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 Carpentier'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

Manuscript received May 11, 2012; revised July 22, 2012, November 18, 2012, January 04, 2013, and March 01, 2013; accepted March 03, 2013. Date of publication April 23, 2013; date of current version July 18, 2013. This work was supported by NSF through NetSE grant CNS 0911041, DoE's ARPA-E through grant DE-AR0000226, the National Science Council of Taiwan (R. O. C.) through grant NSC 101-3113-P-008-001, SCE, the Resnick Institute of Caltech, Cisco, and the Okawa Foundation. A preliminary and abridged version has appeared in [1]. Paper no. TPWRS-00424-2012.

The authors are with the Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2013.2255317

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 nononvex 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

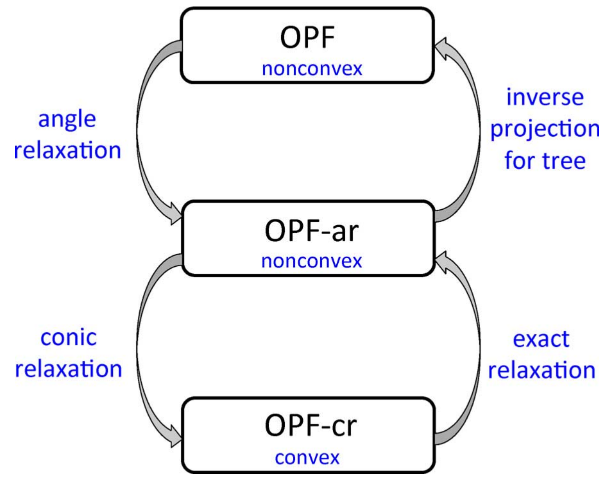


Fig. 1. Proposed solution strategy for solving OPF.

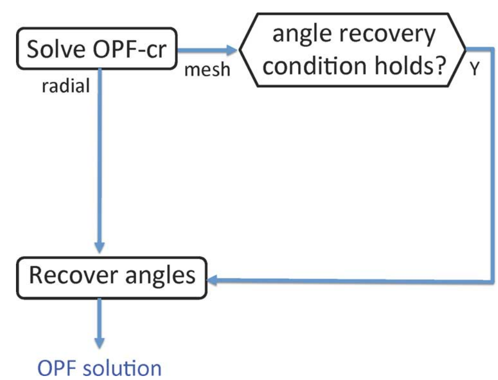


Fig. 2. Proposed algorithm for solving OPF (11)–(12) without phase shifters. The details are explained in Sections II–V.

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 $(\bmod\ 2\pi)$ 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., $s := (s_i, i = 1, \ldots, n)$, $S := (S_{ij}, (i,j) \in E)$. For a vector $a = (a_1, \ldots, a_k)$, a_{-i} denotes $(a_1, \ldots, a_{i-1}, a_{i+1}, a_k)$. For 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 e, (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 $\theta = 0 \ (\mathrm{mod}\ 2\pi)$ if $\theta = 2\pi k$, and $\theta = \phi \ (\mathrm{mod}\ 2\pi)$ if $\theta - \phi = 2\pi k$, for some integer k. For a d-dimensional vector α, $\mathcal{P}(\alpha)$ denotes its projection onto $(-\pi, \pi]^d$ by taking modulo 2π componentwise.

A. Branch Flow Model

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, 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 $E \backslash E_T$ 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.[1] For each link $(i,j) \in E$, let $z_{ij} = r_{ij} + \mathbf{i}x_{ij}$ be the complex impedance on the line, and $y_{ij} := 1/z_{ij} =: g_{ij} - \mathbf{i}b_{ij}$ be the corresponding admittance. For each node $i \in N$, let $z_i = r_i + \mathbf{i}x_i$ be the shunt impedance from i to ground, and $y_i := 1/z_i =: g_i - \mathbf{i}b_i$.[2]

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 injection, which is generation minus load on bus i. 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.

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:

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

the definition of branch power flow:

$$S_{ij} = V_i I_{ij}^*, \quad \forall (i,j) \in E \tag{2}$$

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

$$\sum_{k: j \rightarrow k} S_{jk} - \sum_{i: i \rightarrow j} (S_{ij} - z_{ij}|I_{ij}|^2) + y_j^* |V_j|^2 = s_j. \tag{3}$$

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 $\mathbb{X}(s) \subseteq \mathbb{C}^{2m+n+1}$ be the set of all branch flow solutions with respect to a given s:

$$\mathbb{X}(s) := \{x := (S, I, V, s_0) \mid x \text{ solves (1)-(3) given } s\} \tag{4}$$

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

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

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), de-

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

[2]The shunt admittance y_i represents capacitive devices on bus i only and a line is modeled by a series admittance y_{ij} without shunt elements. If a shunt admittance $\mathbf{i}\tilde{b}_{ij}/2$ is included on each end of line (i,j) in the π-model, then the line flow should be $|S_{ij} - \mathbf{i}\tilde{b}_{ij}|V_i|^2/2|$.

pending 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), s is also an optimization variable. Let $p_i := p_i^g - p_i^c$ and $q_i := q_i^g - q_i^c$ where p_i^g and q_i^g (p_i^c and q_i^c) 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^g represent the placement and sizing of shunt capacitors. In addition to (1)–(3), we impose the following constraints on power generation: for $i \in N$:

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

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 [27] and [28], p_i^g are the fixed (solar) power outputs and the reactive power q_i^g are the control variables. For power consumption, we require, for $i \in N$

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

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

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

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

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

We allow any objective function that is convex and does not depend on the angles $\angle V_i, \angle I_{ij}$ 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 $c_i p_i^g$, and maximize energy savings through conservation voltage reduction (CVR). Then the objective function takes the form (see [27] and [28])

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

for some given constants $c_i, \alpha_i \geq 0$.

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 $s^g - s^c$ or (s^g, 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 $\hat{y} := \hat{y}(s) := (S, \ell, v, s_0)$ 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 $f(\hat{h}(x), s)$. 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^c, and independent of S (condition A3). Let

$$\mathbb{S} := \{(S, v, s_0, s) \,|\, (v, s_0, s) \text{ satisfies } (6) - (9)\}.$$

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

The optimal power flow problem is

OPF:

$$\min_{x,s} \quad f(\hat{h}(x), s) \tag{11}$$

$$\text{subject to} \quad x \in \mathbb{X}, \quad (S, v, s_0, s) \in \mathbb{S} \tag{12}$$

where $\mathbb{X}$ 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 ℓ, nonincreasing in load s^c, 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 f 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 $V_j = V_i - z_{ij} S_{ij}^* / V_i^*$. Taking the magnitude squared, we have $v_j = v_i + |z_{ij}|^2 \ell_{ij} - (z_{ij} S_{ij}^* + z_{ij}^* S_{ij})$. Using (3) and (2) and in terms of real variables, we therefore have

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

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

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

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

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 $v_i := \|V_i\|^2$
$s_i = p_i + \mathbf{i}q_i$ $p_i = p_i^g - p_i^c$ $q_i = q_i^g - q_i^c$	net complex load power on bus i net real power equals generation minus load; net reactive power equals generation minus load
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)
$\mathbb{X}$ $\hat{\mathbb{Y}}$ $\overline{\mathbb{Y}}$	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 (13)–(16) either for a given s or for some s; set of all relaxed branch flow solutions that satisfy (13)–(15) and (22) either for a given s or for some s;
$x = (S, I, V, s_0) \in \mathbb{X}$ $\hat{y} = (S, \ell, v, s_0) \in \hat{\mathbb{Y}}$ $\hat{y} = \hat{h}(x); \quad x = h_\theta(\hat{y})$	vector x of power flow variables and its projection $\hat{y}$; projection mapping $\hat{y}$ and an inverse h_θ
z_{ij}, y_i	impedance on line (i,j) and shunt admittance from bus i to ground
$f = f\left(\hat{h}(x), s\right)$	objective function of **OPF**

equations were first proposed in [25], [26] to model radial distribution circuits. They define a system of equations in the variables $(P, Q, \ell, v, p_0, q_0) := (P_{ij}, Q_{ij}, \ell_{ij}, (i,j) \in E,\ v_i, i = 1, \ldots, n,\ p_0, q_0)$. We often use (S, ℓ, v, s_0) as a shorthand for $(P, Q, \ell, v, p_0, q_0)$. The relaxed model has a solution under A4.

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 $|V_0| \simeq 1$ and r_{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, ℓ, 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 a 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.

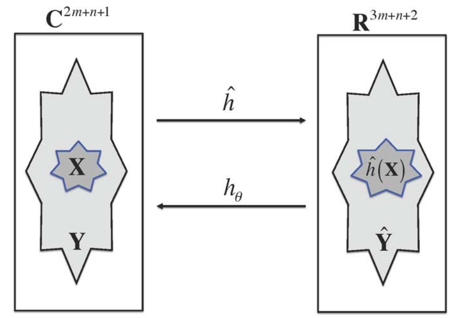


Fig. 3. $\mathbb{X}$ is the set of branch flow solutions and $\hat{\mathbb{Y}} = \hat{h}(\mathbb{Y})$ is the set of relaxed solutions. The inverse projection h_θ is defined in Section V.

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 $\hat{h} : \mathbb{C}^{2m+n+1} \to \mathbb{R}^{3m+n+2}$ by $\hat{h}(S, I, V, s_0) = (P, Q, \ell, v, p_0, q_0)$ where

$$P_{ij} = \mathrm{Re}\, S_{ij}, \quad Q_{ij} = \mathrm{Im}\, S_{ij}, \quad \ell_{ij} = |I_{ij}|^2 \tag{17}$$
$$p_i = \mathrm{Re}\, s_i, \quad q_i = \mathrm{Im}\, s_i, \quad v_i = |V_i|^2. \tag{18}$$

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 solutions:[3]

$$\mathbb{Y} := \{y := (S, I, V, s_0) | \hat{h}(y) \text{ solves } (13) - (16)\}. \tag{19}$$

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

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

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 Fig. 3.

B. Two Relaxations

Consider the OPF with angles relaxed:

$$\min_{x,s} \quad f(\hat{h}(x), s)$$
$$\text{subject to} \quad x \in \mathbb{Y}, \quad (S, v, s_0, s) \in \mathbb{S}.$$

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

[3]As mentioned earlier, the set defined in (19) 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.

$\mathbb{Y}$ involves angles $\angle V_i$, $\angle I_{ij}$, this problem is equivalent to the following

OPF-ar:

$$\min_{\hat{y},s} \quad f(\hat{y}, s) \tag{20}$$

$$\text{subject to} \quad \hat{y} \in \hat{\mathbb{Y}}, \quad (S, v, s_0, s) \in \mathbb{S}. \tag{21}$$

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

$$\ell_{ij} \geq \frac{P_{ij}^2 + Q_{ij}^2}{v_i}, \quad (i,j) \in E. \tag{22}$$

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

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

Consider the following conic relaxation of OPF-ar:

OPF-cr:

$$\min_{\hat{y},s} \quad f(\hat{y}, s) \tag{23}$$

$$\text{subject to} \quad \hat{y} \in \overline{\mathbb{Y}}, \quad (S, v, s_0, s) \in \mathbb{S}. \tag{24}$$

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:

1) OFP-cr is convex. Moreover, if there are no upper bounds on loads, then the conic relaxation is exact so that *any* optimal solution $(\hat{y}_{\rm cr}, s_{\rm cr})$ 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 $(\hat{y}_{\rm ar}, s_{\rm ar})$ 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 $(\hat{y}_{\rm ar}, s_{\rm ar})$ to a feasible solution of OPF. Our characterization can be used to determined if $(\hat{y}_{\rm ar}, s_{\rm ar})$ 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 $(\hat{y}_{\rm ar}, s_{\rm ar})$ 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 $\overline{p}_i^c = \overline{q}_i^c = \infty$, $i \in N$. 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 $\overline{\mathbb{Y}}$ 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.[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 $(\hat{y}_*, s_*) := (S_*, \ell_*, v_*, s_{*0}^g, s_{*0}^c, s_*^g, s_*^c)$ 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 $(\tilde{y}, \tilde{s}) = (\tilde{S}, \tilde{\ell}, \tilde{v}, \tilde{s}_0^g, \tilde{s}_0^c, \tilde{s}^g, \tilde{s}^c)$ defined by

$$\begin{aligned}
\tilde{v} &= v_*, \quad \tilde{s}^g = s_*^g \\
\tilde{\ell}_{ij} &= [\ell_*]_{ij} - \varepsilon, \quad \tilde{\ell}_{-ij} = [\ell_*]_{-ij} \\
\tilde{S}_{ij} &= [S_*]_{ij} - z_{ij}\varepsilon/2, \quad \tilde{S}_{-ij} = [S_*]_{-ij} \\
\tilde{s}_i^c &= [s_*^c]_i + z_{ij}\varepsilon/2, \quad \tilde{s}_j^c = [s_*^c]_j + z_{ij}\varepsilon/2 \\
\tilde{s}_{-i}^c &= [s_*^c]_{-i}, \quad \tilde{s}_{-j}^c = [s_*^c]_{-j}
\end{aligned}$$

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)–(9), (13)–(15) and (22), and hence is indeed a feasible point. Since $(\hat{y}_*, s_*)$ is feasible, (6)–(9) hold for $(\tilde{y}, \tilde{s})$ too. Similarly, $(\tilde{y}, \tilde{s})$ 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 $(\tilde{y}, \tilde{s})$ 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

$$\begin{aligned}
\tilde{s}_i &= \tilde{s}_i^g - \tilde{s}_i^c = [s_*^g]_i - [s_*^c]_i - z_{ij}\varepsilon/2 \\
&= \sum_{i \to j'} [S_*]_{ij'} - \sum_{k \to i} ([S_*]_{ki} - z_{ki}[\ell_*]_{ki}) \\
&\quad + y_i^* v_i - z_{ij}\varepsilon/2 \\
&= \sum_{i \to j', j' \neq j} \tilde{S}_{ij'} + (\tilde{S}_{ij} + z_{ij}\varepsilon/2) \\
&\quad - \sum_{k \to i} (\tilde{S}_{ki} - z_{ki}\tilde{\ell}_{ki}) + y_i^* \tilde{v}_i - z_{ij}\varepsilon/2 \\
&= \sum_{i \to j'} \tilde{S}_{ij'} - \sum_{k \to i} (\tilde{S}_{ki} - z_{ki}\tilde{\ell}_{ki}) + y_i^* \tilde{v}_i.
\end{aligned}$$

[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

$$\begin{aligned}\tilde{s}_j &= \tilde{s}^g_j - \tilde{s}^c_j = [s^g_*]_j - [s^c_*]_j - z_{ij}\varepsilon/2 \\ &= \sum_{j\to k}[S_*]_{jk} - \sum_{i'\to j}([S_*]_{i'j} - z_{i'j}[\ell_*]_{i'j}) \\ &\quad + y_j^* v_j - z_{ij}\varepsilon/2 \\ &= \sum_{j\to k}\tilde{S}_{jk} - \sum_{i'\to j, i'\neq i}(\tilde{S}_{i'j} - z_{i'j}\tilde{\ell}_{i'j}) + y_j^*\tilde{v}_j \\ &\quad - ((\tilde{S}_{ij} + z_{ij}\varepsilon/2) - z_{ij}(\tilde{\ell}_{ij} + \varepsilon)) - z_{ij}\varepsilon/2 \\ &= \sum_{j\to k}\tilde{S}_{jk} - \sum_{i'\to j}(\tilde{S}_{i'j} - z_{i'j}\tilde{\ell}_{i'j}) + y_j^*\tilde{v}_j.\end{aligned}$$

Hence (13)–(14) hold at nodes i, j.

For (15) across link (i,j):

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

For (22) across link (i,j), we have

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

Since $[v_*]_i[\ell_*]_{ij} - [P_*]_{ij}^2 - [Q_*]_{ij}^2 > 0$, 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: Assumption A3 is used in the proof here to contradict the optimality of $(\hat{y}_*, s_*)$. Instead of A3, if $f(\hat{y}, s)$ is nondecreasing in ℓ, the same argument shows that, given an optimal $(\hat{y}_*, s_*)$ with a strict inequality $[v_*]_i[\ell_*]_{ij} > {[P_*]_{ij}}^2 + {[Q_*]_{ij}}^2$, one can choose $\varepsilon > 0$ to obtain another optimal point $(\tilde{y}, \tilde{s})$ that attains equality and has a cost $f(\tilde{y}, \tilde{s}) \leq f(\hat{y}_*, s_*)$. 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^c_* on buses i and j to obtain the alternative feasible solution $(\tilde{y}, \tilde{s})$. 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 $(\hat{y}, s)$ 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 $\hat{y}$, given any s.

Hence, for the rest of Section V, we fix an s. 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

$$C_{ie} = \begin{cases} 1 & \text{if link } e \text{ leaves node } i \\ -1 & \text{if link } e \text{ enters node } i \\ 0 & \text{otherwise.}\end{cases} \tag{25}$$

The first row of C corresponds to node 0 where $V_0 = |V_0|e^{\mathbf{i}\theta_0}$ 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 $e \in E, i = 1, \ldots, n$

$$B_{ei} = \begin{cases} 1 & \text{if link } e \text{ leaves node } i \\ -1 & \text{if link } e \text{ enters node } i \\ 0 & \text{otherwise.}\end{cases}$$

Since G is connected, $m \geq n$ and $\operatorname{rank}(B) = n$ [43]. 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 $e = 1, \ldots, n$. Then B can be partitioned into

$$B = \begin{bmatrix} B_T \\ B_\perp \end{bmatrix} \tag{26}$$

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

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

$$\beta_{ij} := \angle(v_i - z_{ij}^* S_{ij}), \quad (i,j) \in E. \tag{27}$$

Informally, β_{ij} is the phase angle difference across link (i,j) that is implied by the relaxed solution $\hat{y}$. Write β as

$$\beta = \begin{bmatrix} \beta_T \\ \beta_\perp \end{bmatrix} \tag{28}$$

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

Recall the projection mapping $\hat{h} : \mathbb{C}^{2m+n+1} \to \mathbb{R}^{3m+n+2}$ defined in (17)–(18). 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

$$S_{ij} := P_{ij} + \mathbf{i}Q_{ij} \tag{29}$$
$$I_{ij} := \sqrt{\ell_{ij}}\, e^{\mathbf{i}(\theta_i - \angle S_{ij})} \tag{30}$$
$$V_i := \sqrt{v_i}\, e^{\mathbf{i}\theta_i} \tag{31}$$
$$s_0 := p_0 + \mathbf{i}q_0. \tag{32}$$

These mappings are illustrated in Fig. 3.

By definition of $\hat{h}(\mathbb{X})$ and $\hat{\mathbb{Y}}$, a branch flow solution in $\mathbb{X}$ can be recovered from a given relaxed solution $\hat{y}$ if $\hat{y}$ is in $\hat{h}(\mathbb{X})$ and cannot 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 $h_\theta(\hat{y})$ 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 $\angle V_i, \angle I_{ij}$ 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 $\tilde{\beta}$ be the extension of β from directed links to undirected links: if $(i,j) \in E$ then $\tilde{\beta}_{ij} := \beta_{ij}$ and $\tilde{\beta}_{ji} := -\beta_{ij}$. For any d-dimensional vector α, let $\mathcal{P}(\alpha)$ denote its projection onto $(-\pi, \pi]^d$ by taking modulo 2π componentwise.

Theorem 2: Let T be *any* spanning tree of G. Consider a relaxed solution $\hat{y} \in \hat{\mathbb{Y}}$ and the corresponding $\beta = \beta(\hat{y})$ defined in (27)–(28).

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

$$B_\perp B_T^{-1} \beta_T = \beta_\perp \ (\mathrm{mod}\ 2\pi). \tag{33}$$

2) The angle recovery condition (33) holds if and only if for every cycle c in G

$$\sum_{(i,j) \in c} \tilde{\beta}_{ij} = 0 \ (\mathrm{mod}\ 2\pi). \tag{34}$$

3) If (33) holds, then $\theta_* = \mathcal{P}(B_T^{-1} \beta_T)$.

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 B_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 $(\mathrm{mod}\ 2\pi)$ 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 $\mathbb{X}$. Fix any $\hat{y} := (S, \ell, v, s_0)$ in $\hat{\mathbb{Y}}$ and the corresponding $\beta := \beta(\hat{y})$ defined in (27). Consider the equation

$$B\theta = \beta + 2\pi k \tag{35}$$

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

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

[5]Using the connectedness of G and the definition of B, one can argue that α 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 $\hat{\mathbb{Y}}$ and the corresponding $\beta := \beta(\hat{y})$ defined in (27):

1) $h_\theta(\hat{y})$ is a branch flow solution in $\mathbb{X}$ 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 (θ, 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

$$\begin{aligned} \theta_i - \theta_j &= \angle(v_i - z_{ij}^* V_i I_{ij}^*) + 2\pi k_{ij} \\ &= \angle\ V_i (V_i - z_{ij} I_{ij})^* + 2\pi k_{ij}. \end{aligned}$$

Hence

$$\angle V_j = \theta_j = \angle(V_i - z_{ij} I_{ij}) - 2\pi k_{ij}. \tag{36}$$

From (15) and (2), 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 and (36) imply $V_j = V_i - z_{ij} I_{ij}$ which is (1).

Conversely, suppose $h_\theta(\hat{y}) \in \mathbb{X}$. From (1) and (2), we have $V_i V_j^* = |V_i|^2 - z_{ij}^* S_{ij}$. Then $\theta_i - \theta_j = \beta_{ij} + 2\pi k_{ij}$ for some integer $k_{ij} = k_{ij}(\theta)$. Hence (θ, k) solves (35).

The discussion preceding the lemma shows that, given any $k \in \mathbb{N}^m$, there is at most one θ that satisfies (35). If no such θ exists for any $k \in \mathbb{N}^m$, then (35) has no solution (θ, k). If (35) has a solution (θ, 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 $(\tilde{\theta}, \tilde{k}) \notin \sigma(\theta, k)$ with $B\tilde{\theta} = \beta + 2\pi\tilde{k}$. Then $B(\tilde{\theta} - \theta) = 2\pi(\tilde{k} - k)$, or $\tilde{k} = k + B\alpha$ for some α. Since $\tilde{k} \in \mathbb{N}^m$, α is an integer vector; moreover $\tilde{\theta}$ is unique given $\tilde{k}$. This means $(\tilde{\theta}, \tilde{k}) \in \sigma(\theta, k)$, a contradiction. ∎

Proof of Theorem 2: Since $m \geq n$ and $\mathrm{rank}(B) = n$, 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 is such a basis so that B and β are partitioned as in (26) and (28), respectively. Then Lemma 3 implies that $h_{\theta_*}(\hat{y}) \in \mathbb{X}$ with $\theta_* \in (-\pi, \pi]^n$ if and only if $(\theta_*, k_*(\theta_*))$ is the unique solution of

$$\begin{bmatrix} B_T \\ B_\perp \end{bmatrix} \theta = \begin{bmatrix} \beta_T \\ \beta_\perp \end{bmatrix} + 2\pi \begin{bmatrix} k_T \\ k_\perp \end{bmatrix}. \tag{37}$$

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 $B_\perp B_T^{-1}(\beta_T + 2\pi k_T) = \beta_\perp + 2\pi k_\perp$, i.e., $B_\perp B_T^{-1} \beta_T = \beta_\perp + 2\pi \hat{k}_\perp$ where $\hat{k}_\perp := k_\perp - B_\perp B_T^{-1} k_T$. Then (38) below implies that $\hat{k}_\perp$ 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 $T(i \rightsquigarrow j)$ denote the unique path from node i to node j in T; in particular, $T(0 \rightsquigarrow j)$ consists of links all with the same orientation as the path and $T(j \rightsquigarrow 0)$ of links all with the opposite orientation. Then it can be verified directly that

$$\left[B_T^{-1}\right]_{ei} := \begin{cases} -1 & \text{if link } e \text{ is in } T(0 \rightsquigarrow i) \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

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

$$\left[B_T^{-1}\beta_T\right]_i = \sum_e \left[B_T^{-1}\right]_{ie} [\beta_T]_e = -\sum_{e \in T(0 \rightsquigarrow i)} [\beta_T]_e.$$

Hence $B_\perp B_T^{-1}\beta_T$ 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 $e := (i,j) \in E \setminus E_T$:

$$\begin{aligned} \left[B_\perp B_T^{-1}\beta_T\right]_e &= \left[B_T^{-1}\beta_T\right]_i - \left[B_T^{-1}\beta_T\right]_j \\ &= \sum_{e' \in T(0 \rightsquigarrow j)} [\beta_T]_{e'} - \sum_{e' \in T(0 \rightsquigarrow i)} [\beta_T]_{e'}. \end{aligned}$$

Since

$$\sum_{e' \in T(0 \rightsquigarrow j)} [\beta_T]_{e'} = -\sum_{e' \in T(j \rightsquigarrow 0)} [\tilde{\beta}_T]_{e'}$$

the angle recovery condition (33) is equivalent to

$$\begin{aligned} \sum_{e' \in T(0 \rightsquigarrow i)} [\beta_T]_{e'} + [\beta_\perp]_{ij} + \sum_{e' \in T(j \rightsquigarrow 0)} [\tilde{\beta}_T]_{e'} \\ = \sum_{e' \in c(i,j)} \tilde{\beta}_{e'} = 0 \ (\mathrm{mod}\ 2\pi) \end{aligned}$$

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 (θ_*, 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 $\theta_* - 2\pi B_T^{-1}[k_*]_T = \beta_T$. Consider $\alpha := -B_T^{-1}[k_*]_T$ which is in $\mathbb{N}^n$ due to (38). Then $(\theta_* + 2\pi\alpha, k_* + B\alpha) \in \sigma(\theta_*, k_*)$ and hence is also a solution of (37) by Lemma 3. Moreover $\theta_* + 2\pi\alpha = B_T^{-1}\beta_T$ since $[k_*]_T + B_T\alpha = 0$. This means that θ_* 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 $\hat{y} \in \hat{\mathbb{Y}}$:

1) Choose *any* n basis rows of B and form B_T, $B_\perp$.
2) Compute β from $\hat{y}$ and check if $B_\perp B_T^{-1}\beta_T - \beta_\perp = 0 \ (\mathrm{mod}\ 2\pi)$.
3) If not, then $\hat{y} \notin \hat{h}(\mathbb{X})$; stop.
4) Otherwise, compute $\theta_* = \mathcal{P}\left(B_T^{-1}\beta_T\right)$.
5) Compute $h_{\theta_*}(\hat{y}) \in \mathbb{X}$ through (29)–(32).

Theorem 2 guarantees that $h_{\theta_*}(\hat{y})$, 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 $\angle I_{ij}$, $\angle V_i$, and a branch flow solution. This procedure is more amenable to a distributed implementation.

Algorithm 2: distributed angle recovery. Given a relaxed solution $\hat{y} \in \hat{\mathbb{Y}}$:

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

$$\angle I_{jk} := \angle V_j - \angle S_{jk} \tag{39}$$
$$\angle V_k := \angle V_j - \angle(v_j - z_{jk}^* S_{jk}). \tag{40}$$

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 $\angle V_k$ 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 $\theta_k^j \angle V_k \neq 0 \ (\mathrm{mod}\ 2\pi)$, then angle recovery has failed; stop.

If the angle recovery procedure succeeds in Step 3, then $\hat{y}$ together with these angles $\angle V_k$, $\angle I_{jk}$ 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 $\hat{\mathbb{Y}} \setminus \hat{h}(\mathbb{X})$ 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 $\hat{y}$, $\theta_* := \mathcal{P}(B^{-1}\beta)$ always exists and is the unique vector in $(-\pi, \pi]^n$ 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 $\theta_* = \mathcal{P}(B^{-1}\beta) \in (\pi, \pi]^n$ 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 *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

[1] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification," in *Proc. 51st IEEE Conf. Decision and Control*, Dec. 2012.

[2] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin de la Societe Francoise des Electriciens*, vol. 3, no. 8, pp. 431–447, 1962, in French.

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

[4] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 762–770, May 1991.

[5] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. Part I: Nonlinear and quadratic programming approaches," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 96–104, Feb. 1999.

[6] J. A. Momoh, M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. Part II: Newton, linear programming and interior point methods," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 105–111, Feb. 1999.

[7] K. S. Pandya and S. K. Joshi, "A survey of optimal power flow methods," *J. Theoret. Appl. Inf. Technol.*, vol. 4, no. 5, pp. 450–458, 2008.

[8] B. Stott and O. Alsaç, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 859–869, 1974.

[9] O. Alsaç, J. Bright, M. Prais, and B. Stott, "Further developments in LP-based optimal power flow," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 697–711, Aug. 1990.

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

[11] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[12] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Elect. Power Energy Syst.*, vol. 30, no. 6–7, pp. 383–392, 2008.

[13] X. Bai and H. Wei, "Semi-definite programming-based method for security-constrained unit commitment with operational and optimal power flow constraints," *IET Gen., Transm., Distrib.*, vol. 3, no. 2, pp. 182–197, 2009.

[14] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[15] J. Lavaei, "Zero duality gap for classical OPF problem convexifies fundamental nonlinear power problems," in *Proc. Amer. Control Conf.*, 2011.

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

[17] S. Bose, D. Gayme, S. H. Low, and K. M. Chandy, "Optimal power flow over tree networks," in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2011.

[18] B. Zhang and D. Tse, "Geometry of feasible injection region of power networks," in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2011.

[19] S. Bose, D. Gayme, S. H. Low, and K. M. Chandy, Quadratically Constrained Quadratic Programs on Acyclic Graphs With Application to Power Flow, Mar. 2012, arXiv:1203.5599v1.

[20] 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 Conf.*, 2011.

[21] I. A. Hiskens and R. Davy, "Exploring the power flow solution space boundary," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 389–395, Aug. 2001.

[22] B. C. Lesieutre and I. A. Hiskens, "Convexity of the set of feasible injections and revenue adequacy in FTR markets," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1790–1798, Nov. 2005.

[23] Y. V. Makarov, Z. Y. Dong, and D. J. Hill, "On convexity of power flow feasibility boundary," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 811–813, May 2008.

[24] D. T. Phan, "Lagrangian duality and branch-and-bound algorithms for optimal power flow," *Oper. Res.*, vol. 60, no. 2, pp. 275–285, Mar./Apr. 2012.

[25] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan. 1989.

[26] M. E. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 735–743, Jan. 1989.

[27] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy, "Inverter var control for distribution systems with renewables," in *Proc. IEEE SmartGridComm Conf.*, Oct. 2011.

[28] M. Farivar, R. Neal, C. Clarke, and S. H. Low, "Optimal inverter VAR control in distribution systems with high pv penetration," in *Proc. IEEE Power & Energy Society (PES) General Meeting*, Jul. 2012.

[29] J. A. Taylor, "Conic optimization of electric power systems," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, Jun. 2011.

[30] J. A. Taylor and F. S. Hover, "Convex models of distribution system reconfiguration," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1407–1413, Aug. 2012.

[31] R. Cespedes, "New method for the analysis of distribution networks," *IEEE Trans. Power Del.*, vol. 5, no. 1, pp. 391–396, Jan. 1990.

[32] A. G. Expósito and E. R. Ramos, "Reliable load flow technique for radial distribution networks," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1063–1069, Aug. 1999.

[33] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1458–1459, Aug. 2006.

[34] R. A. Jabr, "A conic quadratic format for the load flow equations of meshed networks," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2285–2286, Nov. 2007.

[35] L. Gan, N. Li, U. Topcu, and S. Low, "Branch flow model for radial networks: Convex relaxation," in *Proc. 51st IEEE Conf. Decision and Control*, Dec. 2012.

[36] N. Li, L. Chen, and S. Low, "Exact convex relaxation for radial networks using branch flow models," in *Proc. IEEE Int. Conf. Smart Grid Communications*, Nov. 2012.

[37] S. Bose, S. H. Low, and M. Chandy, "Equivalence of branch flow and bus injection models," in *Proc. 50th Annu. Allerton Conf. Communication, Control, and Computing*, Oct. 2012.

[38] M. H. Nazari and M. Parniani, “Determining and optimizing power loss reduction in distribution feeders due to distributed generation,” in *Proc. 2006 IEEE PES Power Systems Conf. Expo. (PSCE’06)* , 2006, pp. 1914–1918.

[39] M. H. Nazari and M. Illic, “Potential for efficiency improvement of future electric energy systems with distributed generation units,” in *Proc. 2010 IEEE Power and Energy Society General Meeting*, 2010, pp. 1–9.

[40] H.-D. Chiang and M. E. Baran, “On the existence and uniqueness of load flow solution for radial distribution power networks,” *IEEE Trans. Circuits Syst.*, vol. 37, no. 3, pp. 410–416, Mar. 1990.

[41] H.-D. Chiang, “A decoupled load flow method for distribution power networks: Algorithms, analysis and convergence study,” *Int. J. Elect. Power Energy Syst.*, vol. 13, no. 3, pp. 130–138, Jun. 1991.

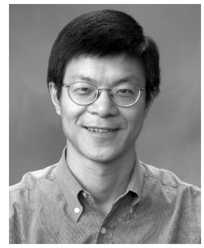
[42] M. Farivar and S. H. Low, “Branch flow model: Relaxations and convexification—Part II,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2565–2572, Aug. 2013.

[43] L. R. Foulds, *Graph Theory Applications*. New York, NY, USA: Springer-Verlag, 1992.

[44] N. Biggs, *Algebraic Graph Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1993, Cambridge Mathematical Library.

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*.