycle \( c \):

\[ \sum_{(i, j) \in c} \hat{\beta}_{ij} = 0 \pmod{2\pi} \]

where \( \hat{\beta}_{ij} = \beta_{ij} \) if \( (i, j) \in E \) and \( \hat{\beta}_{ij} = -\beta_{ji} \) if \( (j, i) \in E \).

B. Model With Phase Shifters

Phase shifters can be traditional transformers or Flexible AC Transmission Systems (FACTS) devices. They can increase transmission capacity and improve stability and power quality [3], [4]. In this paper, 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. Specifically, consider an idealized phase shifter parametrized by \( \phi_{ij} \) across line \( (i, j) \), as shown in Fig. 2. 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 \( \phi_{ij} \) and line impedance \( z_{ij} \). Let \( V_k \) and \( I_k \) be the voltage at \( k \) and current from \( k \) to \( j \), respectively. Then the effect of the idealized phase shifter is summarized by the following modeling assumption:

\[ V_k = V_i e^{i\phi_{ij}}, \quad \text{and} \quad I_k = I_{ij} e^{i\phi_{ij}}. \]

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:

\[ I_{ij} = y_{ij}(V_i - V_j e^{-i\phi_{ij}}) \]  

(9)

\[ S_{ij} = V_i I_{ij}^* \]  

(10)

\[ s_j = \sum_{k:j\rightarrow k} S_{jk} - \sum_{i:i\rightarrow j} (S_{ij} - z_{ij}|I_{ij}|^2) + y_{ij}^2 V_j^2. \]  

(11)
Without phase shifters \((\phi_{ij} = 0)\), (9)–(11) reduce to the branch flow model (1)–(3).

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

\[
\overline{X} := \{x | x \text{solves } (9) - (11) \text{ for some } \phi \}.
\] (12)

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

\[
\overline{X}_T := \{x | x \text{solves } (9) - (11) \text{ for some } \phi \in T^{\perp}\}.
\] (13)

Since (9)–(11) reduce to the branch flow model when \(\phi = 0\), \(\overline{X} \subseteq \overline{X}_T \subseteq \overline{X} \).

### III. Phase Angle Setting

Given a relaxed solution \(\tilde{y}\), there are in general many ways to choose angles \(\phi\) on the phase shifters to recover a feasible branch flow solution \(x \in \overline{X}\) from \(\tilde{y}\). They depend on the number and location of the phase shifters.

#### A. Computing \(\phi\)

For a network with phase shifters, we have from (9) and (10)

\[
S_{ij} = V_iV_j^*e^{i\theta_{ij}} - \frac{1}{z_{ij}^*} = V_iV_j^*e^{i\theta_{ij}} - V_iV_j^*e^{i\phi_{ij}}
\]

leading to \(V_iV_j^*e^{i\theta_{ij}} = V_iV_j^*e^{i\phi_{ij}}\). Hence \(\theta_{ij} - \phi_{ij} = \beta_{ij} - \phi_{ij} + 2\pi k_{ij}\) for some integer \(k_{ij}\). This changes the angle recovery condition in [2, Theorem 2] from whether there exists \((\theta, k)\) that solves (7) to whether there exists \((\theta, \phi, k)\) that solves

\[
B\theta - \beta = \phi + 2\pi k
\] (14)

for some integer vector \(k \in (-2\pi, 2\pi]^m\). The case without phase shifters corresponds to setting \(\phi = 0\).

We now describe two ways to compute \(\phi\): the first minimizes the required number of phase shifters, and the second minimizes the size of phase angles.

1) **Minimize Number of Phase Shifters:** Our first key result implies that, given a relaxed solution \(\tilde{y} := (S, \ell, v, s_0) \in \overline{Y}\), we can always recover a branch flow solution \(x := (S, I, V, s_0) \in \overline{X}\) of the convexified network. Moreover it suffices to use phase shifters in branches only outside a spanning tree. This method requires the smallest number \((m - n)\) of phase shifters.

Given any \(d\)-dimensional vector \(\alpha\), let \(P(\alpha)\) denote its projection onto \((-\pi, \pi]^d\) by taking modulo \(2\pi\) componentwise.

**Theorem 1:** Let \(T\) be any spanning tree of \(G\). Consider a relaxed solution \(\tilde{y} \in \overline{Y}\) and the corresponding \(\beta\) defined by (6) in terms of \(\tilde{y}\).

1) There exists a unique \((\theta^*, \phi^*) \in (-\pi, \pi]^n + m\) with \(\phi^* \in T^\perp\) such that \(h_e(\tilde{y}) \in \overline{X}_T\), i.e., \(h_e(\tilde{y})\) is a branch flow solution of the convexified network. Specifically

\[
\theta^* = P(\begin{bmatrix} B_T^{-1} \beta_T \\ 0 \end{bmatrix})
\]

\[
\phi^* = P(\begin{bmatrix} \beta_T \\ \beta_T - B_T^{-1} \beta_T \end{bmatrix})
\]

2) \(\tilde{y} = \overline{X} = \overline{X}_T\) and hence \(\tilde{y} = \overline{h}(\tilde{X}) = \overline{h}(\overline{X}_T)\).

**Proof:** For the first assertion, write \(\phi = [\phi_T \phi_I]^T\) and set \(\phi_T = 0\). Then (14) becomes

\[
\begin{bmatrix} H_T \theta \\ B_1 \theta \end{bmatrix} = \begin{bmatrix} \beta_T \\ \beta_T - \phi_I \end{bmatrix} + 2\pi \begin{bmatrix} k_T \\ k_I \end{bmatrix}
\] (15)

We now argue that there always exists a unique \((\theta^*, \phi^*, k)\), with \(\theta^* \in (-\pi, \pi]^m\), \(\phi^* \in (-\pi, \pi]^m\) and \(\phi^* \in T^\perp\), that solves (15) for some \(k \in \mathbb{N}^m\).

The same argument as in the proof of [2, Theorem 2] shows that a vector \((\theta^*, \phi^*, k)\) with \(\theta^* \in (-\pi, \pi]^m\) and \(\phi^* \in T^\perp\) is a solution of (15) if and only if

\[
B_1 B_T^{-1} \beta_T - \beta = 2\pi \begin{bmatrix} k_T \\ k_I \end{bmatrix}
\] (16)

Note that given \(\theta_\perp\), \(k_{I, T}\) is uniquely determined since \([\phi_\perp, k_{\perp, T}]\) can be freely chosen to satisfy (16). Hence we can choose the unique \(k_{I, T}\), such that \(\phi_\perp = \theta_\perp\) and solves (16) for some \(k \in \mathbb{N}^m\). Moreover this unique vector \((\theta^*, \phi^*)\) is given by the formulae in the theorem.

The second assertion follows from assertion 1.

2) **Minimize Phase Angles:** The choice of \((\theta^*, \phi^*)\) in Theorem 1 has the advantage that it requires the minimum number of phase shifters (only on links outside an arbitrary spanning tree \(T\)). It might however require relatively large angles \([\phi^*]_e\) at some links \(e\) outside \(T\). On the other extreme, suppose we have phase shifters on every link. Then one can choose \((\theta^*, \phi^*)\), such that the phase shifter angles are minimized.

Specifically we are interested in a solution \((\theta, \phi, k)\) of (14) that minimizes \(\|P(\phi)e\|^2\) where \(\|\cdot\|\) denotes the Euclidean norm of \(\phi\) after taking modulo \(2\pi\) componentwise. Hence we are interested in solving the following problem: given \(B, \beta\)

\[
\min_{\theta, \phi, k, l} \|\phi - 2\pi l\|^2
\] (17)

subject to

\[
B\theta - \beta (\phi + 2\pi k)
\] (18)

where \(k, l \in \mathbb{N}^m\) are integer vectors.

**Theorem 2:** The problem (17), (18) of minimum phase angles is NP-hard.
Fig. 3. Each lattice point corresponds to \(2\pi k\) for a \(k \in \mathbb{N}^m\). The constrained optimization (19) is to find a lattice point that is closest to the range space \(\{B\theta|\theta \in \mathbb{R}^n\}\) of \(B\). The shaded region around the origin is \((-\pi, \pi]^m\) and contains a point \(\beta':=\beta+2\pi k\) for exactly one \(k \in \mathbb{N}^m\). Our approximate solution corresponds to solving (20) for this fixed \(k\).

Proof: Clearly the problem (17), (18) is equivalent to the following unconstrained minimization [eliminate \(\phi\) from (17), (18)]:

\[
\min_{k \in \mathbb{N}^m} \min_{\theta \in \mathbb{R}^n} \|\beta + 2\pi k - B\theta\|^2.
\]  
It thus solves for a lattice point \(\beta + 2\pi k\) that is closest to the range space \(\{B\theta|\theta \in \mathbb{R}^n\}\) of \(B\), as illustrated in Fig. 3.

Fix any \(k \in \mathbb{N}^m\). Consider \(\beta':=\beta + 2\pi k\) and the inner minimization in (19):

\[
\min_{\theta \in \mathbb{R}^n} \|\beta' - B\theta\|^2.
\]  
This is the standard linear least-squares estimation where \(\beta'\) represents an observed vector that is to be estimated by an vector in the range space of \(B\) in order to minimize the normed error squared. The optimal solution is

\[
\theta_* := (B^T B)^{-1} B^T \beta' 
\]
\[
\beta' - B\theta_* - \left( I - B (B^T B)^{-1} B^T \right) \beta'.
\]

Substituting (22) and (20) into (19), (19) becomes

\[
\min_{k \in \mathbb{N}^m} \|\gamma + 2\pi A k\|^2
\]  
where \(\gamma := AB \in \mathbb{R}^m\) and \(A := I - B (B^T B)^{-1} B^T\) is the orthogonal complement of the range space of \(B\). But (23) is the closest lattice vector problem and is known to be NP-hard [5].

Remark 1: Since the objective function is strictly convex, the phase angles \(\phi_* = (\beta + 2\pi k_*) - B\theta_*\) at optimality will lie in \((-\pi, \pi]^m\). Moreover, if an optimal solution exists, then there is always an optimal solution with \(\theta_*\) in \((-\pi, \pi]^m\); if \(\theta, k\) is optimal for (19) with \(\theta \not\in (-\pi, \pi]^m\), then by writing \(k' =: B\alpha + k\) for integer vectors \(\alpha \in \mathbb{N}^n, k' \in \mathbb{N}^m\), the objective function in (19) becomes

\[
(\beta + 2\pi k') - B(\theta - 2\pi \alpha)
\]  
i.e., we can always choose \((k_*, \alpha_*)\) so that \(\theta_* := \theta - 2\pi \alpha_*\) lies in \((-\pi, \pi]^n\) and \(k = H\alpha_* + k_*\). Therefore, given an optimal solution \((\theta, k)\) with \(\theta \not\in (-\pi, \pi]^n\), we can find another point \((\theta_*, k_*)\) with \(\theta_* = P(\theta) \in (-\pi, \pi]^n\) that is also optimal.

Many algorithms have been proposed to solve the closest lattice vector problem. See [6] for state-of-the-art algorithms. Given \(\beta\), there is a unique \(k\) such that \(\beta' := \beta + 2\pi k\) is in \((-\pi, \pi]^m\), as illustrated in the shaded area of Fig. 3. A simple heuristic that provides an upper bound on (19) is to solve (20) for this fixed \(k\). From (21)–(22), the heuristic solution is

\[
\theta_* := \mathcal{P}\left((B^T B)^{-1} B^T \beta'\right)
\]
\[
\phi_* := \left(I - B (B^T B)^{1/2} B^T\right) \beta'.
\]

This approximate solution is illustrated in Section V and seems to be effective in reducing the phase shifter angles (\(k = 0\) in all our test cases).

B. Arbitrary Network of Phase Shifters

More generally, consider a network with phase shifters on an arbitrary subset of links. Given a relaxed solution \(\phi\), under what condition does there exists a \(\theta\) such that the inverse projection \(h_\theta(\phi)\) is a branch flow solution in \(X\)? If there is a spanning tree \(T\) such that all links outside \(T\) have phase shifters, then Theorem 1 says that such a \(\theta\) always exists, with an appropriate choice of phase shifter angles on non-tree links. Suppose no such spanning tree exists, i.e., given any spanning tree \(T\), there is a set \(E_L \subseteq E \setminus E_T\) of links that contain no phase shifters. Let \(H_\lambda\) and \(\beta_\lambda\) denote the submatrix of \(H\) and subvector of \(\beta\), respectively, corresponding to these links. Then a necessary and sufficient condition for angle recovery is: there exists a spanning tree \(T\) such that the associated \(B_{\perp}, \beta_{\perp}\) satisfy

\[
B_{\perp} B_{\perp}^{-1} \beta_{\perp} = \beta_{\perp}. \quad (\text{mod } 2\pi).
\]

This condition reduces to (8) if there are no phase shifters in the network \(\{E_{\perp} = E \setminus E_T\}\) and is always satisfied if every link outside any spanning tree has a phase shifter \((E_{\perp} = \emptyset)\). It requires that the angle differences implied by \(\tilde{\gamma}\) sum to zero (mod \(2\pi\)) around any loop that contains no phase shifter (c.f. [2, Theorem 2(1) and Remark 4]). After such a \(T\) is identified, the above two methods can be used to compute the required phase shifts.

C. Other Properties

We close this section by discussing two properties of \(\phi\). First, the voltage angles are \(\theta = \mathcal{P}(B_{\perp}^{-1}(\beta_{\perp} - \phi_{\perp}))\) and the angle recovery condition (8) becomes

\[
B_{\perp} B_{\perp}^{-1}(\beta_{\perp} - \phi_{\perp}) = \beta_{\perp} - \phi_{\perp} \quad (\text{mod } 2\pi)
\]  
which can always be satisfied by appropriate (nonunique) choices of \(\phi\). A similar argument to the proof of Theorem 2(2) leads to the following interpretation of (25). For any link \((i,j) \in E\), (14) says that the phase angle difference from node \(i\) to node \(j\) is \(\beta_{ij} - \phi_{ij}\) and consists of the voltage angle difference \(\theta_i - \theta_j = \beta_{ij} - \phi_{ij}\) and the phase shifter angle \(\phi_{ij}\). Fix
any link \((i, j) \in E \setminus E_T\) not in tree \(T\). The left-hand side \(H_i^T H_j^T (\beta_T - \phi_T)_{ij}\) of (25) represents the sum of the voltage angle differences from node \(i\) to node \(j\) along the unique path in \(T\), not including the phase shifter angles along the path. This must be equal to the voltage angle difference \(\beta_{ij} - \phi_{ij}\) across (the non-tree) link \((i, j)\), not including the phase shifter angle across \((i, j)\). Therefore (25) has the same interpretation as before that the voltage angle differences sum to zero (mod \(2\pi\)) around any cycle, though, with phase shifters, the voltage angle differences are now \(\beta_{ij} - \phi_{ij}\) instead of \(\beta_{ij}\). This in particular leads to a relationship between any two solutions \((\theta_+, \phi_+)^\ast\) and \((\hat{\theta}, \hat{\phi})\) of (14).

In particular, let \((\hat{\theta}, \hat{\phi})\) be the solution in Theorem 1 where \(\phi_+ \in T^+, \theta_+ \in T^\perp\), and \((\hat{\theta}, \hat{\phi})\) any other solution. Then applying (25) to both \(\phi_+\) and \(\hat{\phi}_+\) leads to a relation between them on every basis cycle. Specifically, let \(i \rightarrow j\) be a link not in the spanning tree \(T\), let \(T'\) (0 \(\rightarrow\) k) be the unique path in \(T\) from node 0 to any node \(k\). Then for each link \(i \rightarrow j\) in \(E\) that is not in \(T\), we have (equalities to be interpreted as mod \(2\pi\))

\[
[\hat{\phi}_+]_{ij} = \hat{\phi}_{ij} - \sum_{(k,l) \in T(0 \rightarrow j)} \hat{\phi}_{kl} + \sum_{(k,l) \in T(0 \rightarrow i)} \hat{\phi}_{kl} - \beta_{ij} - \sum_{(k,l) \in T'(0 \rightarrow j)} \beta_{kl} + \sum_{(k,l) \in T'(0 \rightarrow i)} \beta_{kl},
\]

Second, Theorem 1 implies that given any relaxed solution \(\hat{\phi}\), there exists a \(\phi \in T^\perp\) such that its inverse projection \(x := h_0(\hat{\phi})\) is a branch flow solution, i.e., \((x, \phi)\) satisfies (9)–(11). We now give an alternative direct construction of such a solution \((x, \phi)\) from any given branch flow solution \(x\) and phase shifter setting \(\hat{\phi}\) that may have nonzero angles on some links in \(T\). It exhibits how the effect of phase shifters in a tree is equivalent to changes in voltage angles.

Fix any spanning tree \(T\). Given any \((x, \hat{\phi})\), partition \(\hat{\phi}^T - [\hat{\phi}_T \hat{\phi}_I]\) with respect to \(T\). Define \(\alpha \in (-\pi, \pi]^n\) by \(B_T \alpha = \hat{\phi}_T\) or \(\alpha := B_T^{-1} \hat{\phi}_T\). Then define the mapping \((x, \phi) = g(\hat{x}, \hat{\phi})\) by

\[
\phi_{ij} := \begin{cases} 0 & \text{if } (i, j) \in E_T \\ \hat{\phi}_{ij} - (\alpha_i - \alpha_j) & \text{if } (i, j) \in E \setminus E_T \end{cases}
\]

This mapping \((x, \phi) = g(\hat{x}, \hat{\phi})\) is a one-to-one correspondence between \((\hat{x}, \hat{\phi})\) and \((x, \phi)\) satisfying (10) and (11). Moreover, the effect of phase shifters in \(T\) is equivalent to adding \(\alpha_i\) to the phases of \(V_i\) and \(I_{ij}\).

**Theorem 3** Fix any tree \(T\). If \((x, \phi)\) is a solution of (9)–(11), so is \((x, \phi) = g(\hat{x}, \hat{\phi})\) defined in (26) and (27).

**Proof:** Since \(V_i = \hat{V}_i, I_{ij} = \hat{I}_{ij}\) and \(S = \hat{S}\) satisfies (10) and (11). (26) and (27) imply

\[
V_i - V_j e^{-i\phi_{ij}} - (\hat{V}_i - \hat{V}_j e^{-i(\alpha_\perp - \alpha_j)}) e^{i\alpha_i} = (\hat{V}_i - \hat{V}_j e^{-i\hat{\phi}_{ij}}) e^{i\alpha_i},
\]

and the second equality follows from \(B_T \alpha = \hat{\phi}_T\). For any link \((i, j) \in E \setminus E_T\) not in \(T\), (26) and (27) imply

\[
V_i - V_j e^{-i\phi_{ij}} - (\hat{V}_i - \hat{V}_j e^{-i\hat{\phi}_{ij}}) e^{i\alpha_i}.
\]

But \((\hat{V}_i - \hat{V}_j e^{-i\hat{\phi}_{ij}}) = \hat{I}_{ij}\) since \((\hat{x}, \hat{\phi})\) satisfies (9). Therefore \(V_i - V_j e^{-i\hat{\phi}_{ij}} = I_{ij}\), i.e., \((x, \phi)\) satisfies (9) on every link. 

**IV. OPF IN CONVEXIFIED NETWORK**

Theorem 1 suggests using phase shifters to convexify a mesh network so that any solution of OPF-ar can be mapped into an optimal solution of OPF of the convexified network. Convexification thus modifies an NP-hard problem into a simple problem without loss in optimality; moreover this requires an one-time deployment cost for subsequent operational simplicity, as we now show.

We will compare four OPF problems: the original OPF defined in [2]:

**OPF:**

\[
\min_{x, \alpha} \{f(\hat{h}(x), s) \mid x \in X, (S, v, s_0, s) \in S\}
\]

the relaxed OPF-ar defined in [2]:

**OPF-ar:**

\[
\min_{x, \alpha} \{f(\hat{h}(x), s) \mid x \in Y, (S, v, s_0, s) \in S\}
\]

the following problem where there is a phase shifter on every line \((\phi \in (\pi, \pi]^m)\):

**OPF-ps:**

\[
\min_{x, \alpha, \phi} \{f(\hat{h}(x), s) \mid x \in X, (S, v, s_0, s) \in S, \phi \in T^\perp\}
\]

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

**OPF-ps(T):**

\[
\min_{x, \alpha, \phi} \{f(\hat{h}(x), s) \mid x \in X_T, (S, v, s_0, s) \in S, \phi \in T^\perp\}
\]

Let the optimal values of OPF, OPF-ar, OPF-ps, and OPF-ps(T) be \(f_T, f_T, f_T, f_T\), respectively.

**Theorem 1** implies that \(X \subseteq Y = X_T\) for any spanning tree \(T\). Hence we have
**TABLE I**

<table>
<thead>
<tr>
<th>Test cases</th>
<th># links (m)</th>
<th>No PS</th>
<th>With phase shifters (PS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Min loss (OPF, MW)</td>
<td>Min loss (OPF-CR, MW)</td>
</tr>
<tr>
<td>IEEE 14-Bus</td>
<td>20</td>
<td>0.546</td>
<td>0.545</td>
</tr>
<tr>
<td>IEEE 30-Bus</td>
<td>41</td>
<td>1.372</td>
<td>1.299</td>
</tr>
<tr>
<td>IEEE 57-Bus</td>
<td>80</td>
<td>11.302</td>
<td>10.910</td>
</tr>
<tr>
<td>IEEE 118-Bus</td>
<td>186</td>
<td>9.232</td>
<td>8.728</td>
</tr>
<tr>
<td>IEEE 300-Bus</td>
<td>411</td>
<td>211.871</td>
<td>197.387</td>
</tr>
<tr>
<td>New England 39-Bus</td>
<td>446</td>
<td>29.915</td>
<td>28.901</td>
</tr>
<tr>
<td>Polish (case2383wp)</td>
<td>2106</td>
<td>433.019</td>
<td>385.894</td>
</tr>
<tr>
<td>Polish (case2737 SOP)</td>
<td>3,506</td>
<td>130.145</td>
<td>109.905</td>
</tr>
</tbody>
</table>

**TABLE II**

<table>
<thead>
<tr>
<th>Test cases</th>
<th>Max load. (OPF)</th>
<th>Max load. (OPF-CR)</th>
<th># required PS (m − n)</th>
<th># active PS</th>
<th>Min #PS (%)</th>
<th>[|\phi|^2] (%)</th>
<th>Simu. time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 14-Bus</td>
<td>195.0%</td>
<td>195.2%</td>
<td>7 (35%)</td>
<td>6 (30%)</td>
<td>[−0.51, 1.35]</td>
<td>[-0.28, 0.23]</td>
<td>1.92</td>
</tr>
<tr>
<td>IEEE 30-Bus</td>
<td>156.7%</td>
<td>158.7%</td>
<td>12 (29%)</td>
<td>9 (22%)</td>
<td>[−0.42, 12.4]</td>
<td>[-2.68, 1.86]</td>
<td>3.86</td>
</tr>
<tr>
<td>IEEE 57-Bus</td>
<td>108.2%</td>
<td>118.3%</td>
<td>24 (30%)</td>
<td>24 (30%)</td>
<td>[−13.1, 25.2]</td>
<td>[-4.12, 4.12]</td>
<td>7.13</td>
</tr>
<tr>
<td>IEEE 118-Bus</td>
<td>203.7%</td>
<td>204.9%</td>
<td>69 (37%)</td>
<td>64 (34%)</td>
<td>[−8.47, 17.6]</td>
<td>[-4.61, 5.36]</td>
<td>15.96</td>
</tr>
<tr>
<td>IEEE 300-Bus</td>
<td>106.8%</td>
<td>112.8%</td>
<td>112 (27%)</td>
<td>103 (25%)</td>
<td>[−15.0, 16.5]</td>
<td>[-4.28, 6.31]</td>
<td>34.6</td>
</tr>
<tr>
<td>New England 39-Bus</td>
<td>109.1%</td>
<td>117.0%</td>
<td>8 (17%)</td>
<td>5 (11%)</td>
<td>[−1.02, 1.28]</td>
<td>[-0.28, 0.18]</td>
<td>2.82</td>
</tr>
<tr>
<td>Polish (case2383wp)</td>
<td>101.4%</td>
<td>106.6%</td>
<td>514 (18%)</td>
<td>435 (15%)</td>
<td>[−19.6, 19.4]</td>
<td>[-4.06, 4.32]</td>
<td>434.5</td>
</tr>
<tr>
<td>Polish (case2737 SOP)</td>
<td>127.6%</td>
<td>132.5%</td>
<td>770 (22%)</td>
<td>420 (12%)</td>
<td>[−13.9, 17.1]</td>
<td>[-2.07, 3.62]</td>
<td>483.7</td>
</tr>
</tbody>
</table>

**Corollary 4:** For any spanning tree $T$, $f_* \geq f_{ar} = f_{ps} = f_T$, with equality if there is a solution $(\gamma_{ar}, s_{ar})$ of OPF-ar that satisfies (8).

Corollary 4 has several important implications:
1. Theorem 1 in [2] implies that we can solve OPF-ar efficiently through conic relaxation to obtain a relaxed solution $(\gamma_{ar}, s_{ar})$. An optimal solution of OPF may or may not be recoverable from it. If $\gamma_{ar}$ satisfies the angle recovery condition (8) with respect to $s_{ar}$, then Theorem 2 in [2] guarantees a unique $\theta_* \in (-\pi, \pi)^n$ such that the inverse projection $(h_{\theta_*}, \gamma_{ar}, s_{ar})$ is indeed optimal for OPF.
2. In this case, Corollary 4 implies that adding any phase shifters to the network cannot further reduce the cost since $f_* = f_{ar} = f_{ps}$.
3. If (8) is not satisfied, then $\gamma_{ar} \not\in h(X)$ and there is no inverse projection that can recover an optimal solution of OPF from $(\gamma_{ar}, s_{ar})$. In this case, $f_* \geq f_{ar}$. Theorem 1 implies that if we allow phase shifters, we can always attain $f_{ar} = f_{ps}$ with the relaxed solution $(\gamma_{ar}, s_{ar})$, with potentially strict improvement over the network without phase shifters (when $f_* > f_{ar}$).
4. Moreover, Corollary 4 implies that such benefit can be achieved with phase shifters only in branches outside an arbitrary spanning tree $T$.

**Remark 2:** The choice of the spanning tree $T$ does not affect the conclusion of the theorem. Different choices of $T$ correspond to different choices of $n$ linearly independent rows of $B$ and the resulting decomposition of $B$ and $\beta$ into $B_T$ and $\beta_T$. Therefore $T$ determines the phase angles $\theta_*$ and $\phi_*$ according to the formulae in the theorem. Since the objective $f(h(x), n)$ of OPF is independent of the phase angles $\theta_*$, for the same relaxed solution $\gamma$, OPF-ps achieves the same objective value regardless of the choice of $T$.

**V. SIMULATIONS**

For radial networks, results in Part I (Theorem 4) guarantees that both the angle relaxation and the conic relaxation are exact. For mesh networks, the angle relaxation may be inexact and phase shifters may be needed to implement a solution of the conic relaxation. We now explore through numerical experiments the following questions:

- How many phase shifters are typically needed to convexify a mesh network?
- What are typical phase shifter angles to implement an optimal solution for the convexified network?

**Test Cases:** We explore these questions using the IEEE benchmark systems with 14, 30, 57, 118, and 300 buses, as well as a 39-bus model of a New England power system and two models of a Polish power system with 2383 and 2737 buses. The data for all the test cases were extracted from the library of built-in models of the MATPOWER toolbox [7] in Matlab. The test cases involve constraints on bus voltages as well as limits on the real and reactive power generation at every generator bus. The New England and the Polish power systems also involve MVA limits on branch power flows. All these systems are mesh networks, but very sparse.

**Objectives:** We solve the test cases for two scenarios:
- Loss minimization. In this scenario, the objective is to minimize the total active power loss of the circuit given constant load values, which is equivalent to minimizing the
total active power generation. The results are shown in Table I.

• *Loadability maximization.* In this scenario, the objective is to determine the maximum possible load increase in the system while satisfying the generation, voltage and line constraints. We have assumed all real and reactive loads grow uniformly, i.e., by a constant multiplicative factor called the max loadability in Table II.

**Solution Methods:** We use the “SEDUMI” solver in Matlab [8]. We first solved the SOCP relaxation OPF-cr for a solution $\left(\hat{\mathbf{g}}, \hat{\mathbf{s}}\right)$ of OPF-ar. In all test cases, equality was attained at optimality for the second-order cone constraint, and hence OPF-cr was exact, as [2, Theorem 1] would suggest. Recall however that the load values were constants in all the test cases. Even though this violated our condition that there are no upper bounds on the loads OPF-cr turned out to be exact with respect to OPF-ar in all cases. This confirms that the no-upper-bound condition is sufficient but not necessary for the conic relaxation to be exact.

Using the solution $\left(\hat{\mathbf{g}}, \hat{\mathbf{s}}\right)$ of OPF-ar, we checked if the angle recovery condition (8) was satisfied. In all test cases, the angle recovery condition failed and hence no $\left(\mathbf{h}_{\beta}(\hat{\mathbf{g}}), \hat{\mathbf{s}}\right)$ was feasible for OPF without phase shifters. We computed the phase shifter angles $\phi$ using both methods explained in Section III-A and the corresponding unique $\left(\mathbf{h}_{\beta}(\hat{\mathbf{g}}), \hat{\mathbf{s}}\right)$ that was an optimal solution of OPF for the convexified network. For the first method that minimizes the number of required phase shifters, we have used a minimum spanning tree of the network where the weights on the lines are their reactance values. For the second method, we solve an approximation to the angle minimization that optimizes over $\beta$ for the fixed $k$ that shifts $\beta$ to $(-\pi, \pi]^m$.

In Tables I and II, we report the number $m - n$ of phase shifters potentially required, the number of active phase shifters (i.e., those with a phase angles greater than 0.1°), and the range of the phase angles at optimality using both methods. In Table II, we also report the simulation time on an Intel 1.8 GHz Core i5 CPU.

We report the optimal objective values of OPF with and without phase shifters in Tables I and II. The optimal values of OPF without phase shifters were obtained by implementing the SDP formulation and relaxation proposed in [9] for solving OPF. In all test cases, the solution matrix was of rank one and hence the SDP relaxation was exact. Therefore the values reported here are indeed optimal for OPF.

The SDP relaxation requires the addition of small resistances ($10^{-6}$ pu) to every link that has a zero resistance in the original model, as suggested in [10]. This addition is, on the other hand, not required for the SOCP relaxation: OPF-cr is tight with respect to OPF-ar with or without this addition. For comparison, we report the results where the same resistances are added for both the SDP and SOCP relaxations.

**Summary:** From Tables I and II:

1) Across all test cases, the convexified networks have higher performance (lower minimum loss and higher maximum loadability) than the original networks. More important than the modest performance improvement is design for simplicity: it guarantees an efficient solution for OPF.

2) The networks are (mesh but) very sparse, with the ratios $m/(n + 1)$ of the number of lines to the number of buses varying from 1.2 to 1.6 (Table I). The numbers $m - n$ of phase shifters potentially required on every link outside a spanning tree for convexification vary from 17% of the numbers $m$ of links to 37%.

3) The numbers of active phase shifters in the test cases vary from 7% of the numbers $m$ of links to 25% for loss minimization, and 11% to 34% for loadability maximization. The phase angles required at optimality is no more than 20° in magnitude with the minimum number of phase shifters. With the maximum number of phase shifters, the range of the phase angles is much smaller (less than 7°).

4) The simulation times range from a few seconds to minutes. This is much faster than SDP relaxation. Furthermore they appear linear in network size.

**VI. CONCLUSION**

We have presented a branch flow model and demonstrated how it can be used for the analysis and optimization of mesh as well as radial networks. Our results confirm that radial networks are computationally much simpler than mesh networks. For mesh networks, we have proposed a simple way to convexify them using phase shifters that will render them computationally as simple as radial networks for power flow solution and optimization. The addition of phase shifters thus convert a nonconvex problem into a different, simpler problem.

We have proposed a solution strategy for OPF that consists of two steps:

1) Compute a relaxed solution of OPF-ar by solving its conic relaxation OPF-cr.

2) Recover from a relaxed solution an optimal solution of the original OPF using an angle recovery algorithm.

We have proved that, for radial networks, both steps are always exact, provided there are no upper bounds on loads, so this strategy guarantees a globally optimal solution. For mesh networks the angle recovery condition may not hold but can be used to check if a given relaxed solution is globally optimal.

Since practical power networks are very sparse, the number of required phase shifters may be relatively small. Moreover, their placement depends only on network topology, but not on power flows, generations, loads, or operating constraints. Therefore an one-time deployment cost is required to achieve the subsequent simplicity in network and market operations.

**ACKNOWLEDGMENT**

The authors would like to thank S. Bose, K. M. Chandy, L. Gan, and B. Hassibi of Caltech, and C. Clarke, M. Montoya, and R. Sherick of the Southern California Edison.

**REFERENCES**


**Masoud Farivar** received the B.S. degree with a double major in electrical engineering and computer science from Sharif University of Technology, Tehran, Iran, in 2009. He is pursuing the Ph.D. degree in electrical engineering at California Institute of Technology, Pasadena, CA, USA.

He is a fellow of the Resnick Sustainability Institute at California Institute of Technology. Since June of 2010, he has also been working in the Advanced Technology division of Southern California Edison on a number of smart grid projects.

**Steven H. Low** (F’08) received the B.S. degree from Cornell University, Ithaca, NY, USA, and the Ph.D. from the University of California, Berkeley, CA, USA, both in electrical engineering.

He is a Professor of the Computing and Mathematical Sciences and Electrical Engineering Departments at California Institute of Technology, Pasadena, CA, USA, and an adjunct professor of both Swinburne University, Australia, and Shanghai Jiao Tong University, China. Before that, he was with AT&T Bell Laboratories, Murray Hill, NJ, USA, and the University of Melbourne, Australia.

Prof. Low was a co-recipient of IEEE best paper awards, the R&D 100 Award, and an Okawa Foundation Research Grant. He was a member of the Networking and Information Technology Technical Advisory Group for the US President’s Council of Advisors on Science and Technology (PCAST) in 2006. He is a Senior Editor of the *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS* (and the mentor for the annual JSAC series on Smart Grid Communications), a Senior Editor of the *IEEE TRANSACTIONS ON CONTROL OF NETWORKED SYSTEMS*, and a Steering Committee Member of the *IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING*. He also served on the editorial boards of *IEEE/ACM Transactions on Networking*, *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, *ACM Computing Surveys*, *Computer Networks Journal*, and *NOW Foundations and Trends in Networking*. 