Branch Flow Model: Relaxations and Convexification

Masoud Farivar Steven H. Low
Engineering and Applied Science
Caltech

***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) problems 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 and we characterize when the angle relaxation may fail. We propose a simple method to convexify a mesh network using phase shifters so that both relaxation steps are always exact and OPF for the convexified network can always be solved efficiently for a globally optimal solution. We prove that convexification requires phase shifters only outside a spanning tree of the network graph and their placement depends only on network topology, not on power flows, generation, loads, or operating constraints. Since power networks are sparse, the number of required phase shifters may be relatively small.**

A preliminary version was submitted for conference publication, March 2012.

CONTENTS

I Introduction 2
I-A Motivation 2
I-B Summary 3

II Branch flow model 4
II-A Branch flow equations 4
II-B Optimal power flow 4
II-C Notations and assumptions 5

III Relaxations and solution strategy 5
III-A Relaxed branch flow equations 5
III-B Two relaxations 7
III-C Solution strategy 7

IV Exact conic relaxation 8

V Angle relaxation 9
V-A Angle recovery condition 9
V-B Angle recovery algorithms 10
V-C Radial networks 11

VI Convexification of mesh network 11
VI-A Branch flow solutions 11
VI-B Optimal power flow 12

VII Simulations 13

VIII Conclusion 15

IX Appendix 15
IX-A OPF-ar has zero duality gap 15
IX-B $\hat{h}$ is injective on $\mathbb{X}$ 16

References 16

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. A key advantage is the simple linear relationship $I = YV$ between the nodal current injections I and the bus voltages V through the admittance matrix Y. 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, including optimal power flow, demand response, and Volt/VAR control.

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 Kirchoff'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 Carpentier's first formulation in 1962 [1]; surveys can be found in, e.g., [2]–[6]. 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. [7]–[10]. An important observation was made in [11], [12] 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 [11], [12] 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. A sufficient condition is derived in [13] under which the semidefinite relaxation is exact. Moreover this condition is shown to essentially hold in various IEEE test systems. Instead of solving the OPF problem directly, [13] proposes to solve its convex Lagrangian dual problem since the duality gap is zero when the relaxation is exact. This result is extended in [14] to include other variables and constraints and in [15] to exploit the sparsity of power networks and phase shifters for convexification of OPF. In [16], [17], it is proved that the sufficient condition of [13] always holds for a radial (tree) network, provided the bounds on the power flows satisfy a simple pattern. See also [18] for a generalization. These results confirm that radial networks are computationally much simpler. This is important as most distribution systems are radial.

While semidefinite relaxation for OPF has generated a lot of interest recently, its limitation has been studied in [19] using mesh networks with 3, 5, and 7 buses. They show that as a line-flow constraint is tightened, the sufficient condition in [13] fails to hold for these examples and the duality gap becomes nonzero. Moreover, the solutions produced by the semidefinite relaxation are physically meaningless in those cases. Indeed, examples of nonconvexity have long been discussed in the literature, e.g., [20]–[22]. Hence it is important to develop systematic methods for solving OPF involving mesh networks when convex relaxation fails.

We propose such a method here. While the papers above are all based on the bus injection model, ours uses a branch flow model. It is motivated by the model first proposed by Baran and Wu in [23], [24] for the optimal placement and sizing of switched capacitors in distribution circuits for Volt/VAR control. By recasting their model as a set of linear and quadratic equality constraints, [25] observes that relaxing the quadratic equality constraints to inequality constraints yields a second-order cone program (SOCP). It proves that the SOCP relaxation is exact and the duality gap is zero for radial networks, when there are no upper bounds on the loads. This result is extended here to mesh networks and convex, as opposed to linear, objective functions (Theorem 1). See also [26], [27] for various convex relaxations of approximations of the Baran-Wu model.

Other branch flow models have also been studied, e.g., in [28]–[30], all for radial networks. Indeed, [28] studies a similar model to that in [23], [24], using receiving-end branch powers as variables instead of sending-end branch powers as in [23], [24]. Both [29] and [30] 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 [29] develops a Newton-Raphson algorithm to solve the bus power injections, [30] proposes to solve for the branch flow powers through an SOCP relaxation for radial networks, though no proof of optimality is provided.

This set of papers [23], [24], [28]–[30] all exploit the fact that the power flow models 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 [31] points out for their branch flow model. In this paper we extend the Baran-Wu model from radial to mesh networks. We start with a full branch flow model and treat the extended Baran-Wu model as a projection of the full model that maps each bus voltage from a point in the complex plane to a circle with radius equal to the distance of the point from the origin. We clarify the mapping between the full branch flow model and its angle relaxation (extended Baran-Wu model) and use this characterization to develop a solution strategy for OPF for mesh as well as radial networks, as we now summarize.

B. Summary

We first formulate in Section II the OPF problem using full 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, as illustrated in Figure 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.
- *Conic relaxation*: relax OPF-ar to a cone program OPF-cr that is convex and hence can be solved efficiently.

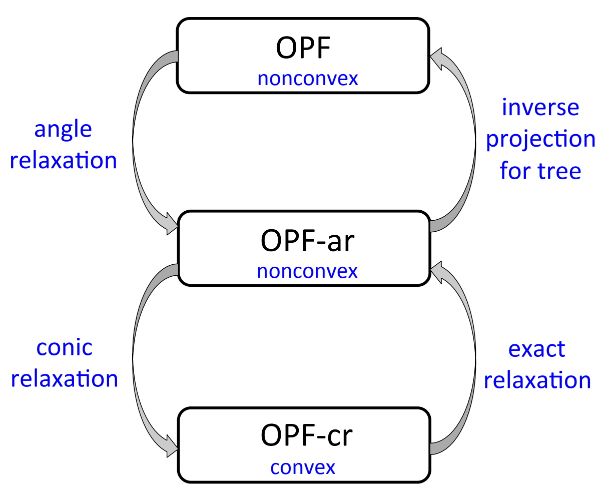

Fig. 1: Proposed solution strategy for solving OPF.

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., an optimal solution of OPF-cr is also optimal for OPF-ar. Given an optimal solution of OPF-ar, can we derive an optimal solution of the original OPF problem? The answer boils down to whether we can recover the voltage and current angles correctly from a relaxed branch flow solution, i.e., whether there exists an inverse of the projection mapping that maps a relaxed solution to a solution of the full branch flow model. In Section V we characterize the exact condition (the angle recovery condition) under which such an inverse projection exists, and present two angle recovery algorithms. It turns out that the angle recovery condition always holds for a radial network and hence the inverse projection always exists and is unique. For a mesh network, the angle recovery condition may not hold, and our characterization can be used to check if a full solution can be recovered from a given relaxed solution.

This suggests the following strategy for solving OPF:

1) Compute a relaxed solution of OPF-ar by solving its conic relaxation OPF-cr;
2) Recover from the relaxed solution an optimal solution of the original OPF using the angle recovery algorithms.

For a radial network, both the angle relaxation and the conic relaxation steps are exact, so this strategy will always yield an optimal solution of OPF. For a mesh network, the second step may fail.

To deal with this case, we propose in Section VI a simple method to convexify the network. 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 simple 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 optimal 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.

Since power networks in practice are very sparse, the number of lines not in a spanning tree can be relatively small, as demonstrated in simulations in Section VII 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 2,000 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.

II. Branch flow model

Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{C}$ denote the set of complex numbers. A variable without a subscript usually denotes a vector with appropriate components, e.g., $s := (s_i, i = 1, \ldots, n)$, $S := (S_{ij}, (i,j) \in E)$. For a complex scalar or vector a, a^* denotes its complex conjugate. For a vector $a = (a_1, \ldots, a_k)$, a_{-i} denotes $(a_1, \ldots, a_{i-1}, a_{i+1}, a_k)$. For a matrix A, A^t denotes its transpose and A^* its complex conjugate transposed.

A. Branch flow equations

Let $G = (N, E)$ 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, node 0 is the root of the tree and represents the substation bus; other nodes in N represent branch buses. 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. 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) if it points from node i to node j. We will use l and (i,j) interchangeably to refer to a link in E. For each link $(i,j) \in E$, we will call node i the *parent* of node j and j the *child* of i. Let $\pi(j) \subseteq N$ be the set of all parents of node j and $\delta(i) \subseteq N$ the set of all children of node i. Henceforth we will assume without loss of generality that G and T are directed graphs as described above.[1]

The basic variables of interest can be defined in terms of G. For each $(i,j) \in E$, let I_{ij} be the complex current from buses i to j and $S_{ij} = P_{ij} + \mathbf{i}Q_{ij}$ be the *sending-end* complex power from buses i to j. For each node $i \in N$, let V_i be the complex voltage on bus i. Let s_i be the net complex power, which is load minus generation on bus i. For power flow analysis, we assume s_i are given. For optimal power flow, VAR control, or demand response, s_i are control variables. We use s_i to denote both the complex number $p_i + \mathbf{i}q_i$ and the pair (p_i, q_i) depending on the context. Finally, let $z_{ij} = r_{ij} + \mathbf{i}x_{ij}$ be the complex impedance on the line connecting buses i and j.

[1]The orientation of G and T are different for different spanning trees T, but we often ignore this subtlety in this paper.

Then these quantities satisfy the law of conservation of power at each bus j:[2]

$$\sum_{i \in \pi(j)} \left(S_{ij} - z_{ij}|I_{ij}|^2 \right) - \sum_{k \in \delta(j)} S_{jk} \;=\; s_j, \quad \forall j \quad (1)$$

the Ohm's law over each link (i,j):

$$V_i - V_j \;=\; z_{ij} I_{ij}, \qquad \forall (i,j) \in E \qquad (2)$$

and the definition of branch power flow:

$$S_{ij} \;=\; V_i I_{ij}^*, \qquad \forall (i,j) \in E \qquad (3)$$

We will refer to (1)–(3) as the *branch flow equations/model*. As customary, we assume that the complex voltage V_0 on the slack/substation bus is given and the corresponding complex net load s_0 is a variable. Recall that the cardinality $|N| = n + 1$ and let $m := |E|$. The branch flow equations (1)–(3) specify $2m + n + 1$ nonlinear equations in $2m + n + 1$ complex variables $(S, I, V, s_0) := (S_{ij}, I_{ij}, (i,j) \in E, \; V_i, i = 1, \ldots, n, \; s_0)$, when other bus power injections $s := (s_i, i = 1, \ldots, n)$ are specified.

We will call a solution of (1)–(3) a *branch flow solution* with respect to a given s, and denote it by $x(s) := (S, I, V, s_0)$. It defines a steady state of a power system. Let $\mathbb{X}(s) \subseteq \mathbb{C}^{2m+n+1}$ be the set of all branch flow solutions with respect to a given s:

$$\mathbb{X}(s) := \{ x := (S, I, V, s_0) \,|\, x \text{ solves (1)–(3) given } s \} \qquad (4)$$

and let $\mathbb{X}$ be the set of all branch flow solutions:

$$\mathbb{X} \;:=\; \bigcup_{s \in \mathbb{C}^n} \mathbb{X}(s) \qquad (5)$$

For simplicity of exposition, we will often abuse notation and use $\mathbb{X}$ to denote either the set defined in (4) or that in (5), depending on the context. For instance, $\mathbb{X}$ 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, s_0), each $s_i = (p_i, q_i)$, $i = 1, \ldots, n$, is also an optimization variable. Let $p_i := p_i^c - p_i^g$ and $q_i := q_i^c - q_i^g$ where p_i^c and q_i^c are the real and reactive power consumption at node i, and p_i^g and q_i^g are

[2]Condition (1) follows from Tellegen's Theorem and (3). Tellegen's Theorem is a direct consequence of Kirchhoff's current and voltage laws.

the real and reactive power generation at node i. The active and reactive power generation and consumption at each node can be either positive or zero depending on whether the node represents a generator, a load, a shunt capacitor, or a storage device, etc. For instance, [23], [24] formulate a Volt/VAR control problem for a distribution circuit where q_i^g represent the placement and sizing of shunt capacitors; [25] uses it for inverter-based Volt/VAR control and formulates it as optimization over reactive power generations q_i^g from inverters that depend on the solar power output p_i^g. In addition to Kirchhoff's laws (1)–(3), we impose the following constraints on power generation: for $i = 0, 1, \ldots, n$,

$$\underline{p}_i^g \le p_i^g \le \overline{p}_i^g, \quad \underline{q}_i^g \le q_i^g \le \overline{q}_i^g \tag{6}$$

In particular, any of p_i^g, q_i^g can be a fixed constant by specifying that $\underline{p}_i^g = \overline{p}_i^g$ and/or $\underline{q}_i^g = \overline{q}_i^g$. For instance, in the inverter-based VAR control problem of [25], p_i^g are the fixed (solar) power outputs and the reactive power q_i^g are the control variables. For power consumption, we only require, for $i = 0, 1, \ldots, n$,

$$\underline{p}_i^c \le p_i^c, \quad \underline{q}_i^c \le q_i^c \tag{7}$$

i.e., there cannot be upper bounds on p_i^c, q_i^c for our proof below to work.[3] Finally the voltage magnitudes must be maintained in a tight range: for $i = 1, \ldots, n$,

$$\underline{v}_i \quad \le \quad |V_i|^2 \quad \le \quad \overline{v}_i \tag{8}$$

Suppose we aim to minimize real power losses $r_{ij}|I_{ij}|^2$ in lines, minimize real power generation costs $c_i p_i^g$, and maximize energy savings through conservation voltage reduction (CVR). As discussed in [25], maximizing CVR savings corresponds to minimizing a weighted sum of the $|V_i|^2$ values in the circuit where the weights are determined by the load models. Then the objective function takes the form

$$\sum_{(i,j)\in E} r_{ij}|I_{ij}|^2 + \sum_{i\in N} c_i p_i^g + \sum_{i\in N} \alpha_i |V_i|^2 \tag{9}$$

for some given constants $r_{ij}, c_i, \alpha_i \ge 0$. More generally, we consider any objective function that does not depend on the angles $\angle V_i, \angle I_{ij}$ of voltages and currents nor on consumptions p_i^c, q_i^c, described as follows.

To simplify notation, let $\ell_{ij} := |I_{ij}|^2$ and $v_i := |V_i|^2$. Let $s^g := (s_i^g,\, i = 1, \ldots, n) = (p_i^g, q_i^g,\, i = 1, \ldots, n)$ 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 $s^c - s^g$ or (s^c, s^g) depending on the context. Given a branch flow solution $x := x(s) := (S, I, V, s_0)$ with respect to a given s, let $\hat{y} := \hat{y}(s) := (S, \ell, v, s_0)$ denote the projection of x that have phase angles $\angle V_i, \angle I_{ij}$ 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 $f\left(\hat{h}(x), s^g\right)$. We assume $f(\hat{y}, s^g)$ is convex (condition A2); in addition, we assume f is strictly increasing in $\ell_{ij}, (i,j) \in E$ (condition A3). Finally, let

$$\mathbb{S} \quad := \quad \{\, (v, s_0, s) \,|\, (v, s_0, s) \text{ satisfies } (6) - (8) \,\}$$

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

The optimal power flow problem is
OPF:

$$\min_{x,s} \qquad f\left(\hat{h}(x), s^g\right) \tag{10}$$

$$\text{subject to} \qquad x \in \mathbb{X}, \quad (v, s_0, s) \in \mathbb{S} \tag{11}$$

where $\mathbb{X}$ is defined in (5). To avoid triviality, we assume the problem is feasible (condition A4).

The feasible set is specified by the nonlinear branch flow equations and hence OPF (10)–(11) 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 $f(\hat{y}, s^g)$ for optimal power flow is convex.
- A3 The cost function $f(\hat{y}, s^g)$ is strictly increasing in $\ell_{ij}, (i,j) \in E$.
- A4 The optimal power flow problem OPF (10)–(11) is feasible.

They are standard and realistic. For instance, the objective function in (9) satisfies conditions A2–A3. A3 holds if the cost function is increasing in power loss in the lines. It is not necessary and can be replaced.

III. Relaxations and solution strategy

We now describe our solution approach.

A. Relaxed branch flow equations

Substituting (3) into (2) yields $V_j = V_i - z_{ij} S_{ij}^* / V_i^*$. Taking the magnitude squared, we have $v_j = v_i +$

[3] This is equivalent to the "over-satisfaction of load" condition in [13], [16]. As we show below and in the simulations, this condition is sufficient but not necessary for the conic relaxation OPF-cr to be exact with respect to OPF-ar. See [32] for exact conic relaxation of OPF-cr under other conditions for radial networks.

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 $v_i := \|V_i\|^2$
$s_i = p_i + \mathbf{i}q_i$ $p_i = p_i^c - p_i^g$ $q_i = q_i^c - q_i^g$	net complex load power on bus i net real power equals load minus generation; net reactive power equals load minus generation
I_{ij}, ℓ_{ij}	complex current from buses i to j with $\ell_{ij} := \|I_{ij}\|^2$
$S_{ij} = P_{ij} + \mathbf{i}Q_{ij}$	complex power from buses i to j (sending-end)
$\mathbb{X}$ $\hat{\mathbb{Y}}$ $\overline{\mathbb{Y}}$ $\overline{X}$, $\overline{X}_T$	set of all branch flow solutions that satisfy (1)–(3) either for some s, or for a given s (sometimes denoted more accurately by $\mathbb{X}(s)$); set of all relaxed branch flow solutions that satisfy (12)–(15) either for a given s or for some s; convex hull of $\hat{\mathbb{Y}}$, i.e., solutions of (12)–(14) and (23); set of branch flow solutions that satisfy (1), (41), (3), for some phase shifter angles ϕ and for some $\phi \in T^{\perp}$;
$x = (S, I, V, s_0) \in \mathbb{X}$ $\hat{y} = (S, \ell, v, s_0) \in \hat{\mathbb{Y}}$ $\hat{y} = \hat{h}(x)$; $x = h_\theta(\hat{y})$	vector x of power flow variables and its projection $\hat{y}$; projection mapping $\hat{y}$ and an inverse h_θ
$z_{ij} = r_{ij} + \mathbf{i}x_{ij}$	impedance on line connecting buses i and j
f	objective function of OPF, of the form $f\left(\hat{h}(x), s^g\right)$

$|z_{ij}|^2\ell_{ij} - (z_{ij}S_{ij}^* + z_{ij}^*S_{ij})$. Using (1) and (3) and in terms of real variables, we therefore have

$$p_j = \sum_{i\in\pi(j)} (P_{ij} - r_{ij}\ell_{ij}) - \sum_{k\in\delta(j)} P_{jk}, \quad \forall j \quad (12)$$

$$q_j = \sum_{i\in\pi(j)} (Q_{ij} - x_{ij}\ell_{ij}) - \sum_{k\in\delta(j)} Q_{jk}, \quad \forall j \quad (13)$$

$$v_j = v_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)\ell_{ij} \quad \forall (i,j) \in E \quad (14)$$

$$\ell_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{v_i}, \quad \forall (i,j) \in E \quad (15)$$

We will refer to (12)–(15) as the *relaxed (branch flow) equations/model* and a solution a *relaxed (branch flow) solution*. These equations were first proposed in [23], [24] to model radial distribution circuits. They define a system of equations in the variables $(P, Q, \ell, v, p_0, q_0) := (P_{ij}, Q_{ij}, \ell_{ij}, (i,j) \in E,\ v_i, i = 1, \ldots, n,\ p_0, q_0)$, which are a subset of the original (complex) variables (S, I, V, s_0), without the angles $\angle V_i, \angle I_{ij}$. Since we assume the original branch flow model has a solution, the relaxed model must also have a solution.

In contrast to the original branch flow equations (1)–(3), the relaxed equations (12)–(15) specifies $2(m+n+1)$ equations in $3m+n+2$ real variables $(P, Q, \ell, v, p_0, q_0)$, for a given s. For a radial network, i.e., G is a tree, $m = |E| = |N| - 1 = n$. Hence the relaxed system (12)–(15) specifies $4n+2$ equations in $4n+2$ real variables. It is shown in [33] that there are generally multiple solutions, but for practical networks where $V_0 \simeq 1$ and r_{ij}, x_{ij} are small p.u., the solution of (12)–(15) is unique. Exploiting structural properties of the Jacobian matrix, efficient algorithms have also been proposed in [34] 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 (12)–(15), and therefore, if a solution exists, it 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 $(P, Q, \ell, v, p_0, q_0)$ 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 (12)–(15) may not correspond to any solution of (1)–(3). The key is whether we can recover the angles $\angle V_i, \angle I_{ij}$ correctly from a relaxed solution. It is then remarkable that, when G is a tree, indeed the solutions of (12)–(15) coincide with those of (1)–(3), as explained in Section V.

To understand the relationship between the branch flow model and the relaxed model and formulate our relaxations precisely, we start by defining some notations. Fix an s. Given a vector $(S, I, V, s_0) \in \mathbb{C}^{2m+n+1}$, define its projection $\hat{h} : \mathbb{C}^{2m+n+1} \to \mathbb{R}^{3m+n+2}$ by $\hat{h}(S, I, V, s_0) = (P, Q, \ell, v, p_0, q_0)$ where

$$P_{ij} = \text{Re } S_{ij}, \qquad Q_{ij} = \text{Im } S_{ij}, \quad \ell_{ij} = |I_{ij}|^2 \quad (16)$$

$$p_i = \text{Re } s_i, \qquad q_i = \text{Im } s_i, \qquad v_i = |V_i|^2 \quad (17)$$

We often abuse notation and write (S, ℓ, v, s_0) as a shorthand for $(P, Q, \ell, v, p_0, q_0)$. Let $\mathbb{Y} \subseteq \mathbb{C}^{2m+n+1}$ denote the set of all $y := (S, I, V, s_0)$ whose projections are the relaxed branch flow solutions: [4]

$$\mathbb{Y} := \left\{ y := (S, I, V, s_0) | \hat{h}(y) \text{ solves (12)–(15)} \right\} \quad (18)$$

[4]As mentioned earlier, the set defined in (18) is strictly speaking $\mathbb{Y}(s)$ with respect to a fixed s. To simplify exposition, we abuse notation and use $\mathbb{Y}$ to denote both $\mathbb{Y}(s)$ and $\bigcup_{s\in\mathbb{C}^n} \mathbb{Y}(s)$, depending on the context. The same applies to $\hat{\mathbb{Y}}$ and $\overline{\mathbb{Y}}$ etc.

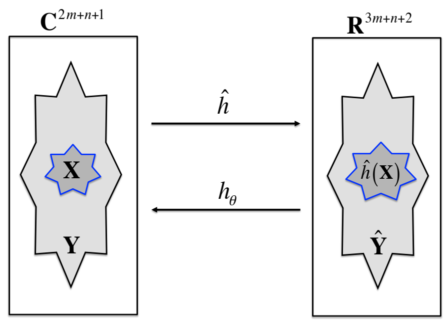

Fig. 2: $\mathbb{X}$ is the set of branch flow solutions and $\hat{\mathbb{Y}} = \hat{h}(\mathbb{Y})$ is the set of relaxed solutions.

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

$$\hat{\mathbb{Y}} \;\; := \;\; \{ \; \hat{y} := (S, \ell, v, s_0) \,|\, \hat{y} \text{ solves (12)–(15)} \; \}$$

Clearly

$$\mathbb{X} \subseteq \mathbb{Y} \quad \text{and} \quad \hat{h}(\mathbb{X}) \subseteq \hat{h}(\mathbb{Y}) = \hat{\mathbb{Y}}$$

Their relationship is illustrated in Figure 2.

B. Two relaxations

Consider the OPF with angles relaxed:
OPF-ar:

$$\min_{x,s} \quad f\left(\hat{h}(x), s^g\right) \tag{19}$$

$$\text{subject to} \quad x \in \mathbb{Y}, \quad (v, s_0, s) \in \mathbb{S} \tag{20}$$

Clearly, this problem provides a lower bound to the original OPF problem since $\mathbb{Y} \supseteq \mathbb{X}$. Since neither $\hat{h}(x)$ nor the constraints in $\mathbb{Y}$ involves angles $\angle V_i, \angle I_{ij}$, this problem is equivalent to the following
OPF-ar:

$$\min_{\hat{y},s} \quad f\left(\hat{y}, s^g\right) \tag{21}$$

$$\text{subject to} \quad \hat{y} \in \hat{\mathbb{Y}}, \quad (v, s_0, s) \in \mathbb{S} \tag{22}$$

The feasible set of OPF-ar is specified by a system of linear-quadratic equations. Hence OPF-ar is in general still nonconvex and hard to solve directly. The key to our solution is the observation that the only source of nonconvexity is the quadratic equalities in (15). Relax them to inequalities:

$$\ell_{ij} \;\; \geq \;\; \frac{P_{ij}^2 + Q_{ij}^2}{v_i}, \qquad (i,j) \in E \tag{23}$$

and define the set $\overline{\mathbb{Y}} \subseteq \mathbb{R}^{2m+n+1}$ as

$$\overline{\mathbb{Y}} := \{\hat{y} := (S, \ell, v, s_0) \,|\, \hat{y} \text{ solves (12)–(14) and (23)}\}$$

Consider the following relaxation of OPF-ar:
OPF-cr:

$$\min_{\hat{y},s} \quad f\left(\hat{y}, s^g\right) \tag{24}$$

$$\text{subject to} \quad \hat{y} \in \overline{\mathbb{Y}}, \quad (v, s_0, s) \in \mathbb{S} \tag{25}$$

Clearly OPF-cr provides a lower bound to OPF-ar since $\overline{\mathbb{Y}} \supseteq \hat{\mathbb{Y}}$.

C. Solution strategy

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

1) OFP-cr is convex. Moreover the conic relaxation is always exact so that an optimal solution $(\hat{y}_{cr}, s_{cr})$ of OPF-cr is also optimal for OPF-ar (Section IV, Theorem 1).
2) Given a solution $(\hat{y}_{ar}, s_{ar})$ of OPF-ar, if the power 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 (10)–(11) through an inverse projection (Section V, Theorems 2 and 4).
3) For a mesh network, angle relaxation may not be exact in the sense that no inverse projection exists to map $(\hat{y}_{ar}, s_{ar})$ to a feasible solution of OPF. In that case, however, the network can be convexified so that $(\hat{y}_{ar}, s_{ar})$ can indeed 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 (Section VI, Theorem 6 and Corollary 7).

These results suggest an algorithm for solving OPF (10)–(11) as summarized in Figure 3.

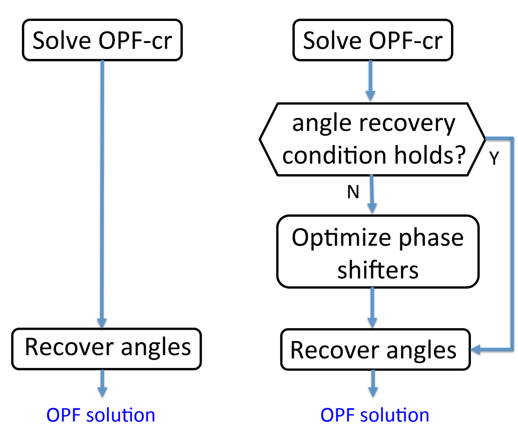

Fig. 3: Proposed algorithm for solving OPF (10)–(11). Left: radial networks. Right: convexified mesh networks. The details are explained in Sections IV–VI.

IV. Exact conic relaxation

Our first key result says that OPF-cr is exact.

Theorem 1: OPF-cr is convex. Moreover, it is exact, i.e., an optimal solution of OPF-cr is also optimal for OPF-ar.

Proof: The feasible set is convex since the nonlinear inequalities in $\overline{\mathbb{Y}}$ can be written as the following second order cone constraint:

$$\left\| \begin{matrix} 2P_{ij} \\ 2Q_{ij} \\ \ell_{ij} - v_i \end{matrix} \right\|_2 \leq \ell_{ij} + v_i$$

Since the objective function is convex, OPF-cr is a conic optimization.[5] To prove that the relaxation is exact, it suffices to show that any optimal solution of OPF-cr attains equality in (23).

Assume for the sake of contradiction that $(\hat{y}_*, s_*) := (S_*, \ell_*, v_*, s^g_{*0}, s^c_{*0}, s^g_*, s^c_*)$ is optimal for OPF-cr, but a link $(i,j) \in E$ has strict inequality, i.e., $[v_*]_i [\ell_*]_{ij} > {[P_*]_{ij}}^2 + {[Q_*]_{ij}}^2$. For some $\varepsilon > 0$ to be determined below, consider another point $(\tilde{y}, \tilde{s}) = (\tilde{S}, \tilde{\ell}, \tilde{v}, \tilde{s}^g_0, \tilde{s}^c_0, \tilde{s}^g, \tilde{s}^c)$ defined by:

$$\begin{aligned} \tilde{v} \;=\; v_*, &\qquad \tilde{s}^g \;=\; s^g_* \\ \tilde{\ell}_{ij} = [\ell_*]_{ij} - \varepsilon, &\qquad \tilde{\ell}_{-ij} = [\ell_*]_{-ij} \\ \tilde{S}_{ij} = [S_*]_{ij} - z_{ij}\varepsilon/2, &\qquad \tilde{S}_{-ij} = [S_*]_{-ij} \\ \tilde{s}^c_i = [s^c_*]_i + z_{ij}\varepsilon/2, &\qquad \tilde{s}^c_j = [s^c_*]_j + z_{ij}\varepsilon/2 \\ \tilde{s}^c_{-(i,j)} = [s^c_*]_{-(i,j)} & \end{aligned}$$

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

It suffices then to check that there exists an $\varepsilon > 0$ such that $(\tilde{y}, \tilde{s})$ satisfies (6)–(8), (12)–(14) and (23), and hence is indeed a feasible point. Since $(\hat{y}_*, s_*)$ is feasible, (6)–(8) hold for $(\tilde{y}, \tilde{s})$ too. Similarly, $(\tilde{y}, \tilde{s})$ satisfies (12)–(13) at all nodes $k \neq i, j$ and (14), (23) across all links $(k,l) \neq (i,j)$. We now show that $(\tilde{y}, \tilde{s})$ satisfies (12)–(13) also at nodes i, j, and (14), (23) across link (i,j):

- Proving (12)–(13) is equivalent to proving (1). At node i, we have

$$\begin{aligned} \tilde{s}_i &= \tilde{s}^c_i - \tilde{s}^g_i \\ &= [s^c_*]_i + z_{ij}\varepsilon/2 - [s^g_*]_i \\ &= \sum_{k \in \pi(i)} ([S_*]_{ki} - z_{ki}[\ell_*]_{ki}) - \sum_{j' \in \delta(i), j' \neq j} [S_*]_{ij'} \\ &\qquad -[S_*]_{ij} + z_{ij}\varepsilon/2 \\ &= \sum_{k \in \pi(i)} \left(\tilde{S}_{ki} - z_{ki}\tilde{\ell}_{ki}\right) - \sum_{j' \in \delta(i), j' \neq j} \tilde{S}_{ij'} \\ &\qquad - \left(\tilde{S}_{ij} + z_{ij}\varepsilon/2\right) + z_{ij}\varepsilon/2 \\ &= \sum_{k \in \pi(i)} \left(\tilde{S}_{ki} - z_{ki}\tilde{\ell}_{ki}\right) - \sum_{j' \in \delta(i)} \tilde{S}_{ij'} \end{aligned}$$

At node j, we have

$$\begin{aligned} \tilde{s}_j &= \tilde{s}^c_j - \tilde{s}^g_j \\ &= [s^c_*]_j + z_{ij}\varepsilon/2 - [s^g_*]_j \\ &= \sum_{i' \in \pi(j), i' \neq i} ([S_*]_{i'j} - z_{i'j}[\ell_*]_{i'j}) + [S_*]_{ij} \\ &\qquad -z_{ij}[\ell_*]_{ij} - \sum_{k \in \delta(j)} [S_*]_{jk} + z_{ij}\varepsilon/2 \\ &= \sum_{i' \in \pi(j), i' \neq i} \left(\tilde{S}_{i'j} - z_{i'j}\tilde{\ell}_{i'j}\right) + \tilde{S}_{ij} + z_{ij}\varepsilon/2 \\ &\qquad -z_{ij}(\tilde{\ell}_{ij} + \varepsilon) - \sum_{k \in \delta(j)} \tilde{S}_{jk} + z_{ij}\varepsilon/2 \\ &= \sum_{i' \in \pi(j)} \left(\tilde{S}_{i'j} - z_{i'j}\tilde{\ell}_{i'j}\right) - \sum_{k \in \delta(j)} \tilde{S}_{jk} \end{aligned}$$

Hence (12)–(13) hold at nodes i, j.

- For (14) across link (i,j), we have

$$\begin{aligned} \tilde{v}_j &= [v_*]_j \\ &= [v_*]_i - 2(r_{ij}[P_*]_{ij} + x_{ij}[Q_*]_{ij}) \\ &\qquad +(r_{ij}^2 + x_{ij}^2)[\ell_*]_{ij} \\ &= \tilde{v}_i - 2(r_{ij}\tilde{P}_{ij} + x_{ij}\tilde{Q}_{ij}) + (r_{ij}^2 + x_{ij}^2)\tilde{\ell}_{ij} \end{aligned}$$

- For (23) across link (i,j), we have

$$\begin{aligned} &\tilde{v}_i\tilde{\ell}_{ij} - \tilde{P}_{ij}^2 - \tilde{Q}_{ij}^2 \\ =\; & [v_*]_i \left([\ell_*]_{ij} - \varepsilon\right) - \left([P_*]_{ij} - r_{ij}\varepsilon/2\right)^2 \\ &\qquad - \left([Q_*]_{ij} - x_{ij}\varepsilon/2\right)^2 \\ =\; & \left([v_*]_i[\ell_*]_{ij} - [P_*]_{ij}^2 - [Q_*]_{ij}^2\right) \\ &\quad -\varepsilon\left([v_*]_i - r_{ij}[P_*]_{ij} - x_{ij}[Q_*]_{ij}\right. \\ &\qquad \left. + \;\varepsilon(r_{ij}^2 + x_{ij}^2)/4\right) \end{aligned}$$

Since

$$[v_*]_i[\ell_*]_{ij} - [P_*]_{ij}^2 - [Q_*]_{ij}^2 \;\; > \;\; 0$$

[5] If the objective function is linear, such as (9), then OPF-cr is a second order cone program. This is the case proved in [25] for radial networks. This result is extended here to mesh networks and convex objective functions.

we can choose an $\varepsilon > 0$ sufficiently small such that $\tilde{\ell}_{ij} \geq (\tilde{P}_{ij}^2 + \tilde{Q}_{ij}^2)/\tilde{v}_i$.

This completes the proof. ■

Remark 1: A3 is not necessary; e.g. [32] replaces assumption A3 with a variety of other conditions and prove that OPF-cr is always an exact relaxation of OPF-ar, as long as G is a tree. Assumption A3 is used in the proof here to contradict the optimality of $(\hat{y}_*, s_*^g)$. Instead of A3, if $f(\hat{y}, s^g)$ is nondecreasing in ℓ, the same argument shows that, given an optimal $(\hat{y}_*, s_*^g)$ with a strict inequality $[v_*]_i[\ell_*]_{ij} > [P_*]_{ij}{}^2 + [Q_*]_{ij}{}^2$, one can choose $\varepsilon > 0$ to obtain another optimal point $(\tilde{y}, \tilde{s}^g)$ that attains equality and has a cost $f(\tilde{y}, \tilde{s}^g) \leq f(\hat{y}_*, s_*^g)$. This implies that, in the absence of A3, there is always an optimal solution of OPF-cr that is also optimal for OPF-ar, even though it does not exclude the possibility that OPF-cr may have other optimal points with strict inequality that are infeasible for OPF-ar.

V. Angle relaxation

Theorem 1 justifies solving the convex problem OPF-cr for an optimal solution of OPF-ar. Given a solution $(\hat{y}, s)$ of OPF-ar, when and how can we recover a solution (x, s) of the original OPF (10)–(11)? The issue boils down to whether we can recover a solution x to the branch flow equations (1)–(3) from $\hat{y}$, given any nodal power injections s.

Hence, for the rest of this section, we fix an s and assume $\mathbb{X}(s)$ is nonempty, i.e., a branch flow solution $x = x(s)$ exists. We abuse notation in this section and write $x, \hat{y}, \theta, \mathbb{X}, \mathbb{Y}, \hat{\mathbb{Y}}$ instead of $x(s), \hat{y}(s), \theta(s), \mathbb{X}(s), \mathbb{Y}(s), \hat{\mathbb{Y}}(s)$ respectively.

A. Angle recovery condition

Fix a relaxed solution $\hat{y} := (S, \ell, v, s_0) \in \hat{\mathbb{Y}}$. Define the $(n+1) \times m$ incidence matrix C of G by

$$C_{il} = \begin{cases} 1 & \text{if link } l \text{ leaves node } i \\ -1 & \text{if link } l \text{ enters node } i \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

The first row of C corresponds to node 0, the reference bus with a given $V_0 = |V_0|e^{\mathbf{i}\theta_0}$. In this section 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 $l \in E, i = 1, \ldots, n$,

$$B_{li} = \begin{cases} 1 & \text{if link } l \text{ leaves node } i \\ -1 & \text{if link } l \text{ enters node } i \ , \\ 0 & \text{otherwise} \end{cases}$$

Since G is connected, $m \geq n$ and $\text{rank}(B) = n$ [35]. Fix any spanning tree $T = (N, E_T)$ of G. We can assume without loss of generality (possibly after re-labeling some of the links) that E_T consists of links $l = 1, \ldots, n$. Then B can be partitioned into

$$B = \begin{bmatrix} B_T \\ B_{T^\perp} \end{bmatrix} \tag{27}$$

where the $n \times n$ submatrix B_T corresponds to links in T and the $(m-n) \times n$ submatrix $B_{T^\perp}$ corresponds to links in $T^\perp := G \setminus T$.

Let $\beta := \beta(\hat{y}) \in [-\pi, \pi]^m$ be defined in terms of the given $\hat{y}$ by

$$\beta_{ij} := \angle\left(v_i - z_{ij}^* S_{ij}\right), \qquad (i,j) \in E \tag{28}$$

Write β as

$$\beta = \begin{bmatrix} \beta_T \\ \beta_{T^\perp} \end{bmatrix} \tag{29}$$

where β_T is $n \times 1$ and $\beta_{T^\perp}$ is $(m-n) \times 1$.

Recall the projection mapping $\hat{h} : \mathbb{C}^{2m+n+1} \to \mathbb{R}^{3m+n+2}$ defined in (16)–(17). For each $\theta := (\theta_i, i = 1, \ldots, n) \in [-\pi, \pi]^n$, define the inverse projection $h_\theta : \mathbb{R}^{3m+n+2} \to \mathbb{C}^{2m+n+1}$ by $h_\theta(P, Q, \ell, v, p_0, q_0) = (S, I, V, s_0)$ where

$$S_{ij} := P_{ij} + \mathbf{i}Q_{ij} \tag{30}$$

$$I_{ij} := \sqrt{\ell_{ij}}\ e^{\mathbf{i}(\theta_i - \angle S_{ij})} \tag{31}$$

$$V_i := \sqrt{v_i}\ e^{\mathbf{i}\theta_i} \tag{32}$$

$$s_0 := p_0 + \mathbf{i}q_0 \tag{33}$$

These mappings are illustrated in Figure 2.

By definition of $\hat{h}(\mathbb{X})$ and $\hat{\mathbb{Y}}$, a branch flow solution in $\mathbb{X}$ can be recovered if a given relaxed solution $\hat{y}$ is in $\hat{h}(\mathbb{X})$ and cannot be recovered if $\hat{y}$ is in $\hat{\mathbb{Y}} \setminus \hat{h}(\mathbb{X})$. In other words, $\hat{h}(\mathbb{X})$ consists of exactly those points $\hat{y} \in \hat{\mathbb{Y}}$ for which there exist θ such that their inverse projections $h_\theta(\hat{h})$ are in $\mathbb{X}$. 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 $\angle V_i, \angle I_{ij}$.

Theorem 2: Let T be *any* spanning tree of G. Consider a relaxed solution $\hat{y} \in \hat{\mathbb{Y}}$ and the corresponding β defined by (28)–(29) in terms of $\hat{y}$.

1) There exists a unique $\theta_* \in [-\pi, \pi]^n$ such that $h_{\theta_*}(\hat{y})$ is a branch flow solution in $\mathbb{X}$ if and only if

$$B_{T^\perp} B_T^{-1} \beta_T = \beta_{T^\perp} \tag{34}$$

2) If (34) holds, then $\theta_* = B_T^{-1}\beta_T$.

Remark 2: Given a relaxed solution $\hat{y} := (S, \ell, v, s_0)$, 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 (34) depends on two factors: the network topology through the reduced incidence matrix B, and the relaxed solution $\hat{y}$ through β. The choice of spanning tree T corresponds to choosing n linearly independent rows of B to form B_T and does not affect the conclusion of the theorem.

The proof of Theorem 2 relies on the following important lemma that describes when any arbitrary inverse projection $h_\theta(\hat{y})$ is a branch flow solution in $\mathbb{X}$.

Lemma 3: Given (S, ℓ, v, s_0) in $\hat{\mathbb{Y}}$ and $\theta \in [-\pi, \pi]^n$, the vector $h_\theta(S, \ell, v, s_0)$ defined by (30)–(33) is in $\mathbb{X}$ if and only if θ satisfies

$$\theta_i - \theta_j \;=\; \angle\left(v_i - z_{ij}^* S_{ij}\right), \qquad (i,j) \in E \quad (35)$$

Moreover such a θ, if it exists, is unique.

Proof: It is necessary since we have from (2) and (3)

$$V_i V_j^* \;=\; |V_i|^2 - z_{ij}^* S_{ij}$$

giving $\theta_i - \theta_j = \angle(v_i - z_{ij}^* S_{ij})$.

For sufficiency, we need to show that (12)–(15) together with (30)–(33) and (35) implies (1)–(3). Now (12) and (13) are equivalent to (1). Moreover (15) and (30)–(32) imply (3). To prove (2), we can substitute (3) into (35) to get

$$\theta_i - \theta_j \;=\; \angle\left(v_i - z_{ij}^* V_i I_{ij}^*\right) \;=\; \angle\, V_i\left(V_i^* - z_{ij}^* I_{ij}^*\right)$$

Hence

$$\angle V_j \;=\; \theta_j \;=\; \angle\left(V_i - z_{ij} I_{ij}\right) \quad (36)$$

From (14) and (3), we have

$$\begin{aligned} |V_j|^2 &= |V_i|^2 + |z_{ij}|^2 |I_{ij}|^2 - (z_{ij} S_{ij}^* + z_{ij}^* S_{ij}) \\ &= |V_i|^2 + |z_{ij}|^2 |I_{ij}|^2 - (z_{ij} V_i^* I_{ij} + z_{ij}^* V_i I_{ij}^*) \\ &= |V_i - z_{ij} I_{ij}|^2 \end{aligned}$$

This together with (36) implies $V_j = V_i - z_{ij} I_{ij}$ which is (2).

The condition (35) implies that a branch flow solution can be recovered from a relaxed solution if and only if there exist θ that solves

$$B\theta \;=\; \beta \quad (37)$$

Since G is connected, $m \geq n$ and $\text{rank}(A) = n$ and hence (37) has at most one solution. ■

Proof of Theorem 2: Since the network is connected, $m \geq n$ and $\text{rank}(B) = n$. Hence 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 [36, Chapter 5]. Assume without loss of generality that the first n rows is such a basis so that B and β are partitioned as in (27) and (29) respectively. Then Lemma 3 implies that there exists a unique $\theta_* \in [-\pi, \pi]^n$ such that $h_{\theta_*}(\hat{y})$ is a branch flow solution in $\mathbb{X}$ if and only if θ_* solves (37), i.e.,

$$\begin{bmatrix} B_T \\ B_{T^\perp} \end{bmatrix} \theta \;=\; \begin{bmatrix} \beta_T \\ \beta_{T^\perp} \end{bmatrix} \quad (38)$$

Since T is a spanning tree, the $n \times n$ submatrix B_T has a full rank and hence is invertible. $B_T \theta_* = \beta_T$ then implies $\theta_* = B_T^{-1} \beta_T$. Moreover (38) has a unique solution if and only if $B_{T^\perp} \theta_* = B_{T^\perp} B_T^{-1} \beta_T = \beta_{T^\perp}$. ■

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 $\hat{y} \in \hat{\mathbb{Y}}$,

1) Choose *any* n basis rows of B and form B_T and $B_{T^\perp}$.
2) Compute β from $\hat{y}$ and check if $B_{T^\perp} B_T^{-1} \beta_T = \beta_{T^\perp}$.
3) If not, then $\hat{y}$ is spurious, i.e., $\hat{y} \notin \hat{h}(\mathbb{X})$; stop.
4) Otherwise, compute $\theta_* = B_T^{-1} \beta_T$.
5) Compute $h_{\theta_*}(\hat{y}) \in \mathbb{X}$ through (30)–(33).

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

We present an alternative iterative procedure to compute the angles $\angle I_{ij}$ and $\angle V_i = \theta_i$, and hence a branch flow solution. This procedure is more amenable to a distributed implementation. Starting from the root node 0, for all children $j \in \delta(0)$, use (3) to determine

$$\angle I_{0j} \;:=\; \angle V_0 - \angle S_{0j}$$

and use (35) to determine

$$\angle V_j \;:=\; \angle V_0 - \angle(v_0 - z_{0j}^* S_{0j})$$

Repeat this process, starting from V_j, $j \in \delta(0)$, to determine the angles $\angle V_k, \angle I_{jk}$ for all children $k \in \delta(j)$, and so on, until all nodes in the radial network are covered. In summary:

Algorithm 2: distributed angle recovery. Given a solution $(S, \ell, v, s_0) \in \hat{\mathbb{Y}}$,

1) Choose *any* spanning tree T of G rooted at node 0.
2) As index j ranges over the tree T, starting from the root $j = 0$ in the order of breadth-first search, for all descendents $k \in \delta(j)$, set

$$\begin{aligned} \angle I_{jk} &:= \angle V_j - \angle S_{jk} \quad (39) \\ \angle V_k &:= \angle V_j - \angle(v_j - z_{jk}^* S_{jk}) \quad (40) \end{aligned}$$

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 $\angle V_k$ has already been computed in Step 2.
   a) Compute current angle $\angle I_{jk}$ using (39).
   b) Compute a new voltage angle θ_k^j using the new parent j and (40). If $\theta_k^j \neq \angle V_k$, then angle recovery has failed and (S, ℓ, v, s_0) is spurious.

If the angle recovery procedure succeeds in Step 3, then (S, ℓ, v, s_0) together with these angles $\angle V_k, \angle I_{jk}$ are indeed a branch flow solution. Otherwise, the angles $\angle V_k$ determined in Step 1 do not satisfy the Kirchhoff voltage law $\sum_i V_i = 0$ around the loop that involves the link (j,k) identified in Step 3(b). This violates the condition $B_{T^\perp} B_T^{-1} \beta_T = \beta_{T^\perp}$ in Theorem 2.

C. Radial networks

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

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

1) $\hat{h}(\mathbb{X}) = \hat{\mathbb{Y}}$.
2) given any $\hat{y}$, $\theta_* := B^{-1}\beta$ always exists and is the unique phase angle vector such that $h_{\theta_*}(\hat{y}) \in \mathbb{X}$.

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

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 an optimal solution $(\hat{y}_*, s_*)$ of OPF-cr, there exists a unique θ_* such that $(h_{\theta_*}(\hat{y}_*), s_*)$ is an optimal solution of the original OPF.

Proof: Suppose $(\hat{y}_*, s_*)$ is optimal for OPF-cr (24)–(25). Theorem 1 implies that it is also optimal for OPF-ar. In particular $\hat{y}_* \in \hat{\mathbb{Y}}(s_*)$. Since G is a tree, $\hat{\mathbb{Y}}(s_*) = \hat{h}(\mathbb{X}(s_*))$ by Theorem 4 and hence there is a unique θ_* such that $h_{\theta_*}(\hat{y}_*)$ is a branch flow solution in $\mathbb{X}(s_*)$. This means $(h_{\theta_*}(\hat{y}_*), s_*)$ is feasible for OPF (10)–(11). Since OPF-ar is a relaxation of OPF, $(h_{\theta_*}(\hat{y}_*), s_*)$ is also optimal for OPF. ■

Remark 3: Theorem 1 implies that we can always solve efficiently a conic relaxation OPF-cr to obtain a solution of OPF-ar, provided there are no upper bounds on the power consumptions p_i^c, q_i^c. From a solution of OPF-ar, Theorem 4 and Corollary 5 prescribe a way to recover an optimal solution of OPF for radial networks. For mesh networks, however, the solution of OPF-ar may be spurious, i.e., there are no angles $\angle V_i, \angle I_{ij}$ that will satisfy the Kirchhoff laws if the angle recovery condition in Theorem 2 fails to hold. To deal with this, we next propose a way to convexify the network.

VI. Convexification of mesh network

In this section, we explain how to use phase shifters to convexify a mesh network so that an extended angle recovery condition can always be satisfied by any relaxed solution and can be mapped to a valid branch flow solution of the convexified network. As a consequence, the OPF for the convexified network can always be solved efficiently (in polynomial time).

A. Branch flow solutions

In this section we study power flow solutions and hence we fix an s. All quantities, such as $x, \hat{y}, \mathbb{X}, \hat{\mathbb{Y}}, \overline{X}, \overline{X}_T$, 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 $\mathbb{X}, \hat{\mathbb{Y}}, \overline{X}, \overline{X}_T$ are for any s; c.f. the more accurate notation in (4) and (5).

Phase shifters can be traditional transformers or FACTS (Flexible AC Transmission Systems) devices. They can increase transmission capacity and improve stability and power quality [37], [38]. 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 ϕ_{ij} across line (i,j), as shown in Figure 4. As before, let V_i denote the sending-end

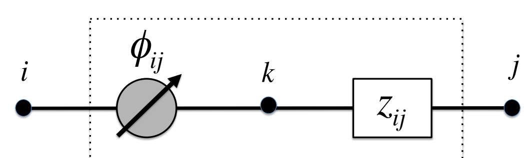

Fig. 4: Model of a phase shifter in line (i,j).

voltage. Define I_{ij} to be the *sending-end* current leaving node i towards node j. Let k be the point between the phase shifter ϕ_{ij} and line impedance z_{ij}. Let V_k and I_k be the voltage at k and current from k to j respectively. Then the effect of the idealized phase shifter is summarized by the following modeling assumption:

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

The power transferred from nodes i to j is still (defined to be) $S_{ij} := V_i I_{ij}^*$ which, as expected, is equal to the power $V_k I_k^*$ from nodes k to j since the phase shifter is assumed to be lossless. Notice however that, while the sending-end current from i is I_{ij}, the receiving-end current to j is now $I_{ij} e^{\mathbf{i}\phi_{ij}}$. Ohm's law is modified from (2) to

$$V_i \, e^{\mathbf{i}\phi_{ij}} - V_j \;=\; z_{ij} I_{ij} \, e^{\mathbf{i}\phi_{ij}} \tag{41}$$

Without phase shifter ($\phi_{ij} = 0$), this reduces to (2). Therefore the new system with phase shifters is defined by (1), (41), (3), and we refer to it as the *convexified* network.

Recall the set $\mathbb{X}$ of branch flow solutions defined in (4) (and (5)). The inclusion of phase shifters modifies the system and enlargers the solution set of the (new) branch flow equations. Formally, let

$$\overline{\mathbb{X}} \;:=\; \{x \,|\, x \text{ solves (1), (41), (3) for some } \phi\} \tag{42}$$

Here and henceforth, $\phi \in [-\pi, \pi]^m$. For any spanning tree T of G, let "$\phi \in T^\perp$" stands for "$\phi_{ij} = 0$ for all $(i,j) \in T$", i.e., ϕ involves only phase shifters in branches not in the spanning tree T. Define

$$\overline{\mathbb{X}}_T \;:=\; \left\{ x \,|\, x \text{ solves (1), (41), (3) for some } \phi \in T^\perp \right\} \tag{43}$$

Since (41) reduces to (2) when $\phi = 0$, $\mathbb{X} \subseteq \overline{\mathbb{X}}_T \subseteq \overline{\mathbb{X}}$.

From (3) and (41), we have

$$S_{ij} \;=\; V_i \, \frac{V_i^* \; e^{-\mathbf{i}\phi_{ij}} - V_j^*}{z_{ij}^* e^{-\mathbf{i}\phi_{ij}}}$$

leading to $V_i V_j^* \, e^{\mathbf{i}\phi_{ij}} = v_i - z_{ij}^* S_{ij}$. Hence $\theta_i - \theta_j = \beta_{ij} - \phi_{ij}$. This changes the angle recovery condition from whether there exists θ that solves $B\theta = \beta$ in (37) to whether there exists (θ, ϕ) that solves

$$B\theta \;=\; \beta - \phi \tag{44}$$

The condition (37) for the case without phase shifter corresponds to setting $\phi = 0$.

Our next key result implies that, given a relaxed solution $\hat{y} := (S, \ell, v, s_0) \in \hat{\mathbb{Y}}$, we can always recover a branch flow solution $x := (S, I, V, s_0) \in \overline{\mathbb{X}}$ of the convexified system. Moreover it suffices to use phase shifters in branches only outside a spanning tree. It extends Theorem 2 to the case with phase shifters.

Theorem 6: Let T be *any* spanning tree of G. Consider a relaxed solution $\hat{y} \in \hat{\mathbb{Y}}$ and the corresponding β defined by (28)–(29) in terms of $\hat{y}$.

1) There exists a unique solution (θ_*, ϕ_*) of (44) with $\phi_* \in T^\perp$. Specifically

$$\begin{aligned} \theta_* &= B_T^{-1} \beta_T \\ \phi_* &= \begin{bmatrix} 0 \\ \beta_{T^\perp} - B_{T^\perp} B_T^{-1} \beta_T \end{bmatrix} \end{aligned}$$

2) $h_{\theta_*}(\hat{y}) \in \overline{\mathbb{X}}_T$, i.e., $h_{\theta_*}(\hat{y})$ is a branch flow solution of the convexified system.
3) $\mathbb{Y} = \overline{\mathbb{X}} = \overline{\mathbb{X}}_T$ and hence $\hat{\mathbb{Y}} = \hat{h}(\overline{X}) = \hat{h}(\overline{\mathbb{X}}_T)$.

Proof:

1) Write $\phi = [\phi_T^t \;\; \phi_{T^\perp}^t]^t$ and set $\phi_T = 0$. Then (44) becomes

$$\begin{bmatrix} B_T \\ B_{T^\perp} \end{bmatrix} \theta \;=\; \begin{bmatrix} \beta_T \\ \beta_{T^\perp} \end{bmatrix} - \begin{bmatrix} 0 \\ \phi_{T^\perp} \end{bmatrix}$$

The same argument as in the proof of Theorem 2 shows that a solution (θ_*, ϕ_*) to (44), with $\theta_* \in T^\perp$, exists and is unique if and only if (since B_T is invertible)

$$B_{T^\perp} B_T^{-1} \beta_T = \beta_{T^\perp} - \phi_{T^\perp}$$

i.e., iff $\phi_{T^\perp} = \beta_{T^\perp} - B_{T^\perp} B_T^{-1} \beta_T$. Hence the (θ_*, ϕ_*) given in the theorem solves (44).

2) Theorem 2 then implies $h_{\theta_*}(S, \ell, v, s_0) \in \overline{\mathbb{X}}_T$ with $\phi_* \in T^\perp$ given in assertion 1.
3) This follows from assertions 1 and 2.

■

B. Optimal power flow

Theorem 6 suggests using phase shifters to convexify a mesh network so that the angle recovery procedure succeeds for *any* solution of OPF-ar. This ensures the success of both the angle relaxation step and the conic relaxation step so that the overall relaxation is exact for the convexified network. Convexification thus modifies a 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 (10)–(11), OPF-ar (19)–(20), the following problem where a phase shifter can be added potentially to every line ($\phi \in [-\pi, \pi]^m$):

OPF-ps:

$$\begin{aligned} \min_{x,s,\phi} \quad & f\left(\hat{h}(x), s^g\right) \\ \text{subject to} \quad & x \in \overline{\mathbb{X}}, \;\; (v, s_0, s) \in \mathbb{S} \end{aligned}$$

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

OPF-psT:

$$\min_{x,s,\phi} \quad f\left(\hat{h}(x), s^g\right)$$
$$\text{subject to} \quad x \in \overline{\mathbb{X}}_T,\ (v, s_0, s) \in \mathbb{S},\ \phi \in T^{\perp}$$

Let the optimal values of OPF, OPF-ar, OPF-ps, and OPF-psT be f_*, f_{ar}, f_{ps}, and f_{psT} respectively.

Theorem 6 implies that $\mathbb{X} \subseteq \mathbb{Y} = \overline{\mathbb{X}} = \overline{\mathbb{X}}_T$ for any spanning tree T. Hence we have

Corollary 7: For any spanning tree T, $f_* \geq f_{ar} = f_{ps} = f_{psT}$, with equality if there is a solution $(\hat{y}_{ar}, s^g_{ar})$ of OPF-ar that satisfies (34).

Corollary 7 has several important implications:

1) Theorem 1 implies that we can solve OPF-ar efficiently through conic relaxation to obtain a relaxed solution $(\hat{y}_{ar}, s_{ar})$. An optimal solution of OPF may or may not be recovered from it. If $\hat{y}_{ar}$ satisfies the angle recovery condition (34) with respect to s_{ar}, then Theorem 2 guarantees a unique θ_* such that the inverse projection $(\hat{h}_{\theta_*}(\hat{y}_{ar}), s_{ar})$ is indeed optimal for OPF.
2) In this case, Corollary 7 implies that adding any phase shifters to the network cannot further reduce the cost since $f_* = f_{ar} = f_{ps}$.
3) If (34) is not satisfied, then $\hat{y}_{ar} \notin \hat{h}(\mathbb{X})$ and there is no inverse projection that can recover an optimal solution of OPF from $(\hat{y}_{ar}, s_{ar})$. In this case, $f_* \geq f_{ar}$. Theorem 6 implies that if we allow phase shifters, we can always attain $f_{ar} = f_{ps}$ with the relaxed solution $(\hat{y}_{ar}, s_{ar})$, with potentially strict improvement over the network without phase shifters (when $f_* > f_{ar}$).
4) Moreover, Corollary 7 implies that such benefit can be achieved with phase shifters only in branches outside an arbitrary spanning tree T.

VII. Simulations

For radial networks, Theorem 4 guarantees that both the angle relaxation and the conic relaxation are exact and hence the solution strategy outlined in Figure 3 will always yield an optimal solution of OPF (10)–(11). For mesh networks, the angle relaxation may be inexact (i.e., the angle recovery condition in Theorem 2 may fail), 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 is a typical range of phase shifter angles to implement an optimal solution of 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 2,383 and 2,737 buses. The data for all the test cases were extracted from the library of built in models of the MATPOWER toolbox [39] 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, even though we have not considered line limits in our analytical model. All these systems are mesh networks, but very sparse (see below).

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 II.
- *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 III.

Solution methods. We use the "SEDUMI" solver of the CVX optimization toolbox in Matlab [40]. We first solved the SOCP relaxation OPF-cr (24)–(25) for a solution $(\hat{y}, s)$ of OPF-ar. We verified the exactness of the SOCP relaxation by checking if equality is attained in (23) at optimality. In all test cases, OPF-cr was exact as 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 (c.f. (7)), 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 $(\hat{y}, s)$ of OPF-ar, we checked if the angle recovery condition in Theorem 2 was satisfied. If it were then we could recover the voltage angles θ and an optimal solution $(h_\theta(\hat{y}), s)$ of the original OPF (10)–(11). In all test cases, however, the angle recovery condition failed and hence no $(h_\theta(\hat{y}), s)$ was feasible for OPF without phase shifters. We computed the phase shifter angles $\phi \in T^{\perp}$ outside a spanning tree T and

Test cases	# links (m)	No PS: Min loss (OPF, MW)	With phase shifters (PS): Min loss (OPF-cr, MW)	# required PS ($m-n$)		# active PS $\lvert\phi_i\rvert > 0.1^\circ$		Angle range ($^\circ$) $[\phi_{\min}, \phi_{\max}]$
IEEE 14-Bus	20	0.546	0.545	7	(35%)	2	(10%)	$[-2.1,\ 0.1]$
IEEE 30-Bus	41	1.372	1.239	12	(29%)	3	(7%)	$[-0.2,\ 4.5]$
IEEE 57-Bus	80	11.302	10.910	24	(30%)	19	(24%)	$[-3.5,\ 3.2]$
IEEE 118-Bus	186	9.232	8.728	69	(37%)	36	(19%)	$[-1.9,\ 2.0]$
IEEE 300-Bus	411	211.871	197.387	112	(27%)	101	(25%)	$[-11.9,\ 9.4]$
New England 39-Bus	46	29.915	28.901	8	(17%)	7	(15%)	$[-0.2,\ 2.2]$
Polish (case2383wp)	2,896	433.019	385.894	514	(18%)	376	(13%)	$[-20.1,\ 16.8]$
Polish (case2737sop)	3,506	130.145	109.905	770	(22%)	433	(12%)	$[-21.9,\ 21.7]$

TABLE II: Loss minimization. Min loss without phase shifters (PS) was computed using SDP relaxation of OPF (10)–(11); min loss with phase shifters was computed using SOCP relaxations OPF-cr (24)–(23) of OPF-ar. The "(%)" indicates the number of PS as a percentage of #links.

Test cases	No PS: Max loadability (OPF)	With phase shifters (PS): Max loadability (OPF-cr)	# required PS ($m-n$)		# active PS $\lvert\phi_i\rvert > 0.1^\circ$		Angle range ($^\circ$) $[\phi_{\min}, \phi_{\max}]$
IEEE 14-Bus	195.0%	195.2%	7	(35%)	6	(30%)	$[-0.5,\ 1.4]$
IEEE 30-Bus	156.7%	158.7%	12	(29%)	9	(22%)	$[-0.4,\ 12.4]$
IEEE 57-Bus	108.2%	118.3%	24	(30%)	24	(30%)	$[-13.1,\ 23.2]$
IEEE 118-Bus	203.7%	204.9%	69	(37%)	64	(34%)	$[-16.5,\ 22.3]$
IEEE 300-Bus	106.8%	112.8%	112	(27%)	103	(25%)	$[-15.0,\ 16.5]$
New England 39-Bus	109.1%	114.8%	8	(17%)	5	(11%)	$[-6.3,\ 10.6]$
Polish (case2383wp)	101.4%	106.6%	514	(18%)	435	(15%)	$[-19.6,\ 19.4]$
Polish (case2737sop)	127.6%	132.5%	770	(22%)	420	(12%)	$[-16.7,\ 17.0]$

TABLE III: Loadability maximization. Max loadability without phase shifters (PS) was computed using SDP relaxation of OPF (10)–(11); max loadability with phase shifters was computed using SOCP relaxations OPF-cr (24)–(23) of OPF-ar. The "(%)" indicates the number of PS as a percentage of #links.

the corresponding unique $(h_\theta(\hat{y}), s)$ that was an optimal solution of OPF for the *convexified* network.

To place the phase shifters, we have used a minimum spanning tree of the network where the weights on the lines are their reactance values. In Tables II and III, we report the number $m-n$ of phase shifters potentially required on every link outside the minimum spanning tree, as well as the number of active phase shifters (i.e., those with a phase angles greater than 0.1°) and the range of their phase angles at optimality. The optimal choice of spanning tree, e.g., to minimize the number of active phase shifters and the range of their angles, remains an open problem.

We also report the optimal objective values of OPF with and without phase shifters in Tables II and III. The optimal values of OPF without phase shifters were obtained by implementing the SDP formulation and relaxation proposed in [16] for solving OPF (10)–(11). We verified the exactness of the SDP relaxation by checking if the solution matrix was of rank one [13], [16]. In all test cases, the SDP relaxation was exact and hence the optimal objective values reported were indeed the optimal value of OPF (10)–(11). As expected, the optimal loss (Table II) and the optimal loadability (Table III) for OPF-ar (equivalently OPF-cr) are, respectively, lower and higher than the corresponding optimal values of OPF. This confirms that the solutions obtained from the SOCP relaxation are infeasible for the original OPF but can be implemented with phase shifters, at a lower loss or higher loadability.

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 [13]. 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 II and III:

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, convexification is design for simplicity: it guarantees an efficient solution for optimal power flow.
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 II). 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 23° in magnitude.

VIII. Conclusion

We have presented a branch flow model that focuses on branch current and power flows instead of nodal injections, and demonstrated how it can be used for analysis and optimization of mesh as well as radial networks. Our results confirm that radial networks are computationally much simpler than mesh networks and we should exploit this advantage whenever we can. 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.

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 so this strategy guarantees a globally optimal solution. For mesh networks the angle recovery condition, which is easily checkable, can fail. However, with judiciously placed phase shifters, both relaxations will always be exact for the convexified network and the optimal cost of the convexified network will be no higher and potentially strictly lower than the original network. Moreover phase shifters are needed only on lines outside an arbitrary spanning tree of the network graph.

Since power networks in practice are very sparse, the number of lines not in a spanning tree may be relatively small. 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 the subsequent simplicity in network operation, market operation, and investment decisions.

Acknowledgment

We are grateful to Mani Chandy and Javad Lavaei of Caltech, Christopher Clarke, Michael Montoya, and Robert Sherick of the Southern California Edison for helpful discussions. We acknowledge the support of the Resnick Institute at Caltech, National Science Foundation through NetSE grant CNS 0911041, Department of Energy's ARPA-E through grant de-ar0000226, the National Science Council of Taiwan (R. O. C.) through grant NSC 101-3113-P-008-001, Southern California Edison, Cisco, and the Okawa Foundation.

IX. Appendix

A. OPF-ar has zero duality gap

Theorem 8: Suppose OPF-cr has a finite optimal value and a strictly feasible point. Then both OPF-cr and OPF-ar have zero duality gap.

Proof:

Recall that the original OPF is feasible by assumption. Hence OPF-ar and OPF-cr are both feasible. In particular the optimal value of OPF-cr is either finite or $-\infty$. We assume without loss of generality that, for all i, $\underline{p}_i^g < \overline{p}_i^g$, $\underline{q}_i^g < \overline{q}_i^g$ in (6) and $\underline{v}_i < \overline{v}_i$ in (8) so that $\mathbb{S}$ has a nonempty interior. [6]

Let f_{cr}, d_{cr} denote the optimal primal and dual value of OPF-ar and f_{ar}, d_{ar} those of OPF-ar. The weak duality theorem and Theorem 1 imply

$$d_{cr} \leq f_{cr} \quad = \quad f_{ar} \geq d_{ar}$$

We will prove strong duality first for OPF-cr and then for OPF-ar.

Theorem 1 implies that OPF-cr is a convex program. By assumption its optimal value f_{cr} is finite and Slater's condition is satisfied, i.e., there exists an $(\hat{y}, s)$ in the relative interior of the feasible set such that

$$\ell_{ij} \quad > \quad \frac{P_{ij}^2 + Q_{ij}^2}{v_i} \tag{45}$$

Hence OPF-cr has a zero duality gap [41, Chapter 5.3.1]. We then have

$$d_{cr} = f_{cr} \quad = \quad f_{ar} \geq d_{ar} \tag{46}$$

To prove that OPF-ar has zero duality gap, it suffices to prove that $d_{ar} \geq d_{cr}$. Notice that the dual variables corresponding to the inequality constraints (23) in OPF-cr are restricted to be nonnegative whereas the dual variables corresponding to the equality constraints (15) are unconstrained. The dual problems of OPF-cr and OPF-ar are otherwise identical. Hence $d_{ar} \geq d_{cr}$ and (46) completes the proof.

■

[6] Otherwise some of these quantities are constants and can be eliminated from the optimization variables.

B. $\hat{h}$ is injective on $\mathbb{X}$

The uniqueness of the θ in Theorem 2 for each $\hat{y}$ implies that the projections into $\hat{\mathbb{Y}}$ of two branch flow solutions in $\mathbb{X}$ must be distinct. Clearly this does not hold for points in $\mathbb{Y} \setminus \mathbb{X}$ that are not branch flow solutions. Formally

Fact 9: The mapping $\hat{h}$ is one-to-one on $\mathbb{X}$, i.e., if $x, x' \in \mathbb{X}$ and $x \neq x'$, then $\hat{h}(x) \neq \hat{h}(x')$.

Proof: Suppose not and there are x and x' in $\mathbb{X}$ such that $\hat{h}(x) = \hat{h}(x') = \hat{y}$. Let $\beta = \beta(\hat{y})$ be the vector defined by $\hat{y}$ through (28). Theorem 2 implies that there is a unique θ_* such that $h_{\theta_*}(\hat{y})$ is the only branch flow solution in $\mathbb{X}$ whose projection is $\hat{y}$. Hence $x = x' = h_{\theta_*}(\hat{y})$. ■

To visualize this, we identify $\mathbb{C}^{2m+n+1}$ with $\mathbb{R}^{3m+n+2} \times [-\pi, \pi]^n$. Then $(S, I, V, s_0) \in \mathbb{C}^{2m+n+1}$ and $(S, \ell, v, s_0; \theta) \in \mathbb{R}^{3m+n+2} \times [-\pi, \pi]^n$ refer to the same vector if

$$h_\theta(S, \ell, v, s_0) \;\; = \;\; (S, I, V, s_0)$$

and we use them interchangeably in that case. Then an equivalent description of $\mathbb{Y}$ is

$$\mathbb{Y} \;\; := \;\; \{ \; (S, \ell, v, s_0; \theta) \,|\, (S, \ell, v, s_0) \text{ solves (12)–(15)}, \; \theta \in [-\pi, \pi]^n \; \}$$

Clearly $\mathbb{X} \subseteq \mathbb{Y}$ and $\hat{h}(\mathbb{X}) \subseteq \hat{h}(\mathbb{Y}) = \hat{\mathbb{Y}}$. Their relationship is illustrated in Figure 5.

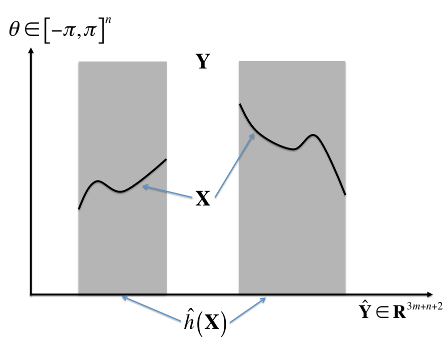

Fig. 5: Fact 9: $\hat{h}$ is injective on $\mathbb{X}$. $\mathbb{X}$ can be represented by curves in the space $\mathbb{R}^{3m+n+2} \times [-\pi, \pi]^n$, and $\mathbb{Y}$ by the shaded areas (higher dimensional).

References

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

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

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

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

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

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

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

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

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

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

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

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

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

[14] J. Lavaei. Zero duality gap for classical OPF problem convexifies fundamental nonlinear power problems. In *Proc.of the American Control Conf.*, 2011.

[15] S. Sojoudi and J. Lavaei. Network topologies guaranteeing zero duality gap for optimal power flow problem. *Submitted for publication*, 2011.

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

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

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

[19] B. Lesieutre, D. Molzahn, A. Borden, and C. L. DeMarco. Examining the limits of the application of semidefinite programming to power flow problems. In *Proc. Allerton Conference*, 2011.

[20] I. A. Hiskens and R. Davy. Exploring the power flow solution space boundary. *IEEE Trans. Power Systems*, 16(3):389–395, 2001.

[21] B. C. Lesieutre and I. A. Hiskens. Convexity of the set of feasible injections and revenue adequacy in FTR markets. *IEEE Trans. Power Systems*, 20(4):1790–1798, 2005.

[22] Yuri V. Makarov, Zhao Yang Dong, and David J. Hill. On convexity of power flow feasibility boundary. *IEEE Trans. Power Systems*, 23(2):811–813, May 2008.

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

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

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

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

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

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

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

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

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

[32] Lingwen Gan, Na Li, Ufuk Topcu, and Steven Low. Branch flow model for radial networks: convex relaxation. Submitted for publication, March 2012.

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

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

[35] L. R. Foulds. *Graph Theory Applications*. Springer-Verlag, 1992.

[36] Norman Biggs. *Algebraic graph theory*. Cambridge University Press, 1993. Cambridge Mathematical Library.

[37] Narain G. Hingoranl and Laszlo Gyugyi. *Understanding FACTS*. Wiley-Interscience, 2000.

[38] Xiao-Ping Zhang, Christian Rehtanz, and Bikash Pal. *Flexible AC Transmission Systems: Modelling and Control*. Springer, 2006.

[39] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas. MATPOWER's extensible optimal power flow architecture. In *Proc. IEEE PES General Meeting*, pages 1–7, 2009.

[40] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. http://cvxr.com/cvx, April 2011.

[41] Dimitri P. Bertsekas. *Nonlinear programming*. Athena Scientific, 1995.