Exact Convex Relaxation of Optimal Power Flow in Radial Networks

Lingwen Gan, Na Li, Ufuk Topcu, and Steven H. Low

***Abstract*—The optimal power flow (OPF) problem determines a network operating point that minimizes a certain objective such as generation cost or power loss. It is nonconvex. We prove that a global optimum of OPF can be obtained by solving a second-order cone program, under a mild condition after shrinking the OPF feasible set slightly, for radial power networks. The condition can be checked a priori, and holds for the IEEE 13, 34, 37, 123-bus networks and two real-world networks.**

I. INTRODUCTION

The optimal power flow (OPF) problem determines a network operating point that minimizes a certain objective such as generation cost or power loss. It has been one of the fundamental problems in power system operation since 1962. As distributed generation (e.g., photovoltaic panels) and controllable loads (e.g., electric vehicles) proliferate, OPF problems for distribution networks become increasingly important. To use controllable loads to integrate volatile renewable generation, solving the OPF problem in real-time will be inevitable. Power distribution networks are usually radial (have a tree topology).

The OPF problem is difficult because power flow is governed by nonlinear Kirchhoff's laws. There are three ways to deal with this challenge: 1) approximate the power flow equations; 2) look for a local optimum of the OPF problem; and 3) convexify the constraints imposed by the Kirchhoff's laws. After a brief discussion of the first two approaches, we will focus on the third. See extensive surveys in e.g. [1]–[12].

Power flow equations can be approximated by linear equations known as the DC power flow equations [13]–[15] if 1) power losses on the lines are small; 2) voltages are close to their nominal values; 3) voltage angle differences between adjacent buses are small. With the DC approximation, the OPF problem, called DC OPF, reduces to a linear program. For transmission networks, these three assumptions are reasonable and DC OPF is widely used in practice. It, however, has three limitations. First, it is not applicable for applications such as power routing (e.g. [16]) and volt/var control (e.g. [17]) since it assumes fixed voltage magnitudes and ignores reactive powers. Second, a solution of the DC approximation may not be feasible (may not satisfy the nonlinear power flow equations). In this case an operator typically tightens some constraints in DC OPF and solves again. This may not only reduce efficiency but also relies on heuristics that are hard to scale to larger systems or faster control in the future. Finally, DC approximation is unsuitable for distribution systems where loss is much higher than in transmission systems, voltages can fluctuate significantly, and reactive powers are used to stabilize voltages [17]. See [18] for a more accurate power flow linearization that addresses these shortcomings of the DC approximation.

Many nonlinear algorithms that seek a local optimum of the OPF problem have also been developed to avoid these shortcomings. Representative algorithms include successive linear/quadratic programming [19], trust-region based methods [20], [21], Lagrangian Newton method [22], and interior-point methods [23]–[25]. Some of them, especially those based on Newton-Ralphson, are quite successful empirically. However, when they converge, these algorithms converge to a local minimum without assurance on the suboptimality gap.

In this paper we focus on the convexification approach (see [26], [27] for a tutorial). Solving OPF through semidefinite relaxation is first proposed in [28] as a second-order cone program (SOCP) for radial networks and in [29] as a semidefinite program (SDP) for general networks in a bus injection model. It is first proposed in [30], [31] as an SOCP for radial networks in the branch flow model of [32], [33]. While these convex relaxations have been illustrated numerically in [28] and [29], whether or when they are exact is first studied in [34] (i.e., when an optimal solution of the original OPF problem can be recovered from every optimal solution of an SDP relaxation). Exploiting graph sparsity to simplify the SDP relaxation of OPF is first proposed in [35], [36] and analyzed in [37], [38]. These relaxations are equivalent for radial networks in the sense that there is a bijective map between their feasible sets [38]. The SOCP relaxation, however, has a much lower computational complexity. We will hence focus on the SOCP relaxation in this paper.

Solving OPF through convex relaxation offers several advantages. It provides the ability to check if a solution is globally optimal. If it is not, the solution provides a lower bound on the minimum cost and hence a bound on how far any feasible solution is from optimality. Unlike approximations, if a relaxation is infeasible, it is a certificate that the original OPF is infeasible.

Convex relaxations may not be exact [39]–[41]. For radial networks, three types of sufficient conditions have been developed in the literature that guarantee their exactness. They are

This work was supported by NSF NetSE grant CNS 0911041, ARPA-E grant DE-AR0000226, Southern California Edison, National Science Council of Taiwan, R.O.C, grant NSC 103-3113-P-008-001, Los Alamos National Lab through a DoE grant, Resnick Institute, and AFOSR award number FA9550-12-1-0302.

Lingwen Gan, and Steven H. Low are with the Engineering and Applied Science Division, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: lgan@caltech.edu). Na Li is with the Department of Electrical Engineering, Harvard University, Boston, MA 02138 USA. Ufuk Topcu is with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA.

not necessary in general and have implications on allowable power injections, voltage magnitudes, or voltage angles:

A *Power injections:* These conditions require that not both constraints on real and reactive power injections be binding at both ends of a line [30], [31], [42]–[47].
B *Voltage angles:* These conditions require that the voltage angles across each line be sufficiently close [48], [49]. This is needed also for stability reasons.
C *Voltages magnitudes:* These conditions require that the upper bounds on voltage magnitudes not be binding [50]–[52]. They can be enforced through affine constraints on power injections. This paper generalizes these results.

Summary of contributions: The goal of this paper is to show that after modifying the OPF problem for radial networks slightly, the corresponding SOCP relaxation is exact under a mild condition that can be checked a priori. In particular, contributions of this paper are threefold.

First, we prove in Theorem 1 that *if voltage upper bounds do not bind at optimality, then the SOCP relaxation is exact under a mild condition*. The condition can be checked a priori and holds for the IEEE 13, 34, 37, 123-bus networks and two real-world networks. The condition has a physical interpretation that all upstream reverse power flows increase if the power loss on a line is reduced.

Second, in Section IV we *modify the OPF problem by limiting power injections to a region where voltage upper bounds do not bind* so that the SOCP relaxation is exact under the aforementioned condition. We illustrate that this only eliminates power injections from the original feasible set that are close to voltage upper bounds. Examples exist where the SOCP relaxation is not exact without this modification.

Third, we prove in Theorem 4 that *the result in this paper unifies and generalizes the results in [50], [51].*

The rest of this paper is organized as follows. The OPF problem and the SOCP relaxation are introduced in Section II, and a sufficient condition for exactness is provided in Section III. The condition consists of two parts, C1 and C2. Since C2 cannot be checked a priori, we propose in Section IV a modified OPF problem that always satisfies C2 and therefore its SOCP relaxation is exact under C1. We compare C1 with prior works in Section V and show that C1 holds with large margin for a number of test networks in Section VI.

II. The optimal power flow problem

A. Power flow model

A distribution network is composed of buses and lines connecting these buses, and is usually radial. The root of the network is a *substation bus* that connects to the transmission network. It has a fixed voltage and redistributes the bulk power it receives from the transmission network to other buses. Index the substation bus by 0 and the other buses by $1, \ldots, n$. Let $\mathcal{N} := \{0, \ldots, n\}$ denote the collection of all buses and define $\mathcal{N}^+ := \mathcal{N} \backslash \{0\}$. Each line connects an ordered pair (i, j) of buses where bus j lies on the unique path from bus i to bus 0. Let $\mathcal{E}$ denote the collection of all lines, and abbreviate $(i, j) \in \mathcal{E}$ by $i \to j$ whenever convenient.

For each bus $i \in \mathcal{N}$, let v_i denote the square of the magnitude of its complex voltage, e.g., if the voltage is $1.05\angle 120°$ per unit, then $v_i = 1.05^2$. The substation voltage v_0 is fixed and given. Let $s_i = p_i + \mathbf{i}q_i$ denote the power injection of bus i where p_i and q_i denote the real and reactive power injections respectively. Let $\mathcal{P}_i$ denote the unique path from bus i to bus 0. Since the network is radial, the path $\mathcal{P}_i$ is well-defined. For each line $(i, j) \in \mathcal{E}$, let $z_{ij} = r_{ij} + \mathbf{i}x_{ij}$ denote its impedance. Let ℓ_{ij} denote the square of the magnitude of the complex current from bus i to bus j, e.g., if the current is $0.5\angle 10°$, then $\ell_{ij} = 0.5^2$. Let $S_{ij} = P_{ij} + \mathbf{i}Q_{ij}$ denote the sending-end power flow from bus i to bus j where P_{ij} and Q_{ij} denote the real and reactive power flow respectively. Some of the notations are summarized in Fig. 1. We use a letter without subscripts to denote a vector of the corresponding quantities, e.g., $v = (v_i)_{i \in \mathcal{N}^+}$, $\ell = (\ell_{ij})_{(i,j) \in \mathcal{E}}$. Note that subscript 0 is not included in nodal quantities such as v and s. For a complex number $a \in \mathbb{C}$, let $\bar{a}$ denote the conjugate of a.

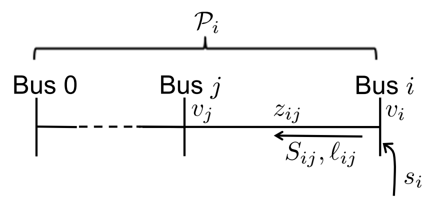

Fig. 1. Some of the notations.

Given the network $(\mathcal{N}, \mathcal{E})$, the impedance z, and the substation voltage v_0, the other variables (s, S, v, ℓ, s_0) are described by the *branch flow model* for radial networks [32], [33]:

$$S_{ij} = s_i + \sum_{h:\, h \to i} (S_{hi} - z_{hi}\ell_{hi}), \qquad \forall (i, j) \in \mathcal{E}; \tag{1a}$$

$$0 = s_0 + \sum_{h:\, h \to 0} (S_{h0} - z_{h0}\ell_{h0}); \tag{1b}$$

$$v_i - v_j = 2\mathrm{Re}(\bar{z}_{ij} S_{ij}) - |z_{ij}|^2 \ell_{ij}, \qquad \forall (i, j) \in \mathcal{E}; \tag{1c}$$

$$\ell_{ij} = \frac{|S_{ij}|^2}{v_i}, \qquad \forall (i, j) \in \mathcal{E} \tag{1d}$$

B. The OPF problem

We consider the following controllable devices in a distribution network: distributed generators, inverters, controllable loads such as electric vehicles and smart appliances, and shunt capacitors. For application examples, in volt/var control, reactive power injection of inverters and shunt capacitors are controlled to regulate voltages; in demand response, real power consumption of controllable loads is reduced or shifted. Mathematically, power injection s is the control variable, after specifying which the other variables (S, v, ℓ, s_0) are determined by the power flow laws in (1).

The power injection s_i of a bus $i \in \mathcal{N}^+$ is constrained to be in a pre-specified set $\mathcal{S}_i$, i.e.,

$$s_i \in \mathcal{S}_i, \qquad i \in \mathcal{N}^+. \tag{2}$$

The set $\mathcal{S}_i$ for some controllable devices are:

- If s_i represents a shunt capacitor with nameplate capacity $\overline{q}_i$, then $\mathcal{S}_i = \{s \in \mathbb{C} \mid \mathrm{Re}(s) = 0,\ \mathrm{Im}(s) = 0 \text{ or } \overline{q}_i\}$. Note that $\mathcal{S}_i$ is nonconvex and disconnected in this case.

- If s_i represents a solar panel with generation capacity $\overline{p}_i$, that is connected to the grid through an inverter with nameplate capacity $\overline{s}_i$, then $\mathcal{S}_i = \{s \in \mathbb{C} \mid 0 \le \text{Re}(s) \le \overline{p}_i,\ |s| \le \overline{s}_i\}$.
- If s_i represents a controllable load with constant power factor η, whose real power consumption can vary continuously from $-\overline{p}_i$ to $-\underline{p}_i$ (here $\underline{p}_i \le \overline{p}_i \le 0$), then $\mathcal{S}_i = \left\{s \in \mathbb{C} \mid \underline{p}_i \le \text{Re}(s) \le \overline{p}_i,\ \text{Im}(s) = \sqrt{1-\eta^2}\text{Re}(s)/\eta\right\}$.

Note that s_i can represent the aggregate power injection of multiple such devices with an appropriate $\mathcal{S}_i$, and that the set $\mathcal{S}_i$ is not necessarily convex or connected.

An important goal of control is to regulate the voltages to lie within pre-specified lower and upper bounds $\underline{v}_i$ and $\overline{v}_i$, i.e.,

$$\underline{v}_i \le v_i \le \overline{v}_i, \qquad i \in \mathcal{N}^+. \tag{3}$$

For example, if voltages must not deviate by more than 5% from their nominal values, then $0.95^2 \le v_i \le 1.05^2$ per unit. We consider the control objective

$$C(s, s_0) = \sum_{i \in \mathcal{N}} f_i(\text{Re}(s_i)) \tag{4}$$

where $f_i : \mathbb{R} \to \mathbb{R}$ denotes the generation cost at bus i for $i \in \mathcal{N}$. If $f_i(x) = x$ for $i \in \mathcal{N}$, then C is the total power loss on the network.

The OPF problem seeks to minimize the generation cost (4), subject to power flow constraints (1), power injection constraints (2), and voltage constraints (3):

$$\textbf{OPF:} \quad \min \sum_{i \in \mathcal{N}} f_i(\text{Re}(s_i))$$
$$\text{over} \quad s, S, v, \ell, s_0$$
$$\text{s.t.} \quad S_{ij} = s_i + \sum_{h:\, h \to i} (S_{hi} - z_{hi}\ell_{hi}), \quad \forall (i,j) \in \mathcal{E}; \tag{5a}$$
$$0 = s_0 + \sum_{h:\, h \to 0} (S_{h0} - z_{h0}\ell_{h0}); \tag{5b}$$
$$v_i - v_j = 2\text{Re}(\bar{z}_{ij}S_{ij}) - |z_{ij}|^2\ell_{ij}, \quad \forall (i,j) \in \mathcal{E}; \tag{5c}$$
$$\ell_{ij} = \frac{|S_{ij}|^2}{v_i}, \quad \forall (i,j) \in \mathcal{E}; \tag{5d}$$
$$s_i \in \mathcal{S}_i, \quad i \in \mathcal{N}^+; \tag{5e}$$
$$\underline{v}_i \le v_i \le \overline{v}_i, \quad i \in \mathcal{N}^+. \tag{5f}$$

The following assumptions are made throughout this paper.

A1 The network $(\mathcal{N}, \mathcal{E})$ is a tree. Distribution networks are usually radial.

A2 The substation voltage v_0 is fixed and given. In practice, v_0 can be modified several times a day, and therefore can be considered as a given constant at the timescale of OPF.

A3 Line resistances and reactances are strictly positive, i.e., $r_{ij} > 0$ and $x_{ij} > 0$ for $(i,j) \in \mathcal{E}$. This holds in practice because lines are passive (consume power) and inductive.

A4 Voltage lower bounds are strictly positive, i.e., $\underline{v}_i > 0$ for $i \in \mathcal{N}^+$. In practice, $\underline{v}_i$ is slightly below 1 per unit.

The equality constraint (5d) is nonconvex, and one can relax it to inequality constraints to obtain the following second-order cone programming (SOCP) relaxation [31]:

$$\textbf{SOCP:} \quad \min \sum_{i \in \mathcal{N}} f_i(\text{Re}(s_i))$$
$$\text{over} \quad s, S, v, \ell, s_0$$
$$\text{s.t.} \quad \text{(5a)} - \text{(5c)}, \ \text{(5e)} - \text{(5f)};$$
$$\ell_{ij} \ge \frac{|S_{ij}|^2}{v_i}, \quad \forall (i,j) \in \mathcal{E}. \tag{6}$$

Note that SOCP is not necessarily convex, since we allow f_i to be nonconvex and $\mathcal{S}_i$ to be nonconvex. Nonetheless, we call it SOCP for brevity.

If an optimal SOCP solution $w = (s, S, v, \ell, s_0)$ is feasible for OPF, i.e., w satisfies (5d), then w is a global optimum of OPF. This motivates the following definition.

Definition 1. *SOCP is exact if every of its optimal solutions satisfies* (5d).

III. A SUFFICIENT CONDITION

We now provide a sufficient condition that ensures SOCP is exact. It motivates a modified OPF problem in Section IV.

A. Statement of the condition

We start with introducing the notations that will be used in the statement of the condition. One can ignore the ℓ terms in (1a) and (1c) to obtain the *Linear DistFlow Model* [32], [33]:

$$S_{ij} = s_i + \sum_{h:\, h \to i} S_{hi}, \qquad \forall (i,j) \in \mathcal{E};$$
$$v_i - v_j = 2\text{Re}(\bar{z}_{ij}S_{ij}), \qquad \forall (i,j) \in \mathcal{E}.$$

Let $(\hat{S}, \hat{v})$ denote the solution of the Linear DistFlow model, then

$$\hat{S}_{ij}(s) = \sum_{h:\, i \in \mathcal{P}_h} s_h, \qquad \forall (i,j) \in \mathcal{E};$$
$$\hat{v}_i(s) := v_0 + 2 \sum_{(j,k) \in \mathcal{P}_i} \text{Re}\left(\bar{z}_{jk}\hat{S}_{jk}(s)\right), \qquad \forall i \in \mathcal{N}$$

as in Fig. 2. Physically, $\hat{S}_{ij}(s)$ denotes the sum of power injections s_h towards bus 0 that go through line (i,j). Note that $(\hat{S}(s), \hat{v}(s))$ is affine in s, and equals (S, v) if and only if line loss $z_{ij}\ell_{ij}$ is 0 for $(i,j) \in \mathcal{E}$. For two complex numbers

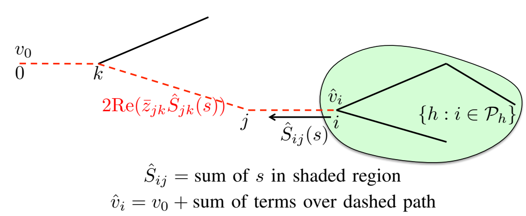

Fig. 2. Illustration of $\hat{S}_{ij}$ and $\hat{v}_i$. The shaded region is downstream of bus i, and contains the buses $\{h : i \in \mathcal{P}_h\}$. Quantity $\hat{S}_{ij}(s)$ is defined to be the sum of bus injections in the shaded region. The dashed lines constitute the path $\mathcal{P}_i$ from bus i to bus 0. Quantity $\hat{v}_i(s)$ is defined as v_0 plus the terms $2\text{Re}(\bar{z}_{jk}\hat{S}_{jk}(s))$ over the dashed path.

$a, b \in \mathbb{C}$, let $a \le b$ denote $\text{Re}(a) \le \text{Re}(b)$ and $\text{Im}(a) \le \text{Im}(b)$. For two vectors a, b of the same dimension, let $a \le b$ denote componentwise inequality. Define $<$, $>$, and $\ge$ similarly.

Lemma 1. *If (s, S, v, ℓ, s_0) satisfies (1a)–(1c) and $\ell \geq 0$ componentwise, then $S \leq \hat{S}(s)$ and $v \leq \hat{v}(s)$.*

Lemma 1 implies that $\hat{v}(s)$ and $\hat{S}(s)$ provide upper bounds on v and S. It is proved in Appendix A. Let $\hat{P}(s)$ and $\hat{Q}(s)$ denote the real and imaginary parts of $\hat{S}(s)$ respectively. Then

$$\hat{P}_{ij}(s = p + \mathbf{i}q) = \hat{P}_{ij}(p) = \sum_{h:\, i \in \mathcal{P}_h} p_h, \qquad (i,j) \in \mathcal{E};$$
$$\hat{Q}_{ij}(s = p + \mathbf{i}q) = \hat{Q}_{ij}(q) = \sum_{h:\, i \in \mathcal{P}_h} q_h, \qquad (i,j) \in \mathcal{E}.$$

Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that

$$\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \leq \overline{p}_i,\ \mathrm{Im}(s) \leq \overline{q}_i\}$$

for $i \in \mathcal{N}^+$ as in Fig. 3, i.e., $\mathrm{Re}(s_i)$ and $\mathrm{Im}(s_i)$ are upper bounded by $\overline{p}_i$ and $\overline{q}_i$ respectively. Define $a^+ := \max\{a, 0\}$

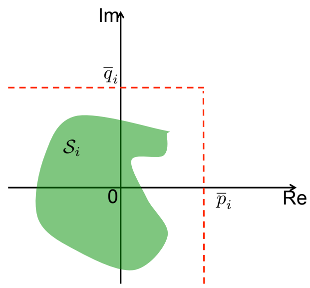

Fig. 3. We assume that $\mathcal{S}_i$ lies in the left bottom corner of $(\overline{p}_i, \overline{q}_i)$, but do not assume that $\mathcal{S}_i$ is convex or connected.

for $a \in \mathbb{R}$. Let $I := \mathrm{diag}(1,1)$ denote the 2×2 identity matrix, and define

$$u_{ij} := \begin{pmatrix} r_{ij} \\ x_{ij} \end{pmatrix}, \qquad \underline{A}_{ij} := I - \frac{2}{\underline{v}_i} \begin{pmatrix} r_{ij} \\ x_{ij} \end{pmatrix} \begin{pmatrix} \hat{P}^+_{ij}(\overline{p}) & \hat{Q}^+_{ij}(\overline{q}) \end{pmatrix}$$

for $(i,j) \in \mathcal{E}$. For each $i \in \mathcal{N}^+$, $(i, j_1) \in \mathcal{E}$ and $(i, j_2) \in \mathcal{E}$ implies $j_1 = j_2$, and therefore we can abbreviate u_{ij} and $\underline{A}_{ij}$ by u_i and $\underline{A}_i$ respectively without ambiguity.

Further, let $\mathcal{L} := \{l \in \mathcal{N} \mid \nexists k \in \mathcal{N} \text{ such that } k \to l\}$ denote the collection of leaf buses in the network. For a leaf bus $l \in \mathcal{L}$, let $n_l + 1$ denote the number of buses on path $\mathcal{P}_l$, and suppose

$$\mathcal{P}_l = \{l_{n_l} \to l_{n_l - 1} \to \ldots \to l_1 \to l_0\}$$

with $l_{n_l} = l$ and $l_0 = 0$ as in Fig. 4. Let

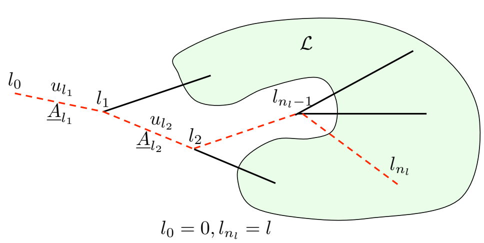

Fig. 4. The shaded region denotes the collection $\mathcal{L}$ of leaf buses, and the path $\mathcal{P}_l$ of a leaf bus $l \in \mathcal{L}$ is illustrated by a dashed line.

$$\mathcal{S}_{\mathrm{volt}} := \{s \in \mathbb{C}^n \mid \hat{v}_i(s) \leq \overline{v}_i \text{ for } i \in \mathcal{N}^+\}$$

denote the power injection region where $\hat{v}(s)$ is upper bounded by $\overline{v}$. Since $v \leq \hat{v}(s)$ (Lemma 1), the set $\mathcal{S}_{\mathrm{volt}}$ is a power injection region where voltage upper bounds do not bind.

The following theorem provides a sufficient condition that guarantees the exactness of SOCP.

Theorem 1. *Assume that f_0 is strictly increasing, and that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \leq \overline{p}_i,\ \mathrm{Im}(s) \leq \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then SOCP is exact if the following conditions hold:*

- C1 *$\underline{A}_{l_s} \underline{A}_{l_{s+1}} \cdots \underline{A}_{l_{t-1}} u_{l_t} > 0$ for any $l \in \mathcal{L}$ and any s, t such that $1 \leq s \leq t \leq n_l$;*
- C2 *every optimal SOCP solution $w = (s, S, v, \ell, s_0)$ satisfies $s \in \mathcal{S}_{\mathrm{volt}}$.*

Theorem 1 implies that if C2 holds, i.e., optimal power injections lie in the region $\mathcal{S}_{\mathrm{volt}}$ where voltage upper bounds do not bind, then SOCP is exact under C1. C2 depends on SOCP solutions and cannot be checked a priori. This drawback motivates us to modify OPF such that C2 always holds and therefore the corresponding SOCP is exact under C1, as will be discussed in Section IV.

We illustrate the proof idea of Theorem 1 via a 3-bus linear network in Fig. 5. The proof for general radial networks is

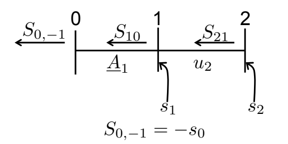

Fig. 5. A 3-bus linear network.

provided in Appendix B. Assume C1 and C2 hold. If SOCP is not exact, then there exists an optimal SOCP solution $w = (s, S, v, \ell, s_0)$ that violates (5d). We will construct another feasible point $w' = (s', S', v', \ell', s'_0)$ of SOCP that has a smaller objective value than w, contradicting the optimality of w and implying SOCP is exact.

There are two ways (5d) gets violated: 1) (5d) is violated on line $(1,0)$; or 2) (5d) is satisfied on line $(1,0)$ but violated on line $(2,1)$. To illustrate the proof idea, we focus on the second case, i.e., the case where $\ell_{10} = |S_{10}|^2/v_1$ and $\ell_{21} > |S_{21}|^2/v_2$. In this case, the construction of w' is

$$\begin{aligned}
&\textbf{Initialization:} && s' = s,\ S'_{21} = S_{21};\\
&\textbf{Forward sweep:} && \ell'_{21} = |S'_{21}|^2/v_2,\\
& && S'_{10} = S'_{21} - z_{21}\ell'_{21} + s'_1;\\
& && \ell'_{10} = |S'_{10}|^2/v_1,\\
& && S'_{0,-1} = S'_{10} - z_{10}\ell'_{10};\\
&\textbf{Backward sweep:} && v'_1 = v_0 + 2\mathrm{Re}(\bar{z}_{10}S'_{10}) - |z_{10}|^2\ell'_{10};\\
& && v'_2 = v'_1 + 2\mathrm{Re}(\bar{z}_{21}S'_{21}) - |z_{21}|^2\ell'_{21}
\end{aligned}$$

where $S'_{0,-1} = -s'_0$. The construction consists of three steps:

- S1 In the initialization step, s' and S'_{21} are initialized as the corresponding values in w.
- S2 In the forward sweep step, $\ell'_{k,k-1}$ and $S'_{k-1,k-2}$ are recursively constructed for $k = 2, 1$ by alternatively applying (5d) (with v' replaced by v) and (5a)/(5b).

This recursive construction updates ℓ' and S' alternatively along the path $\mathcal{P}_2$ from bus 2 to bus 0, and is therefore called a *forward sweep*.

S3 In the backward sweep step, v_k' is recursively constructed for $k = 1, 2$ by applying (5c). This recursive construction updates v' along the negative direction of $\mathcal{P}_2$ from bus 0 to bus 2, and is therefore called a *backward sweep*.

One can show that w' is feasible for SOCP and has a smaller objective value than w. This contradicts the optimality of w, and therefore SOCP is exact.

Remark 1. *Theorem 1 still holds if there is an additional power injection constraint $s \in \mathcal{S}$ in OPF, where $\mathcal{S}$ can be an arbitrary set. This is because we set $s' = s$ in the construction of w', and therefore $s \in \mathcal{S}$ implies $s' \in \mathcal{S}$. Hence, an additional constraint $s \in \mathcal{S}$ does not affect the fact that w' is feasible for SOCP and has a smaller objective value than w.*

B. Interpretation of C1

We illustrate C1 through a linear network as in Fig. 6. The collection of leaf buses is a singleton $\mathcal{L} = \{n\}$, and the path from the only leaf bus n to bus 0 is $\mathcal{P}_n = \{n \to n-1 \to \cdots \to 1 \to 0\}$. Then, C1 takes the form

$$\underline{A}_s \underline{A}_{s+1} \cdots \underline{A}_{t-1} u_t > 0, \qquad 1 \le s \le t \le n.$$

That is, given any network segment $(s-1, t)$ where $1 \le s \le t \le n$, the multiplication $\underline{A}_s \underline{A}_{s+1} \cdots \underline{A}_{t-1}$ of $\underline{A}$ over the segment $(s-1, t-1)$ times u_t is strictly positive.

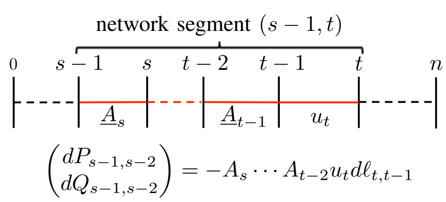

Fig. 6. In the above linear network, $\mathcal{L} = \{n\}$ and $\mathcal{P}_n = \{n \to n-1 \to \cdots \to 1 \to 0\}$. C1 requires that given any highlighted segment $(s-1, t)$ where $1 \le s \le t \le n$, the multiplication of $\underline{A}$ over $(s-1, t-1)$ times u_t is strictly positive (componentwise).

C1 only depends on SOCP parameters $(r, x, \overline{p}, \overline{q}, \underline{v})$. It can be checked a priori and efficiently since $\underline{A}$ and u are simple functions of $(r, x, \overline{p}, \overline{q}, \underline{v})$ that can be computed in $O(n)$ time and there are no more than $n(n+1)/2$ inequalities in C1.

Proposition 1. *If $(\overline{p}, \overline{q}) \le (\overline{p}', \overline{q}')$ and C1 holds for $(r, x, \overline{p}', \overline{q}', \underline{v})$, then C1 also holds for $(r, x, \overline{p}, \overline{q}, \underline{v})$.*

Proposition 1 implies that if C1 holds a set of power injections, then C1 also holds for smaller power injections. It is proved in Appendix C.

Proposition 2. *If $(\overline{p}, \overline{q}) \le 0$, then C1 holds.*

Proposition 2 implies that if every bus only consumes real and reactive power, then C1 holds. This is because when $(\overline{p}, \overline{q}) \le 0$, the quantities $\hat{P}_{ij}(\overline{p}) \le 0$, $\hat{Q}_{ij}(\overline{q}) \le 0$ for $(i,j) \in \mathcal{E}$. It follows that $\underline{A}_i = I$ for $i \in \mathcal{N}^+$. Hence, $\underline{A}_{l_s} \cdots \underline{A}_{l_{t-1}} u_{l_t} = u_{l_t} > 0$ for any $l \in \mathcal{L}$ and any s, t such that $1 \le s \le t \le n_l$.

For practical parameter ranges of $(r, x, \overline{p}, \overline{q}, \underline{v})$, line resistance and reactance $r_{ij}, x_{ij} \ll 1$ per unit for $(i,j) \in \mathcal{E}$, line flows $\hat{P}_{ij}(\overline{p}), \hat{Q}_{ij}(\overline{q})$ are on the order of 1 per unit for $(i,j) \in \mathcal{E}$, and voltage lower bound $\underline{v}_i \approx 1$ per unit for $i \in \mathcal{N}^+$. Hence, $\underline{A}_i$ is close to I for $i \in \mathcal{N}^+$, and therefore C1 is likely to hold. As will be seen in Section VI, C1 holds for several test networks, including those with big $(\overline{p}, \overline{q})$ (high penetration of distributed generation).

C1 has a physical interpretation. Recall that $S_{k,k-1}$ denotes the reverse power flow on line $(k, k-1)$ for $k = 1, \ldots, n$ and introduce $S_{0,-1} := -s_0$ for convenience. If the power loss on a line is reduced, it is natural that all upstream reverse power flows will increase. More specifically, the power loss on line $(t, t-1)$ where $t \in \{1, 2, \ldots, n\}$ is reduced if the current $\ell_{t,t-1}$ is reduced by $-d\ell_{t,t-1} > 0$. When power loss gets smaller, reverse power flow $S_{s-1,s-2}$ is likely to increase, i.e., $dS_{s-1,s-2} > 0$, for $s = 1, 2, \ldots, t$.

Let $dS_{s-1,s-2} = dP_{s-1,s-2} + \mathbf{i} dQ_{s-1,s-2} > 0$ for $s = 1, \ldots, t$. It can be verified that $(dP_{t-1,t-2} \;\; dQ_{t-1,t-2})^T = -u_t d\ell_{t,t-1}$, and one can compute from (1) the Jacobian matrix

$$\begin{aligned} A_k &:= \begin{pmatrix} \frac{\partial P_{k-1,k-2}}{\partial P_{k,k-1}} & \frac{\partial P_{k-1,k-2}}{\partial Q_{k,k-1}} \\ \frac{\partial Q_{k-1,k-2}}{\partial P_{k,k-1}} & \frac{\partial Q_{k-1,k-2}}{\partial Q_{k,k-1}} \end{pmatrix} \\ &= I - \frac{2}{v_k} \begin{pmatrix} r_{k,k-1} \\ x_{k,k-1} \end{pmatrix} \begin{pmatrix} P_{k,k-1} & Q_{k,k-1} \end{pmatrix} \end{aligned}$$

for $k = 1, \ldots, n$. Therefore

$$(dP_{s-1,s-2} \;\; dQ_{s-1,s-2})^T = -A_s A_{s+1} \cdots A_{t-1} u_t d\ell_{t,t-1}$$

for $s = 1, \ldots, t$. Then, $dS_{s-1,s-2} > 0$ implies

$$A_s A_{s+1} \cdots A_{t-1} u_t > 0 \tag{7}$$

for $s = 1, 2, \ldots, t$. Note that $\underline{A}_k$ is obtained by replacing (P, Q, v) in A_k by $(\hat{P}^+(\overline{p}), \hat{Q}^+(\overline{q}), \underline{v})$ (so that $\underline{A}_k$ only depends on SOCP parameters), and then (7) becomes C1.

IV. A MODIFIED OPF PROBLEM

The condition C2 in Theorem 1 depends on SOCP solutions and cannot be checked a priori. It can however be enforced by the additional constraint

$$s \in \mathcal{S}_{\text{volt}} \tag{8}$$

on OPF. Condition (8) is equivalent to n affine constraints on s, $\hat{v}_i(s) \le \overline{v}_i$ for $i \in \mathcal{N}^+$. Since $v_i \le \hat{v}_i(s)$ (Lemma 1), the constraints $v_i \le \overline{v}_i$ in (5f) become redundant after imposing (8). To summarize, the modified OPF problem is

$$\begin{aligned} \textbf{OPF-m:} \quad \min \quad & \sum_{i \in \mathcal{N}} f_i(\mathrm{Re}(s_i)) \\ \text{over} \quad & s, S, v, \ell, s_0 \\ \text{s.t.} \quad & \text{(5a)} - \text{(5e)}; \\ & \underline{v}_i \le v_i, \;\; \hat{v}_i(s) \le \overline{v}_i, \quad i \in \mathcal{N}^+. \end{aligned} \tag{9}$$

A modification to OPF is necessary to ensure an exact SOCP, since otherwise examples exist where SOCP is not exact. Remarkably, the feasible sets of OPF-m and OPF are similar since $\hat{v}_i(s)$ is close to v_i in practice [17], [32], [33].

One can relax (5d) to (6) to obtain the corresponding SOCP relaxation for OPF-m:

$$\begin{aligned}\textbf{SOCP-m:}\ \min\ & \sum_{i\in\mathcal{N}} f_i(\mathrm{Re}(s_i))\\ \text{over}\ & s, S, v, \ell, s_0\\ \text{s.t.}\ & (5\text{a})-(5\text{c}),\ (6),\ (5\text{e}),\ (9).\end{aligned}$$

Note again that SOCP-m is not necessarily convex, since we allow f_i and $\mathcal{S}_i$ to be nonconvex.

Since OPF-m is obtained by imposing additional constraint (8) on OPF, it follows immediately from Remark 1 that SOCP-m relaxation is exact under C1 – a mild condition that can be checked a priori.

Theorem 2. *Assume that f_0 is strictly increasing, and that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then SOCP-m is exact if C1 holds.*

The next result implies that SOCP-m has at most one optimal solution if it is convex and exact. The theorem is proved in Appendix D. The proof also implies that, under the condition of Theorem 3, the feasible set is hollow.

Theorem 3. *If f_i is convex for $i \in \mathcal{N}$, $\mathcal{S}_i$ is convex for $i \in \mathcal{N}^+$, and SOCP-m is exact, then SOCP-m has at most one optimal solution.*

V. Connection with prior results

Theorem 1 unifies and generalizes the results in [50], [51] due to Theorem 4 proved in Appendix E. Theorem 4 below says that C1 holds if at least one of the followings hold: 1) Every bus only consumes real and reactive power; 2) lines share the same resistance to reactance ratio; 3) The buses only consume real power and the resistance to reactance ratio increases as lines branch out from the substation; 4) The buses only consume reactive power and the resistance to reactance ratio decreases as lines branch out from the substation; 5) upper bounds $\hat{P}^+(\bar{p})$, $\hat{Q}^+(\bar{q})$ on reverse power flows are sufficiently small. Let

$$\mathcal{E}' := \{(i,j) \in \mathcal{E} \mid i \notin \mathcal{L}\}$$

denote the set of all non-leaf lines.

Theorem 4. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if any one of the following statements is true:*

1) *$\hat{S}_{ij}(\overline{p} + \boldsymbol{i}\overline{q}) \le 0$ for all $(i,j) \in \mathcal{E}'$.*
2) *r_{ij}/x_{ij} is identical for all $(i,j) \in \mathcal{E}$; and $\underline{v}_i - 2r_{ij}\hat{P}^+_{ij}(\overline{p}) - 2x_{ij}\hat{Q}^+_{ij}(\overline{q}) > 0$ for all $(i,j) \in \mathcal{E}'$.*
3) *$r_{ij}/x_{ij} \ge r_{jk}/x_{jk}$ whenever $(i,j),(j,k) \in \mathcal{E}$; and $\hat{P}_{ij}(\overline{p}) \le 0$, $\underline{v}_i - 2x_{ij}\hat{Q}^+_{ij}(\overline{q}) > 0$ for all $(i,j) \in \mathcal{E}'$.*
4) *$r_{ij}/x_{ij} \le r_{jk}/x_{jk}$ whenever $(i,j),(j,k) \in \mathcal{E}$; and $\hat{Q}_{ij}(\overline{q}) \le 0$, $\underline{v}_i - 2r_{ij}\hat{P}^+_{ij}(\overline{p}) > 0$ for all $(i,j) \in \mathcal{E}'$.*
5) $$\begin{bmatrix}\prod_{(k,l)\in\mathcal{P}_j} c_{kl} & -\sum_{(k,l)\in\mathcal{P}_j} d_{kl}\\ -\sum_{(k,l)\in\mathcal{P}_j} e_{kl} & \prod_{(k,l)\in\mathcal{P}_j} f_{kl}\end{bmatrix}\begin{bmatrix} r_{ij}\\ x_{ij}\end{bmatrix} > 0 \text{ for all } (i,j) \in \mathcal{E}$$
   where $c_{kl} := 1 - 2r_{kl}\hat{P}^+_{kl}(\bar{p})/\underline{v}_k$, $d_{kl} := 2r_{kl}\hat{Q}^+_{kl}(\bar{q})/\underline{v}_k$, $e_{kl} := 2x_{kl}\hat{P}^+_{kl}(\bar{p})/\underline{v}_k$, and $f_{kl} := 1 - 2x_{kl}\hat{Q}^+_{kl}(\bar{q})/\underline{v}_k$.

The results in [50], [51] say that, if there are no voltage upper bounds, i.e., $\overline{v} = \infty$, then SOCP is exact if any one of 1)–5) holds. Since C2 holds automatically when $\overline{v} = \infty$ and C1 holds if any one of 1)–5) holds (Theorem 4), the results in [50], [51] follow from Theorem 1. Besides, the following corollary follows immediately from Theorems 2 and 4.

Corollary 1. *Assume that f_0 is strictly increasing, and that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then SOCP-m is exact if any one of 1)–5) holds.*

VI. Case Studies

In this section, we use six test networks to demonstrate that

1) SOCP is simpler computationally than SDP.
2) C1 holds. We define *C1 margin* that quantifies how well C1 is satisfied, and show that the margin is big.
3) The feasible sets of OPF and OPF-m are similar. We define *modification gap* that quantifies the difference between the feasible sets of OPF and OPF-m, and show that this gap is small.

A. Test networks

The test networks include IEEE 13, 34, 37, 123-bus networks [53] and two real-world networks [30], [54] in the service territory of Southern California Edison (SCE), a utility company in California, USA [55].

The IEEE networks are unbalanced three-phase radial networks with some devices (regulators, circuit switches, transformers, and distributed loads) not modeled in (1). Therefore we modify the IEEE networks as follows.

1) Assume that each bus has three phases and split its load uniformly among the three phases.
2) Assume that the three phases are decoupled so that the network becomes three identical single phase networks.
3) Model closed circuit switches as shorted lines and ignore open circuit switches. Model regulators as multiplying the voltages by fixed constants (set to 1.08 in the simulations). Model transformers as lines with appropriate impedances. Model the distributed load on a line as two identical spot loads located at two ends of the line.

The SCE networks, a 47-bus network and a 56-bus network, are shown in Fig. 7 with parameters given in Tables I and II.

These networks have increasing penetration of distributed generation (DG) as listed in Table III. While the IEEE networks do not have any DG, the SCE 47-bus network has 56.6% DG penetration (6.4MW nameplate DG capacity against 11.3MVA peak spot load), and the SCE 56-bus network has 130.4% DG penetration (5MW nameplate DG capacity against 3.835MVA peak spot load).

B. SOCP is more efficient to compute than SDP

We compare the computation times of SOCP and SDP for the test networks, and summarize the results in Table III. All

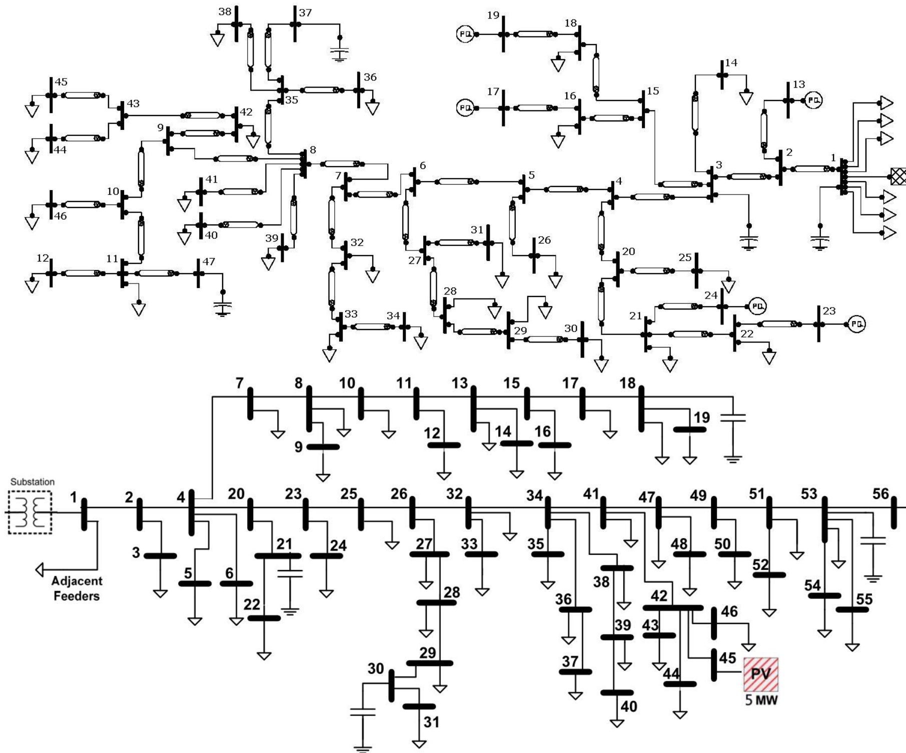

Fig. 7. Topologies of the SCE 47-bus and 56-bus networks [30], [54].

TABLE I
LINE IMPEDANCES, PEAK SPOT LOAD, AND NAMEPLATE RATINGS OF CAPACITORS AND PV GENERATORS OF THE 47-BUS NETWORK.

Network Data															
Line Data				Line Data				Line Data				Load Data		Load Data		PV Generators	
From Bus	To Bus	R (Ω)	X (Ω)	From Bus	To Bus	R (Ω)	X (Ω)	From Bus	To Bus	R (Ω)	X (Ω)	Bus No	Peak MVA	Bus No	Peak MVA	Bus No	Nameplate Capacity
1	2	0.259	0.808	8	41	0.107	0.031	21	22	0.198	0.046	1	30	34	0.2		
2	13	0	0	8	35	0.076	0.015	22	23	0	0	11	0.67	36	0.27	13	1.5MW
2	3	0.031	0.092	8	9	0.031	0.031	27	31	0.046	0.015	12	0.45	38	0.45	17	0.4MW
3	4	0.046	0.092	9	10	0.015	0.015	27	28	0.107	0.031	14	0.89	39	1.34	19	1.5 MW
3	14	0.092	0.031	9	42	0.153	0.046	28	29	0.107	0.031	16	0.07	40	0.13	23	1 MW
3	15	0.214	0.046	10	11	0.107	0.076	29	30	0.061	0.015	18	0.67	41	0.67	24	2 MW
4	20	0.336	0.061	10	46	0.229	0.122	32	33	0.046	0.015	21	0.45	42	0.13		
4	5	0.107	0.183	11	47	0.031	0.015	33	34	0.031	0.010	22	2.23	44	0.45	Shunt Capacitors	
5	26	0.061	0.015	11	12	0.076	0.046	35	36	0.076	0.015	25	0.45	45	0.2	Bus No.	Nameplate Capacity
5	6	0.015	0.031	15	18	0.046	0.015	35	37	0.076	0.046	26	0.2	46	0.45		
6	27	0.168	0.061	15	16	0.107	0.015	35	38	0.107	0.015	28	0.13				
6	7	0.031	0.046	16	17	0	0	42	43	0.061	0.015	29	0.13	Base Voltage (kV) = 12.35		1	6000 kVAR
7	32	0.076	0.015	18	19	0	0	43	44	0.061	0.015	30	0.2	Base kVA = 1000		3	1200 kVAR
7	8	0.015	0.015	20	21	0.122	0.092	43	45	0.061	0.015	31	0.07	Substation Voltage = 12.35		37	1800 kVAR
8	40	0.046	0.015	20	25	0.214	0.046					32	0.13			47	1800 kVAR
8	39	0.244	0.046	21	24	0	0					33	0.27				

simulations in this paper use matlab 7.9.0.529 (64-bit) with toolbox cvx 1.21 on Mac OS X 10.7.5 with 2.66GHz Intel Core 2 Due CPU and 4GB 1067MHz DDR3 memory.

We use the following OPF setup throughout the simulations.

1) The objective is to minimize power loss in the network.
2) The power injection constraints are as follows. For each bus $i \in \mathcal{N}^+$, there may be multiple devices including loads, capacitors, and PV panels. Assume that there is a total of D_i such devices and label them by $1, 2, \ldots, D_i$. Let $s_{i,d}$ denote the power injection of device d for $d = 1, 2, \ldots, D_i$. If device d is a load with given real and reactive power consumptions p and q, then we impose
$$s_{i,d} = -p - \mathbf{i}q. \tag{10}$$
If device d is a load with given peak apparent power s_{peak}, then we impose
$$s_{i,d} = -s_{\text{peak}} \exp(j\theta) \tag{11}$$
where $\theta = \cos^{-1}(0.9)$, i.e, power injection $s_{i,d}$ is considered to be a constant, obtained by assuming a power factor of 0.9 at peak apparent power. If device d is a capacitor with nameplate $\overline{q}$, then we impose
$$\text{Re}(s_{i,d}) = 0 \text{ and } 0 \le \text{Im}(s_{i,d}) \le \overline{q}. \tag{12}$$
If device d is a PV panel with nameplate $\overline{s}$, then we impose
$$\text{Re}(s_{i,d}) \ge 0 \text{ and } |s_{i,d}| \le \overline{s}. \tag{13}$$

TABLE II

LINE IMPEDANCES, PEAK SPOT LOAD, AND NAMEPLATE RATINGS OF CAPACITORS AND PV GENERATORS OF THE 56-BUS NETWORK.

Network Data																	
Line Data				Line Data				Line Data				Load Data		Load Data		Load Data	
From Bus.	To Bus.	R (Ω)	X (Ω)	From Bus.	To Bus.	R (Ω)	X (Ω)	From Bus.	To Bus.	R (Ω)	X (Ω)	Bus No.	Peak MVA	Bus No.	Peak MVA	Bus No.	Peak MVA
1	2	0.160	0.388	20	21	0.251	0.096	39	40	2.349	0.964	3	0.057	29	0.044	52	0.315
2	3	0.824	0.315	21	22	1.818	0.695	34	41	0.115	0.278	5	0.121	31	0.053	54	0.061
2	4	0.144	0.349	20	23	0.225	0.542	41	42	0.159	0.384	6	0.049	32	0.223	55	0.055
4	5	1.026	0.421	23	24	0.127	0.028	42	43	0.934	0.383	7	0.053	33	0.123	56	0.130
4	6	0.741	0.466	23	25	0.284	0.687	42	44	0.506	0.163	8	0.047	34	0.067	Shunt Cap	
4	7	0.528	0.468	25	26	0.171	0.414	42	45	0.095	0.195	9	0.068	35	0.094	Bus	Mvar
7	8	0.358	0.314	26	27	0.414	0.386	42	46	1.915	0.769	10	0.048	36	0.097	19	0.6
8	9	2.032	0.798	27	28	0.210	0.196	41	47	0.157	0.379	11	0.067	37	0.281	21	0.6
8	10	0.502	0.441	28	29	0.395	0.369	47	48	1.641	0.670	12	0.094	38	0.117	30	0.6
10	11	0.372	0.327	29	30	0.248	0.232	47	49	0.081	0.196	14	0.057	39	0.131	53	0.6
11	12	1.431	0.999	30	31	0.279	0.260	49	50	1.727	0.709	16	0.053	40	0.030	Photovoltaic	
11	13	0.429	0.377	26	32	0.205	0.495	49	51	0.112	0.270	17	0.057	41	0.046	Bus	Capacity
13	14	0.671	0.257	32	33	0.263	0.073	51	52	0.674	0.275	18	0.112	42	0.054		
13	15	0.457	0.401	32	34	0.071	0.171	51	53	0.070	0.170	19	0.087	43	0.083	45	5MW
15	16	1.008	0.385	34	35	0.625	0.273	53	54	2.041	0.780	22	0.063	44	0.057		
15	17	0.153	0.134	34	36	0.510	0.209	53	55	0.813	0.334	24	0.135	46	0.134	V_{base} = 12kV	
17	18	0.971	0.722	36	37	2.018	0.829	53	56	0.141	0.340	25	0.100	47	0.045	S_{base} = 1MVA	
18	19	1.885	0.721	34	38	1.062	0.406					27	0.048	48	0.196	Z_{base} = 144Ω	
4	20	0.138	0.334	38	39	0.610	0.238					28	0.038	50	0.045		

TABLE III

DG PENETRATION, C1 MARGINS, MODIFICATION GAPS, AND COMPUTATION TIMES FOR DIFFERENT TEST NETWORKS.

	DG penetration	numerical precision	SOCP time	SDP time	C1 margin	estimated modification gap
IEEE 13-bus	0%	10^{-10}	0.5162s	0.3842s	27.6762	0.0362
IEEE 34-bus	0%	10^{-10}	0.5772s	0.5157s	20.8747	0.0232
IEEE 37-bus	0%	10^{-9}	0.5663s	1.6790s	$+\infty$	0.0002
IEEE 123-bus	0%	10^{-8}	2.9731s	32.6526s	52.9636	0.0157
SCE 47-bus	56.6%	10^{-8}	0.7265s	2.5932s	2.5416	0.0082
SCE 56-bus	130.4%	10^{-9}	1.0599s	6.0573s	1.2972	0.0053

The power injection at bus i is

$$s_i = \sum_{d=1}^{D_i} s_{i,d}$$

where $s_{i,d}$ satisfies one of (10)–(13).

3) The voltage regulation constraint is considered to be $0.9^2 \le v_i \le 1.1^2$ for $i \in \mathcal{N}^+$. Note that we choose a small voltage lower bound 0.9 so that OPF is feasible for all test networks. We choose a big voltage upper bound 1.1 such that Condition C2 holds and therefore SDP/SOCP is exact under C1.

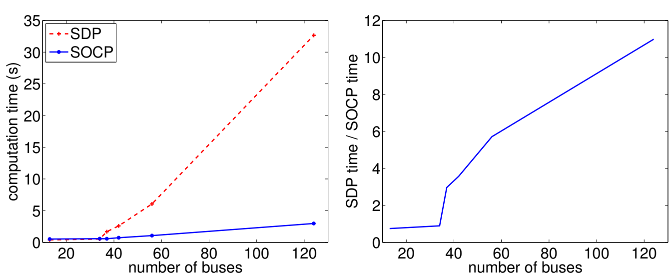

Fig. 8. Comparison of the computation times for SOCP and SDP.

The computation times of SDP and SOCP for different test networks are summarized in Fig. 8. The number of buses determines the number of constraints and variables in the optimization, and therefore reflects the problem size. Network topology also affects the computation time. As the number of buses increases, the computation time of SOCP scales up much more slowly than that of SDP and their ratio increases dramatically. Hence SOCP is much more efficient than SDP for medium to large networks.

SOCP and SDP can only be solved to certain numerical precisions. The best numerical precision we obtain without applying pre-conditioning techniques are listed in Table III.

C. C1 holds with a large margin

In this section, we show that C1 holds with a large margin for all test networks. Noting that C1 becomes more difficult to hold as $(\overline{p}, \overline{q})$ increases (Proposition 1), one can increase $\overline{p}$, $\overline{q}$ until C1 fails. More specifically, let p_i^{fix} and q_i^{fix} denote the fixed real and reactive loads at bus $i \in \mathcal{N}^+$, let PV_i and Cap_i denote the nameplate capacities of the photovoltaic panels and the shunt capacitors at bus $i \in \mathcal{N}^+$, and define

$$\begin{aligned} \overline{p}_i(\eta) &:= p_i^{\text{fix}} + \eta \cdot \text{PV}_i, && i \in \mathcal{N}^+, \ \eta \ge 0; \\ \overline{q}_i(\eta) &:= q_i^{\text{fix}} + \eta \cdot (\text{PV}_i + \text{Cap}_i), && i \in \mathcal{N}^+, \ \eta \ge 0. \end{aligned}$$

When $\eta = 0$, one has $(\overline{p}(\eta), \overline{q}(\eta)) \le 0$ and therefore C1 holds according to Proposition 2. According to Proposition 1, there exists a unique $\eta^* \in \mathbb{R}^+ \cup \{+\infty\}$ such that

$$\eta < \eta^* \Rightarrow \text{C1 holds for } (r, x, \overline{p}(\eta), \overline{q}(\eta), \underline{v}); \quad (14a)$$

$$\eta > \eta^* \Rightarrow \text{C1 does not hold for } (r, x, \overline{p}(\eta), \overline{q}(\eta), \underline{v}). \quad (14b)$$

Definition 2. *C1 margin is defined as the unique* $\eta^* \ge 0$ *that satisfies* (14).

Physically, η^* is the multiple by which one can scale up distributed generation (PVs) and shunt capacitors before C1 fails to hold. Noting that $\overline{p} = \overline{p}(1)$ and $\overline{q} = \overline{q}(1)$, C1 holds for $(r, x, \overline{p}, \overline{q}, \underline{v})$ if and only if $\eta^* > 1$ (ignore the corner case where $\eta^* = 1$). The larger η^* is, the "more easily" C1 holds.

The C1 margins of different test networks are summarized in Table III. The minimum C1 margin is 1.30, meaning that one can scale up distributed generation and shunt capacitors by 1.30 times before C1 fails to hold. C1 margin of the IEEE 37-bus network is $+\infty$, and this is because there is neither distributed generation nor shunt capacitors in the network.

The C1 margin is above 20 for all IEEE networks, but much smaller for SCE networks. This is because SCE networks have big $\overline{p}$ and $\overline{q}$ (due to big PV_i and Cap_i) that make C1 more difficult to hold. However, note that the SCE 56-bus network already has a DG penetration of over 130%, and that one can still scale up its DG by a factor of 1.30 times before C1 breaks down. This highlights that C1 is a mild condition.

D. The feasible sets of OPF and OPF-m are similar

In this section, we show that OPF-m eliminates some feasible points of OPF that are close to the voltage upper bounds for all test networks. To present the result, let $\mathcal{F}_{\text{OPF}}$ denote the feasible set of OPF, let $\|\cdot\|_\infty$ denote the ℓ_∞ norm,[1] and let

$$\varepsilon \ := \ \max \ \|\hat{v}(s) - v\|_\infty \ \ \text{s.t.} \ \ (s, S, v, \ell, s_0) \in \mathcal{F}_{\text{OPF}} \quad (15)$$

denote the maximum deviation of v from its linear approximation $\hat{v}(s)$ over all OPF feasible points (s, S, v, ℓ, s_0).

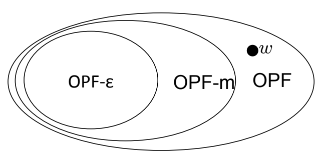

Fig. 9. Feasible sets of OPF-ε, OPF-m, and OPF. The point w is feasible for OPF but not for OPF-m.

The value ε quantifies the difference between the feasible sets of OPF and OPF-m. Consider the OPF problem with a stricter voltage upper bound constraint:

$$\begin{aligned} \textbf{OPF-}\varepsilon\textbf{:} \ \min \quad & \sum_{i \in \mathcal{N}} f_i(\mathrm{Re}(s_i)) \\ \text{over} \quad & s, S, v, \ell, s_0 \\ \text{s.t.} \quad & \text{(5a)} - \text{(5e)}; \\ & \underline{v}_i \le v_i \le \overline{v}_i - \varepsilon, \quad i \in \mathcal{N}^+. \end{aligned}$$

The feasible set $\mathcal{F}_{\text{OPF-}\varepsilon}$ of OPF-ε is contained in $\mathcal{F}_{\text{OPF}}$. Hence, for every $(s, S, \ell, v, s_0) \in \mathcal{F}_{\text{OPF-}\varepsilon} \subseteq \mathcal{F}_{\text{OPF}}$, one has

$$\hat{v}_i(s) \le v_i + \|\hat{v}(s) - v\|_\infty \le \overline{v}_i - \varepsilon + \varepsilon = \overline{v}_i, \quad i \in \mathcal{N}^+$$

by (15). It follows that $\mathcal{F}_{\text{OPF-}\varepsilon} \subseteq \mathcal{F}_{\text{OPF-m}}$ and therefore

$$\mathcal{F}_{\text{OPF-}\varepsilon} \subseteq \mathcal{F}_{\text{OPF-m}} \subseteq \mathcal{F}_{\text{OPF}}$$

as illustrated in Fig. 9.

If ε is small, then $\mathcal{F}_{\text{OPF-m}}$ is similar to $\mathcal{F}_{\text{OPF}}$. Any point w that is feasible for OPF but infeasible for OPF-m is close to the voltage upper bound since $v_i > \overline{v}_i - \varepsilon$ for some $i \in \mathcal{N}^+$. Such points are perhaps undesirable for robust operation.

Definition 3. *The value ε defined in* (15) *is called the modification gap.*

We demonstrate that the modification gap ε is small for all test networks through Monte-Carlo simulations. Note that ε is difficult to compute since the objective function in (15) is not concave and the constraints in (15) are not convex. We choose 1000 samples of s, calculate the corresponding (S, v, ℓ, s_0) by solving the power flow equations (1a)–(1d) (using the *forward backward sweep* algorithm [56]) for each s, and compute $\varepsilon(s) := \|\hat{v}(s) - v\|_\infty$ if $(s, S, v, \ell, s_0) \in \mathcal{F}_{\text{OPF}}$. We use the maximum $\varepsilon(s)$ over the 1000 samples as an estimate for ε. The estimated modification gap ε^{set} we obtained for different test networks are listed in Table III. For example, $\varepsilon^{\text{set}} = 0.0362$ for the IEEE 13-bus network, in which case the voltage constraints are $0.81 \le v_i \le 1.21$ for OPF and $0.81 \le v_i \le 1.1738$ for OPF-ε (assuming $\varepsilon = \varepsilon^{\text{set}}$).

VII. Conclusion

We have proved that SOCP is exact if conditions C1 and C2 hold. C1 can be checked a priori, and has the physical interpretation that upstream power flows should increase if the power loss on a line is reduced. C2 requires that optimal power injections lie in a region $\mathcal{S}_{\text{volt}}$ where voltage upper bounds do not bind. We have proposed a modified OPF problem that includes the additional constraint that power injections lie in $\mathcal{S}_{\text{volt}}$, such that the corresponding SOCP relaxation is exact under C1. We have also proved that SOCP has at most one optimal solution if it it convex and exact. These results unify and generalize our prior works [50], [51]. Empirical studies show that C1 holds with large margin and that the feasible sets of OPF and OPF-m are close, for the IEEE 13, 34, 37, 123-bus networks and two real-world networks.

[1]The ℓ_∞ norm of a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is defined as $\|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}$.

References

[1] J. A. Momoh, *Electric Power System Applications of Optimization*, ser. Power Engineering, H. L. Willis, Ed. Markel Dekker Inc.: New York, USA, 2001.
[2] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *IEEE Trans. on Power Systems*, vol. 6, no. 2, pp. 762–770, 1991.
[3] 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*, vol. 14, no. 1, pp. 96–104, 1999.
[4] ——, "A review of selected optimal power flow literature to 1993. Part II: Newton, linear programming and interior point methods," *IEEE Trans. on Power Systems*, vol. 14, no. 1, pp. 105 – 111, 1999.
[5] K. S. Pandya and S. K. Joshi, "A survey of optimal power flow methods," *J. of Theoretical and Applied Information Technology*, vol. 4, no. 5, pp. 450–458, 2008.
[6] S. Frank, I. Steponavice, and S. Rebennack, "Optimal power flow: a bibliographic survey, I: formulations and deterministic methods," *Energy Systems*, vol. 3, pp. 221–258, September 2012.
[7] ——, "Optimal power flow: a bibliographic survey, II: nondeterministic and hybrid methods," *Energy Systems*, vol. 3, pp. 259–289, September 2013.

[8] M. B. Cain, R. P. O'Neill, and A. Castillo, "History of optimal power flow and formulations (OPF Paper 1)," US FERC, Tech. Rep., December 2012.
[9] R. P. O'Neill, A. Castillo, and M. B. Cain, "The IV formulation and linear approximations of the AC optimal power flow problem (OPF Paper 2)," US FERC, Tech. Rep., December 2012.
[10] ——, "The computational testing of AC optimal power flow using the current voltage formulations (OPF Paper 3)," US FERC, Tech. Rep., December 2012.
[11] A. Castillo and R. P. O'Neill, "Survey of approaches to solving the ACOPF (OPF Paper 4)," US FERC, Tech. Rep., March 2013.
[12] ——, "Computational performance of solution techniques applied to the ACOPF (OPF Paper 5)," US FERC, Tech. Rep., March 2013.
[13] B. Stott and O. Alsac, "Fast decoupled load flow," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 859–869, 1974.
[14] O. Alsac, J. Bright, M. Prais, and B. Stott, "Further developments in lp-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 697–711, 1990.
[15] B. Stott, J. Jardim, and O. Alsac, "Dc power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.
[16] Y. Xiao, Y. Song, and Y. Sun, "Power flow control approach to power systems with embedded FACTS devices," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 943–950, 2002.
[17] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Local control of reactive power by distributed photovoltaic generators," in *IEEE SmartGridComm*, 2010, pp. 79–84.
[18] C. Coffrin and P. Van Hentenryck, "A linear-programming approximation of ac power flows," *arXiv:1206.3614*, 2012.
[19] G. C. Contaxis, C. Delkis, and G. Korres, "Decoupled optimal power flow using linear or quadratic programming," *IEEE Transactions on Power Systems*, vol. 1, no. 2, pp. 1–7, 1986.
[20] W. Min and L. Shengsong, "A trust region interior point algorithm for optimal power flow problems," *International Journal on Electrical Power and Energy Systems*, vol. 27, no. 4, pp. 293–300, 2005.
[21] A. A. Sousa and G. L. Torres, "Robust optimal power flow solution using trust region and interior methods," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 487–499, 2011.
[22] E. C. Baptista, E. A. Belati, and G. R. M. da Costa, "Logarithmic barrier-augmented lagrangian function to the optimal power flow problem," *International Journal on Electrical Power and Energy Systems*, vol. 27, no. 7, pp. 528–532, 2005.
[23] G. L. Torres and V. H. Quintana, "An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1211–1218, 1998.
[24] R. A. Jabr, "A primal-dual interior-point method to solve the optimal power flow dispatching problem," *Optimization and Engineering*, vol. 4, no. 4, pp. 309–336, 2003.
[25] F. Capitanescu, M. Glavic, D. Ernst, and L. Wehenkel, "Interior-point based algorithms for the solution of optimal power flow problems," *Electric Power System Research*, vol. 77, no. 5-6, pp. 508–517, 2007.
[26] S. H. Low, "Convex relaxation of optimal power flow, I: formulations and relaxations," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.
[27] ——, "Convex relaxation of optimal power flow, II: exactness," *IEEE Trans. on Control of Network Systems*, vol. 1, no. 2, June 2014.
[28] R. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, 2006.
[29] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 6-7, pp. 383–392, 2008.
[30] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy, "Inverter VAR control for distribution systems with renewables," in *IEEE SmartGridComm*, 2011, pp. 457–462.
[31] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification (parts I, II)," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2572, 2013.
[32] M. E. Baran and F. F. Wu, "Optimal Capacitor Placement on radial distribution systems," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 725–734, 1989.
[33] ——, "Optimal Sizing of Capacitors Placed on A Radial Distribution System," *IEEE Trans. Power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.
[34] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.
[35] X. Bai and H. Wei, "A semidefinite programming method with graph partitioning technique for optimal power flow problems," *International Journal on Electrical Power and Energy Systems*, vol. 33, no. 7, pp. 1309–1314, 2011.
[36] R. A. Jabr, "Exploiting sparsity in sdp relaxations of the opf problem," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1138–1139, 2012.
[37] D. Molzahn, J. Holzer, B. Lesieutre, and C. DeMarco, "Implementation of a large-scale optimal power flow solver based on semidefinite programming," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3987–3998, 2013.
[38] S. Bose, S. H. Low, T. Teeraratkul, and B. Hassibi, "Equivalent relaxations of optimal power flow," *IEEE Trans. Automatic Control*, 2014.
[39] B. Lesieutre, D. Molzahn, A. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *Allerton*, 2011, pp. 1492–1499.
[40] W. A. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, "Local solutions of optimal power flow," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4780–4788, 2013.
[41] D. K. Molzahn, B. C. Lesieutre, and C. L. DeMarco, "Investigation of non-zero duality gap solutions to a semidefinite relaxation of the optimal power flow problem," in *47th Hawaii International Conference on System Sciences (HICSS), 2014*, January 6-9 2014.
[42] S. Bose, D. F. Gayme, S. Low, and K. M. Chandy, "Optimal power flow over tree networks," in *Allerton*, 2011, pp. 1342–1348.
[43] S. Bose, D. Gayme, K. M. Chandy, and S. H. Low, "Quadratically constrained quadratic programs on acyclic graphs with application to power flow," March 2012, arXiv:1203.5599v1.
[44] B. Zhang and D. Tse, "Geometry of feasible injection region of power networks," in *Proc. Allerton Conf. on Comm., Ctrl. and Computing*, October 2011.
[45] ——, "Geometry of injection regions of power networks," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 788–797, 2013.
[46] S. Sojoudi and J. Lavaei, "Physics of power networks makes hard optimization problems easy to solve," in *IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.
[47] ——, "Semidefinite relaxation for nonlinear optimization over graphs with application to power systems," 2013, preprint.
[48] J. Lavaei, D. Tse, and B. Zhang, "Geometry of power flows and optimization in distribution networks," *arXiv*, November 2012.
[49] A. Y. Lam, B. Zhang, A. Domínguez-García, and D. Tse, "Optimal distributed voltage regulation in power distribution networks," *arXiv*, April 2012.
[50] L. Gan, N. Li, U. Topcu, and S. H. Low, "On the exactness of convex relaxation for optimal power flow in tree networks," in *IEEE Conference on Decision and Control*, 2012, pp. 465–471.
[51] ——, "Optimal power flow in distribution networks," in *IEEE Conference on Decision and Control*, 2013.
[52] N. Li, L. Chen, and S. Low, "Exact convex relaxation of OPF for radial networks using branch flow models," in *IEEE International Conference on Smart Grid Communications*, November 2012.
[53] "IEEE distribution test feeders," online at available at http://ewh.ieee.org/soc/pes/dsacom/testfeeders/.
[54] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter var control in distribution systems with high pv penetration," in *PES General Meeting*, 2012, pp. 1–7.
[55] "Southern california edison," online at http://www.sce.com/.
[56] W. H. Kersting, *Distribution System Modeling and Analysis*. CRC Press, 2006.

Appendix A
Proof of Lemma 1

Let (s, S, v, ℓ, s_0) satisfy (1a)–(1c) and $\ell \geq 0$ componentwise. It follows from (1a) that

$$S_{ij} = s_i + \sum_{h:\, h \to i} (S_{hi} - z_{hi}\ell_{hi}) \leq s_i + \sum_{h:\, h \to i} S_{hi}$$

for $(i,j) \in \mathcal{E}$. On the other hand, $\hat{S}_{ij}(s)$ is the solution of

$$\hat{S}_{ij} = s_i + \sum_{h:\, h \to i} \hat{S}_{hi}, \quad (i,j) \in \mathcal{E}.$$

By induction from the leaf lines, one can show that

$$S_{ij} \leq \hat{S}_{ij}(s), \quad (i,j) \in \mathcal{E}.$$

It follows from (1c) that

$$\begin{aligned} v_i - v_j &= 2\mathrm{Re}(\bar{z}_{ij} S_{ij}) - |z_{ij}|^2 \ell_{ij} \\ &\le 2\mathrm{Re}(\bar{z}_{ij} S_{ij}) \\ &\le 2\mathrm{Re}(\bar{z}_{ij} \hat{S}_{ij}(s)) \end{aligned}$$

for $(i,j) \in \mathcal{E}$. Sum up the inequalities over $\mathcal{P}_i$ to obtain

$$v_i - v_0 \le 2 \sum_{(j,k)\in\mathcal{P}_i} \mathrm{Re}(\bar{z}_{jk} \hat{S}_{jk}(s)),$$

i.e., $v_i \le \hat{v}_i(s)$, for $i \in \mathcal{N}$.

Appendix B
Proof of Theorem 1

The proof idea of Theorem 1 has been illustrated via a 3-bus linear network in Section III-A. Now we present the proof of Theorem 1 for general radial networks. Assume that f_0 is strictly increasing, and that C1 and C2 hold. If SOCP is not exact, then there exists an optimal SOCP solution $w = (s, S, v, \ell, s_0)$ that violates (5d). We will construct another feasible point $w' = (s', S', v', \ell', s_0')$ of SOCP that has a smaller objective value than w. This contradicts the optimality of w, and therefore SOCP is exact.

Construction of w'

The construction of w' is as follows. Since w violates (5d), there exists a leaf bus $l \in \mathcal{L}$ with $m \in \{1, \ldots, n_l\}$ such that w satisfies (5d) on $(l_1, l_0), \ldots, (l_{m-1}, l_{m-2})$ and violates (5d) on (l_m, l_{m-1}). Without loss of generality, assume $l_k = k$ for $k = 0, \ldots, m$ as in Fig. 10. Then

$$\ell_{m,m-1} > \frac{|S_{m,m-1}|^2}{v_m}, \tag{16a}$$

$$\ell_{k,k-1} = \frac{|S_{k,k-1}|^2}{v_k}, \quad k = 1, \ldots, m-1. \tag{16b}$$

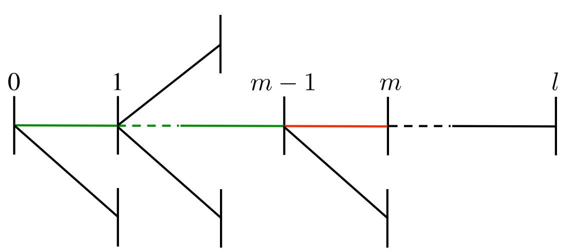

Fig. 10. Bus l is a leaf bus, with $l_k = k$ for $k = 0, \ldots, m$. Equality (5d) is satisfied on $[0, m-1]$, but violated on $[m-1, m]$.

One can then construct $w' = (s', S', v', \ell', s_0')$ as in Algorithm 1. The construction consists of three steps:

- S1 In the initialization step, s', ℓ' outside path $\mathcal{P}_m$, and S' outside path $\mathcal{P}_{m-1}$ are initialized as the corresponding values in w. Since $s' = s$, the point w' satisfies (5e). Furthermore, since $\ell'_{ij} = \ell_{ij}$ for $(i,j) \notin \mathcal{P}_m$ and $S'_{ij} = S_{ij}$ for $(i,j) \notin \mathcal{P}_{m-1}$, the point w' also satisfies (5a) for $(i,j) \notin \mathcal{P}_{m-1}$.
- S2 In the forward sweep step, $\ell'_{k,k-1}$ and $S'_{k-1,k-2}$ are recursively constructed for $k = m, \ldots, 1$ by alternatively applying (5d) (with v' replaced by v) and (5a)/(5b). Hence, w' satisfies (5a) for $(i,j) \in \mathcal{P}_{m-1}$ and (5b).
- S3 In the backward sweep step, v'_i is recursively constructed from bus 0 to leaf buses by applying (5c) consecutively. Hence, the point w' satisfies (5c).

Algorithm 1 Construct a feasible point

Input: an optimal SOCP solution $w = (s, S, v, \ell, s_0)$ that violates (5d), a leaf bus $l \in \mathcal{L}$ with $1 \le m \le n_l$ such that (16) holds (assume $l_k = k$ for $k = 0, \ldots, m$ without loss of generality).

Output: $w' = (s', S', v', \ell', s_0')$.

1: Initialization.
(Construct s', ℓ' outside $\mathcal{P}_m$, and S' outside $\mathcal{P}_{m-1}$.)
keep s:

$$s' \leftarrow s;$$

keep ℓ outside path $\mathcal{P}_m$:

$$\ell'_{ij} \leftarrow \ell_{ij}, \quad (i,j) \notin \mathcal{P}_m;$$

keep S outside path $\mathcal{P}_{m-1}$:

$$S'_{ij} \leftarrow S_{ij}, \quad (i,j) \notin \mathcal{P}_{m-1};$$

2: Forward sweep.
(Construct ℓ' on $\mathcal{P}_m$, S' on $\mathcal{P}_{m-1}$, and s_0'.)
for $k = m, m-1, \ldots, 1$ **do**

$$\begin{aligned} \ell'_{k,k-1} &\leftarrow \frac{|S'_{k,k-1}|^2}{v_k}; \\ S'_{k-1,k-2} &\leftarrow s_{k-1} \mathbb{1}_{k \ne 1} \\ &\quad + \sum_{j:\, j \to k-1} (S'_{j,k-1} - z_{j,k-1} \ell'_{j,k-1}); \end{aligned}$$

end for

$$s_0' \leftarrow -S'_{0,-1};$$

3: Backward sweep.
(Construct v'.)

$$v_0' \leftarrow v_0, \ \mathcal{N}_{\text{visit}} = \{0\};$$

while $\mathcal{N}_{\text{visit}} \ne \mathcal{N}$ **do**
find $i \notin \mathcal{N}_{\text{visit}}$ and $j \in \mathcal{N}_{\text{visit}}$ such that $i \to j$;

$$\begin{aligned} v_i' &\leftarrow v_j' + 2\mathrm{Re}(\bar{z}_{ij} S'_{ij}) - |z_{ij}|^2 \ell'_{ij}; \\ \mathcal{N}_{\text{visit}} &\leftarrow \mathcal{N}_{\text{visit}} \cup \{i\}; \end{aligned}$$

end while

The point w' satisfies another important property given below.

Lemma 2. *The point w' satisfies $\ell'_{ij} \ge |S'_{ij}|^2 / v_i$ for $(i,j) \in \mathcal{E}$.*

Proof. When $(i,j) \notin \mathcal{P}_m$, it follows from Step S1 that $\ell'_{ij} = \ell_{ij} \ge |S_{ij}|^2/v_i = |S'_{ij}|^2/v_i$. When $(i,j) \in \mathcal{P}_m$, it follows from Step S2 that $\ell'_{ij} = |S'_{ij}|^2/v_i$. This completes the proof of Lemma 2. □

Lemma 2 implies that if $v' \ge v$, then w' satisfies (6).

Feasibility and Superiority of w'

We will show that w' is feasible for SOCP and has a smaller objective value than w. This result follows from Claims 1 and 2.

Claim 1. *If C1 holds, then $S'_{k,k-1} > S_{k,k-1}$ for $k = 0,\dots,m-1$ and $v' \ge v$.*

Claim 1 is proved later in this appendix. Here we illustrate with Fig. 11 that $S'_{k,k-1} > S_{k,k-1}$ for $k = 0,\dots,m-1$ seems natural to hold. Note that $S'_{m,m-1} = S_{m,m-1}$ and

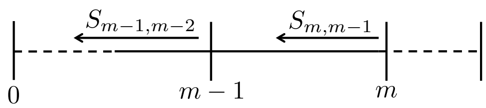

Fig. 11. Illustration of $S'_{k,k-1} > S_{k,k-1}$ for $k = 0,\dots,m-1$.

that $\ell'_{m,m-1} = |S'_{m,m-1}|^2/v_m = |S_{m,m-1}|^2/v_m < \ell_{m,m-1}$. Define $\Delta w = (\Delta s, \Delta S, \Delta v, \Delta \ell, \Delta s_0) = w' - w$, then $\Delta \ell_{m,m-1} < 0$ and therefore

$$\begin{aligned}\Delta S_{m-1,m-2} &= \Delta S_{m,m-1} - z_{m,m-1}\Delta\ell_{m,m-1} \\ &= -z_{m,m-1}\Delta\ell_{m,m-1} > 0. \end{aligned} \quad (17)$$

Intuitively, after increasing $S_{m-1,m-2}$, upstream reverse power flow $S_{k,k-1}$ is likely to increase for $k = 0,\dots,m-2$. C1 is a condition that ensures $S_{k,k-1}$ to increase for $k = 0,\dots,m-1$.

Claim 2. *If C2 holds, then $v' \le \overline{v}$.*

Proof. If C2 holds, then it follows from Lemma 1 that $v' \le \hat{v}(s') = \hat{v}(s) \le \overline{v}$. □

It follows from Claims 1 and 2 that $\underline{v} \le v \le v' \le \overline{v}$, and therefore w' satisfies (5f). Besides, it follows from Lemma 2 that $\ell'_{ij} \ge |S'_{ij}|^2/v_i \ge |S'_{ij}|^2/v'_i$ for $(i,j) \in \mathcal{E}$, i.e., w' satisfies (6). Hence, w' is feasible for SOCP. Furthermore, w' has a smaller objective value than w because

$$\begin{aligned}&\sum_{i\in\mathcal{N}} f_i(\mathrm{Re}(s'_i)) - \sum_{i\in\mathcal{N}} f_i(\mathrm{Re}(s_i)) \\ &= f_0(-\mathrm{Re}(S'_{0,-1})) - f_0(-\mathrm{Re}(S_{0,-1})) < 0.\end{aligned}$$

This contradicts with the optimality of w, and therefore SOCP is exact. To complete the proof, we are left to prove Claim 1.

Proof of Claim 1

Assume C1 holds. First show that $\Delta S_{k,k-1} > 0$ for $k = 0,\dots,m-1$. Recall that $S = P + \mathbf{i}Q$ and that $u_i = (r_{ij}\ x_{ij})^T$. It follows from (17) that

$$\begin{bmatrix}\Delta P_{m-1,m-2}\\ \Delta Q_{m-1,m-2}\end{bmatrix} = -u_m \Delta\ell_{m,m-1} > 0.$$

For any $k \in \{1,\dots,m-1\}$, one has

$$\begin{aligned}\Delta S_{k-1,k-2} &= \Delta S_{k,k-1} - z_{k,k-1}\Delta\ell_{k,k-1} \\ &= \Delta S_{k,k-1} - z_{k,k-1}\frac{|S'_{k,k-1}|^2 - |S_{k,k-1}|^2}{v_k},\end{aligned}$$

which is equivalent to

$$\begin{bmatrix}\Delta P_{k-1,k-2}\\ \Delta Q_{k-1,k-2}\end{bmatrix} = B_k \begin{bmatrix}\Delta P_{k,k-1}\\ \Delta Q_{k,k-1}\end{bmatrix}$$

where

$$B_k = I - \frac{2}{v_k}\begin{bmatrix} r_{k,k-1}\\ x_{k,k-1}\end{bmatrix}\begin{bmatrix}\frac{P_{k,k-1}+P'_{k,k-1}}{2} & \frac{Q_{k,k-1}+Q'_{k,k-1}}{2}\end{bmatrix}.$$

Hence, one has

$$\begin{bmatrix}\Delta P_{k-1,k-2}\\ \Delta Q_{k-1,k-2}\end{bmatrix} = -B_k B_{k+1}\cdots B_{m-1} u_m \Delta\ell_{m,m-1}$$

for $k = 1,\dots,m$. To prove $\Delta S_{k,k-1} > 0$ for $k = 0,\dots,m-1$, it suffices to show that $B_k \cdots B_{m-1} u_m > 0$ for $k = 1,\dots,m$.

C1 implies that $\underline{A}_s \cdots \underline{A}_{t-1} u_t > 0$ when $1 \le s \le t \le m$. One also has $B_k - \underline{A}_k = u_k b_k^T$ where

$$b_k = \begin{bmatrix}\frac{2\hat{P}^+_{k,k-1}(\overline{p})}{\underline{v}_k} - \frac{P_{k,k-1}+P'_{k,k-1}}{v_k}\\ \frac{2\hat{Q}^+_{k,k-1}(\overline{q})}{\underline{v}_k} - \frac{Q_{k,k-1}+Q'_{k,k-1}}{v_k}\end{bmatrix} \ge 0$$

for $k = 1,\dots,m-1$. To show that $B_k \cdots B_{m-1} u_m > 0$ for $k = 1,\dots,m$, we prove the following lemma.

Lemma 3. *Given $m \ge 1$ and $d \ge 1$. Let $\underline{A}_1,\dots,\underline{A}_{m-1}, A_1,\dots,A_{m-1} \in \mathbb{R}^{d\times d}$ and $u_1,\dots,u_m \in \mathbb{R}^d$ satisfy*

- *$\underline{A}_s \cdots \underline{A}_{t-1} u_t > 0$ when $1 \le s \le t \le m$;*
- *there exists $b_k \in \mathbb{R}^d$ that satisfies $b_k \ge 0$ and $A_k - \underline{A}_k = u_k b_k^T$, for $k = 1,\dots,m-1$.*

Then

$$A_s \cdots A_{t-1} u_t > 0 \quad (18)$$

when $1 \le s \le t \le m$.

Proof. We prove that (18) holds when $1 \le t \le s \le m$ by mathematical induction on $t - s$.

i) When $t - s = 0$, one has $A_s \cdots A_{t-1} u_t = u_t = \underline{A}_s \cdots \underline{A}_{t-1} u_t > 0$.

ii) Assume that (18) holds when $t - s = 0, 1, \dots, K$ ($0 \le K \le m-2$). When $t - s = K+1$, one has

$$\begin{aligned}&A_s \cdots A_k \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \\ &= A_s \cdots A_{k-1} \underline{A}_k \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \\ &\quad + A_s \cdots A_{k-1} (A_k - \underline{A}_k) \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \\ &= A_s \cdots A_{k-1} \underline{A}_k \cdots \underline{A}_{t-1} u_t \\ &\quad + A_s \cdots A_{k-1} u_k b_k^T \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \\ &= A_s \cdots A_{k-1} \underline{A}_k \cdots \underline{A}_{t-1} u_t \\ &\quad + \left(b_k^T \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t\right) A_s \cdots A_{k-1} u_k\end{aligned}$$

for $k = s,\dots,t-1$. Since $b_k \ge 0$ and $\underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t > 0$, the term $b_k^T \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \ge 0$. According to induction hypothesis, $A_s \cdots A_{k-1} u_k > 0$. Hence,

$$A_s \cdots A_k \underline{A}_{k+1} \cdots \underline{A}_{t-1} u_t \ge A_s \cdots A_{k-1} \underline{A}_k \cdots \underline{A}_{t-1} u_t$$

for $k = s, \ldots, t-1$. By substituting $k = t-1, \ldots, s$ in turn, one obtains

$$\begin{aligned} A_s \cdots A_{t-1} u_t &\geq A_s \cdots A_{t-2} \underline{A}_{t-1} u_t \\ &\geq \cdots \\ &\geq \underline{A}_s \cdots \underline{A}_{t-1} u_t > 0, \end{aligned}$$

i.e., (18) holds when $t - s = K + 1$.

According to i) and ii), (18) holds when $t - s = 0, \ldots, m-1$. This completes the proof of Lemma 3. □

Lemma 3 implies that $B_s \cdots B_{t-1} u_t > 0$ when $1 \leq s \leq t \leq m$. In particular, $B_k \cdots B_{m-1} u_m > 0$ for $k = 1, \ldots, m$, and therefore $\Delta S_{k,k-1} > 0$ for $k = 0, \ldots, m-1$.

Next show that $v' \geq v$. Noting that $\Delta S_{ij} = 0$ for $(i,j) \notin \mathcal{P}_{m-1}$ and $\Delta \ell_{ij} = 0$ for $(i,j) \notin \mathcal{P}_m$, it follows from (5c) that

$$\Delta v_i - \Delta v_j = 2\mathrm{Re}(\bar{z}_{ij} \Delta S_{ij}) - |z_{ij}|^2 \Delta \ell_{ij} = 0$$

for $(i,j) \notin \mathcal{P}_m$. When $(i,j) \in \mathcal{P}_m$, one has $(i,j) = (k, k-1)$ for some $k \in \{1, \ldots, m\}$, and therefore

$$\begin{aligned} \Delta v_i - \Delta v_j &= 2\mathrm{Re}(\bar{z}_{k,k-1} \Delta S_{k,k-1}) - |z_{k,k-1}|^2 \Delta \ell_{k,k-1} \\ &\geq \mathrm{Re}(\bar{z}_{k,k-1} \Delta S_{k,k-1}) - |z_{k,k-1}|^2 \Delta \ell_{k,k-1} \\ &= \mathrm{Re}(\bar{z}_{k,k-1} (\Delta S_{k,k-1} - z_{k,k-1} \Delta \ell_{k,k-1})) \\ &= \mathrm{Re}(\bar{z}_{k,k-1} \Delta S_{k-1,k-2}) > 0. \end{aligned}$$

Hence, $\Delta v_i \geq \Delta v_j$ whenever $(i,j) \in \mathcal{E}$. Add the inequalities over path $\mathcal{P}_i$ to obtain $\Delta v_i \geq \Delta v_0 = 0$ for $i \in \mathcal{N}^+$, i.e., $v' \geq v$. This completes the proof of Claim 1.

Appendix C
Proof of Proposition 1

Let $\underline{A}$ and $\underline{A}'$ denote the matrices with respect to $(\overline{p}, \overline{q})$ and $(\overline{p}', \overline{q}')$ respectively, i.e., let

$$\begin{aligned} \underline{A}_i' &= I - \frac{2}{\underline{v}_i} u_i \left(\hat{P}_{ij}^+(\overline{p}') \ \hat{Q}_{ij}^+(\overline{q}') \right), \quad (i,j) \in \mathcal{E}; \\ \underline{A}_i &= I - \frac{2}{\underline{v}_i} u_i \left(\hat{P}_{ij}^+(\overline{p}) \ \hat{Q}_{ij}^+(\overline{q}) \right), \quad (i,j) \in \mathcal{E}. \end{aligned}$$

When $(\overline{p}, \overline{q}) \leq (\overline{p}', \overline{q}')$, one has $\underline{A}_{l_k} - \underline{A}_{l_k}' = u_{l_k} b_{l_k}^T$ where

$$b_{l_k} = \frac{2}{\underline{v}_{l_k}} \begin{bmatrix} \hat{P}_{l_k l_{k-1}}^+(\overline{p}') - \hat{P}_{l_k l_{k-1}}^+(\overline{p}) \\ \hat{Q}_{l_k l_{k-1}}^+(\overline{q}') - \hat{Q}_{l_k l_{k-1}}^+(\overline{q}) \end{bmatrix} \geq 0$$

for any $l \in \mathcal{L}$ and any $k \in \{1 \ldots, n_l\}$.

If $\underline{A}_{l_s}' \cdots \underline{A}_{l_{t-1}}' u_{l_t} > 0$ for any $l \in \mathcal{L}$ and any s, t such that $1 \leq s \leq t \leq n_l$, then it follows from Lemma 3 that $\underline{A}_{l_s} \cdots \underline{A}_{l_{t-1}} u_{l_t} > 0$ for any $l \in \mathcal{L}$ any s, t such that $1 \leq s \leq t \leq n_l$. This completes the proof of Proposition 1.

Appendix D
Proof of Theorem 3

Assume that f_i is convex for $i \in \mathcal{N}$, that $\mathcal{S}_i$ is convex for $i \in \mathcal{N}^+$, that SOCP-m is exact, and that SOCP-m has at least one solution. Let $\tilde{w} = (\tilde{s}, \tilde{S}, \tilde{v}, \tilde{\ell}, \tilde{s}_0)$ and $\hat{w} = (\hat{s}, \hat{S}, \hat{v}, \hat{\ell}, \hat{s}_0)$ denote two arbitrary SOCP-m solutions. It suffices to show that $\tilde{w} = \hat{w}$.

Since SOCP-m is exact, $\tilde{v}_i \tilde{\ell}_{ij} = |\tilde{S}_{ij}|^2$ and $\hat{v}_i \hat{\ell}_{ij} = |\hat{S}_{ij}|^2$ for $(i,j) \in \mathcal{E}$. Define $w := (\tilde{w} + \hat{w})/2$. Since SOCP-m is convex, w also solves SOCP-m. Hence, $v_i \ell_{ij} = |S_{ij}|^2$ for $(i,j) \in \mathcal{E}$. Substitute $v_i = (\tilde{v}_i + \hat{v}_i)/2$, $\ell_{ij} = (\tilde{\ell}_{ij} + \hat{\ell}_{ij})/2$, and $S_{ij} = (\tilde{S}_{ij} + \hat{S}_{ij})/2$ to obtain

$$\hat{S}_{ij} \tilde{S}_{ij}^H + \tilde{S}_{ij} \hat{S}_{ij}^H = \hat{v}_i \tilde{\ell}_{ij} + \tilde{v}_i \hat{\ell}_{ij}$$

for $(i,j) \in \mathcal{E}$ where the superscript H stands for hermitian transpose. The right hand side

$$\hat{v}_i \tilde{\ell}_{ij} + \tilde{v}_i \hat{\ell}_{ij} = \hat{v}_i \frac{|\tilde{S}_{ij}|^2}{\tilde{v}_i} + \tilde{v}_i \frac{|\hat{S}_{ij}|^2}{\hat{v}_i} \geq 2 |\tilde{S}_{ij}| |\hat{S}_{ij}|,$$

and the equality is attained if and only if $|\tilde{S}_{ij}|/\tilde{v}_i = |\hat{S}_{ij}|/\hat{v}_i$. The left hand side

$$\hat{S}_{ij} \tilde{S}_{ij}^H + \tilde{S}_{ij} \hat{S}_{ij}^H \leq 2 |\tilde{S}_{ij}| |\hat{S}_{ij}|,$$

and the equality is attained if and only if $\angle \hat{S}_{ij} = \angle \tilde{S}_{ij}$. Hence, $\tilde{S}_{ij}/\tilde{v}_i = \hat{S}_{ij}/\hat{v}_i$ for $(i,j) \in \mathcal{E}$.

Introduce $\hat{v}_0 := \tilde{v}_0 := v_0$ and define $\eta_i := \hat{v}_i/\tilde{v}_i$ for $i \in \mathcal{N}$, then $\eta_0 = 1$ and $\hat{S}_{ij} = \eta_i \tilde{S}_{ij}$ for $(i,j) \in \mathcal{E}$. Hence,

$$\hat{\ell}_{ij} = \frac{|\hat{S}_{ij}|^2}{\hat{v}_i} = \frac{|\eta_i \tilde{S}_{ij}|^2}{\eta_i \tilde{v}_i} = \eta_i \frac{|\tilde{S}_{ij}|^2}{\tilde{v}_i} = \eta_i \tilde{\ell}_{ij}$$

and therefore

$$\eta_j = \frac{\hat{v}_j}{\tilde{v}_j} = \frac{\hat{v}_i - 2\mathrm{Re}(z_{ij}^H \hat{S}_{ij}) + |z_{ij}|^2 \hat{\ell}_{ij}}{\tilde{v}_i - 2\mathrm{Re}(z_{ij}^H \tilde{S}_{ij}) + |z_{ij}|^2 \tilde{\ell}_{ij}} = \eta_i$$

for $(i,j) \in \mathcal{E}$. Since the network $(\mathcal{N}, \mathcal{E})$ is connected, $\eta_i = \eta_0 = 1$ for $i \in \mathcal{N}$. This implies $\hat{w} = \tilde{w}$ and completes the proof of Theorem 3.

Appendix E
Proof of Theorem 4

Theorem 4 follows from Claims 3–7.

Claim 3. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \leq \overline{p}_i,\ \mathrm{Im}(s) \leq \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if $\hat{S}_{ij}(\overline{p} + \mathbf{i}\overline{q}) \leq 0$ for all $(i,j) \in \mathcal{E}'$.*

Proof. If $\hat{S}_{ij}(\overline{p} + \mathbf{i}\overline{q}) \leq 0$ for all $(i,j) \in \mathcal{E}'$, then $\underline{A}_{l_k} = I$ for all $l \in \mathcal{L}$ and all $k \in \{1 \ldots, n_l - 1\}$. It follows that $\underline{A}_{l_s} \cdots \underline{A}_{l_{t-1}} u_{l_t} = u_{l_t} > 0$ for all $l \in \mathcal{L}$ and all s, t such that $1 \leq s \leq t \leq n_l$, i.e., C1 holds. □

Claim 4. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \leq \overline{p}_i,\ \mathrm{Im}(s) \leq \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if 1) r_{ij}/x_{ij} is identical for all $(i,j) \in \mathcal{E}$; and 2) $\underline{v}_i - 2r_{ij}\hat{P}_{ij}^+(\overline{p}) - 2x_{ij}\hat{Q}_{ij}^+(\overline{q}) > 0$ for all $(i,j) \in \mathcal{E}'$.*

Proof. Assume the conditions in Claim 4 hold. Fix an arbitrary $l \in \mathcal{L}$, and assume $l_k = k$ for $k = 0, \ldots, n_l$ without loss of generality. Fix an arbitrary $t \in \{1, \ldots, n_l\}$, and define $(\alpha_s \ \beta_s)^T := \underline{A}_s \cdots \underline{A}_{t-1} u_t$ for $s = 1, \ldots, t$. Then it suffices to prove that $\alpha_s > 0$ and $\beta_s > 0$ for $s = 1, \ldots, t$. In particular, we prove

$$\alpha_s > 0, \ \beta_s > 0, \ \alpha_s/\beta_s = r_{10}/x_{10} \tag{19}$$

inductively for $s = t, t-1, \ldots, 1$. Define $\eta := r_{10}/x_{10}$ and note that $r_{ij}/x_{ij} = \eta$ for all $(i,j) \in \mathcal{E}$.

i) When $s = t$, one has $\alpha_s = r_{t,t-1}$, $\beta_s = x_{t,t-1}$, and $\alpha_s/\beta_s = \eta$. Therefore (19) holds.

ii) Assume that (19) holds for $s = k$ $(2 \le k \le t)$, then

$$\begin{bmatrix}\alpha_k & \beta_k\end{bmatrix}^T = c\begin{bmatrix}\eta & 1\end{bmatrix}^T$$

for some $c \in \{c \in \mathbb{R} \mid c > 0\}$. Abbreviate $r_{k-1,k-2}$ by r, $x_{k-1,k-2}$ by x, $\hat{P}^+_{k-1,k-2}(\bar{p})$ by P, and $\hat{Q}^+_{k-1,k-2}(\bar{q})$ by Q for convenience. Then

$$\underline{v}_{k-1} - 2rP - 2xQ > 0$$

and it follows that

$$\begin{aligned}\begin{bmatrix}\alpha_{k-1}\\ \beta_{k-1}\end{bmatrix} &= \left(I - \frac{2}{\underline{v}_{k-1}}\begin{bmatrix}r\\ x\end{bmatrix}\begin{bmatrix}P & Q\end{bmatrix}\right)\begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix}\\ &= \left(I - \frac{2}{\underline{v}_{k-1}}x\begin{bmatrix}\eta\\ 1\end{bmatrix}\begin{bmatrix}P & Q\end{bmatrix}\right)c\begin{bmatrix}\eta\\ 1\end{bmatrix}\\ &= c\begin{bmatrix}\eta\\ 1\end{bmatrix} - \frac{2}{\underline{v}_{k-1}}c\begin{bmatrix}\eta\\ 1\end{bmatrix}\begin{bmatrix}P & Q\end{bmatrix}x\begin{bmatrix}\eta\\ 1\end{bmatrix}\\ &= \begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix} - \frac{2}{\underline{v}_{k-1}}\begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix}\begin{bmatrix}P & Q\end{bmatrix}\begin{bmatrix}r\\ x\end{bmatrix}\\ &= \left(1 - \frac{2}{\underline{v}_{k-1}}(rP + xQ)\right)\begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix}\\ &= \frac{1}{\underline{v}_{k-1}}\left(\underline{v}_{k-1} - 2rP - 2xQ\right)\begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix} > 0\end{aligned}$$

and $\alpha_{k-1}/\beta_{k-1} = \alpha_k/\beta_k = \eta$. Hence, (19) holds for $s = k-1$.

According to i) and ii), (19) holds for $s = t, t-1\dots, 1$. This completes the proof of Claim 4. □

Claim 5. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if 1) $r_{ij}/x_{ij} \ge r_{jk}/x_{jk}$ whenever $(i,j),(j,k) \in \mathcal{E}$; and 2) $\hat{P}_{ij}(\overline{p}) \le 0$, $\underline{v}_i - 2x_{ij}\hat{Q}^+_{ij}(\overline{q}) > 0$ for all $(i,j) \in \mathcal{E}'$.*

Proof. Assume the conditions in Claim 5 hold. Fix an arbitrary $l \in \mathcal{L}$, and assume $l_k = k$ for $k = 0,\dots,n_l$ without loss of generality. Fix an arbitrary $t \in \{1,\dots,n_l\}$, and define $(\alpha_s\ \beta_s)^T := \underline{A}_s \cdots \underline{A}_{t-1}u_t$ for $s = 1,\dots,t$. Then it suffices to prove that $\alpha_s > 0$ and $\beta_s > 0$ for $s = 1,\dots,t$. In particular, we prove

$$\alpha_s > 0,\ \beta_s > 0,\ \alpha_s/\beta_s \ge r_{t,t-1}/x_{t,t-1} \tag{20}$$

inductively for $s = t, t-1,\dots,1$. Define $\eta := r_{t,t-1}/x_{t,t-1}$ and note that $r_{s,s-1}/x_{s,s-1} \le \eta$ for $s = 1,2,\dots,t$.

i) When $s = t$, one has $\alpha_s = r_{t,t-1}$, $\beta_s = x_{t,t-1}$, and $\alpha_s/\beta_s = \eta$. Therefore (20) holds.

ii) Assume that (20) holds for $s = k$ $(2 \le k \le t)$, then

$$\alpha_k \ge \eta\beta_k > 0.$$

Abbreviate $r_{k-1,k-2}$ by r, $x_{k-1,k-2}$ by x, $\hat{P}^+_{k-1,k-2}(\bar{p})$ by P, and $\hat{Q}^+_{k-1,k-2}(\bar{q})$ by Q for convenience. Then

$$P = 0, \quad \underline{v}_k - 2xQ > 0$$

and it follows that

$$\begin{aligned}\begin{bmatrix}\alpha_{k-1}\\ \beta_{k-1}\end{bmatrix} &= \left(I - \frac{2}{\underline{v}_{k-1}}\begin{bmatrix}r\\ x\end{bmatrix}\begin{bmatrix}P & Q\end{bmatrix}\right)\begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix}\\ &= \begin{bmatrix}\alpha_k\\ \beta_k\end{bmatrix} - \frac{2}{\underline{v}_{k-1}}\begin{bmatrix}r\\ x\end{bmatrix}Q\beta_k.\end{aligned}$$

Hence,

$$\beta_{k-1} = \beta_k - \frac{2xQ}{\underline{v}_{k-1}}\beta_k = \frac{1}{\underline{v}_{k-1}}\left(\underline{v}_{k-1} - 2xQ\right)\beta_k > 0.$$

Then,

$$\begin{aligned}\alpha_{k-1} &= \alpha_k - \frac{2rQ}{\underline{v}_{k-1}}\beta_k\\ &\ge \left(\eta - \frac{2rQ}{\underline{v}_{k-1}}\right)\beta_k\\ &\ge \eta\left(1 - \frac{2xQ}{\underline{v}_{k-1}}\right)\beta_k\\ &= \eta\beta_{k-1} > 0.\end{aligned}$$

The second inequality is due to $r/x \le \eta$. Hence, (20) holds for $s = k-1$.

According to i) and ii), (20) holds for $s = t, t-1,\dots,1$. This completes the proof of Claim 5. □

Claim 6. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if 1) $r_{ij}/x_{ij} \le r_{jk}/x_{jk}$ whenever $(i,j),(j,k) \in \mathcal{E}$; and 2) $\hat{Q}_{ij}(\overline{q}) \le 0$, $\underline{v}_i - 2r_{ij}\hat{P}^+_{ij}(\overline{p}) > 0$ for all $(i,j) \in \mathcal{E}'$.*

Proof. The proof of Claim 6 is similar to that of Claim 5 and omitted for brevity. □

Claim 7. *Assume that there exists $\overline{p}_i$ and $\overline{q}_i$ such that $\mathcal{S}_i \subseteq \{s \in \mathbb{C} \mid \mathrm{Re}(s) \le \overline{p}_i,\ \mathrm{Im}(s) \le \overline{q}_i\}$ for $i \in \mathcal{N}^+$. Then C1 holds if*

$$\begin{bmatrix}\prod\limits_{(k,l)\in\mathcal{P}_j} c_{kl} & -\sum\limits_{(k,l)\in\mathcal{P}_j} d_{kl}\\ -\sum\limits_{(k,l)\in\mathcal{P}_j} e_{kl} & \prod\limits_{(k,l)\in\mathcal{P}_j} f_{kl}\end{bmatrix}\begin{bmatrix}r_{ij}\\ x_{ij}\end{bmatrix} > 0, \quad (i,j) \in \mathcal{E} \tag{21}$$

where $c_{kl} := 1 - 2r_{kl}\hat{P}^+_{kl}(\overline{p})/\underline{v}_k$, $d_{kl} := 2r_{kl}\hat{Q}^+_{kl}(\overline{q})/\underline{v}_k$, $e_{kl} := 2x_{kl}\hat{P}^+_{kl}(\overline{p})/\underline{v}_k$, and $f_{kl} := 1 - 2x_{kl}\hat{Q}^+_{kl}(\overline{q})/\underline{v}_k$.

The following lemma is used in the proof of Claim 7.

Lemma 4. *Given $i \ge 1$; c, d, e, $f \in \mathbb{R}^i$ such that $0 < c \le 1$, $d \ge 0$, $e \ge 0$, and $0 < f \le 1$ componentwise; and $u \in \mathbb{R}^2$ that satisfies $u > 0$. If*

$$\begin{bmatrix}\prod\limits_{j=1}^{i} c_j & -\sum\limits_{j=1}^{i} d_j\\ -\sum\limits_{j=1}^{i} e_j & \prod\limits_{j=1}^{i} f_j\end{bmatrix}u > 0, \tag{22}$$

then

$$\begin{bmatrix}c_j & -d_j\\ -e_j & f_j\end{bmatrix}\cdots\begin{bmatrix}c_i & -d_i\\ -e_i & f_i\end{bmatrix}u > 0 \tag{23}$$

for $j = 1,\dots,i$.

Proof. Lemma 4 can be proved by mathematical induction on i.

i) When $i = 1$, Lemma 4 is trivial.

ii) Assume that Lemma 4 holds for $i = K$ ($K \geq 1$). When $i = K+1$, if

$$\begin{bmatrix} \prod_{j=1}^{i} c_j & -\sum_{j=1}^{i} d_j \\ -\sum_{j=1}^{i} e_j & \prod_{j=1}^{i} f_j \end{bmatrix} u > 0,$$

one can prove that (23) holds for $j = 1, \ldots, K+1$ as follows.

First prove that (23) holds for $j = 2, \ldots, K+1$. The idea is to construct some $c', d', e', f' \in \mathbb{R}^K$ and apply the induction hypothesis. The construction is

$$\begin{aligned} c' &= (c_2,\ c_3,\ \ldots,\ c_{K+1}), \\ d' &= (d_2,\ d_3,\ \ldots,\ d_{K+1}), \\ e' &= (e_2,\ e_3,\ \ldots,\ e_{K+1}), \\ f' &= (f_2,\ f_3,\ \ldots,\ f_{K+1}). \end{aligned}$$

Clearly, c', d', e', f' satisfies $0 < c' \leq 1$, $d' \geq 0$, $e' \geq 0$, $0 < f' \leq 1$ componentwise and

$$\begin{aligned} \begin{bmatrix} \prod_{j=1}^{K} c'_j & -\sum_{j=1}^{K} d'_j \\ -\sum_{j=1}^{K} e'_j & \prod_{j=1}^{K} f'_j \end{bmatrix} u &= \begin{bmatrix} \prod_{j=2}^{K+1} c_j & -\sum_{j=2}^{K+1} d_j \\ -\sum_{j=2}^{K+1} e_j & \prod_{j=2}^{K+1} f_j \end{bmatrix} u \\ &\geq \begin{bmatrix} \prod_{j=1}^{K+1} c_j & -\sum_{j=1}^{K+1} d_j \\ -\sum_{j=1}^{K+1} e_j & \prod_{j=1}^{K+1} f_j \end{bmatrix} u \\ &> 0. \end{aligned}$$

Apply the induction hypothesis to obtain that

$$\begin{bmatrix} c'_j & -d'_j \\ -e'_j & f'_j \end{bmatrix} \cdots \begin{bmatrix} c'_K & -d'_K \\ -e'_K & f'_K \end{bmatrix} u > 0$$

for $j = 1, \ldots, K$, i.e., (23) holds for $j = 2, \ldots, K+1$.

Next prove that (23) holds for $j = 1$. The idea is still to construct some $c', d', e', f' \in \mathbb{R}^K$ and apply the induction hypothesis. The construction is

$$\begin{aligned} c' &= (c_1 c_2,\ c_3,\ \ldots,\ c_{K+1}), \\ d' &= (d_1 + d_2,\ d_3,\ \ldots,\ d_{K+1}), \\ e' &= (e_1 + e_2,\ e_3,\ \ldots,\ e_{K+1}), \\ f' &= (f_1 f_2,\ f_3,\ \ldots,\ f_{K+1}). \end{aligned}$$

Clearly, c', d', e', f' satisfies $0 < c' \leq 1$, $d' \geq 0$, $e' \geq 0$, $0 < f' \leq 1$ componentwise and

$$\begin{bmatrix} \prod_{j=1}^{K} c'_j & -\sum_{j=1}^{K} d'_j \\ -\sum_{j=1}^{K} e'_j & \prod_{j=1}^{K} f'_j \end{bmatrix} u = \begin{bmatrix} \prod_{j=1}^{K+1} c_j & -\sum_{j=1}^{K+1} d_j \\ -\sum_{j=1}^{K+1} e_j & \prod_{j=1}^{K+1} f_j \end{bmatrix} u > 0.$$

Apply the induction hypothesis to obtain

$$\begin{aligned} v'_2 &:= \begin{bmatrix} c'_2 & -d'_2 \\ -e'_2 & f'_2 \end{bmatrix} \cdots \begin{bmatrix} c'_K & -d'_K \\ -e'_K & f'_K \end{bmatrix} u > 0, \\ v'_1 &:= \begin{bmatrix} c'_1 & -d'_1 \\ -e'_1 & f'_1 \end{bmatrix} \cdots \begin{bmatrix} c'_K & -d'_K \\ -e'_K & f'_K \end{bmatrix} u > 0. \end{aligned}$$

It follows that

$$\begin{aligned} &\begin{bmatrix} c_1 & -d_1 \\ -e_1 & f_1 \end{bmatrix} \cdots \begin{bmatrix} c_{K+1} & -d_{K+1} \\ -e_{K+1} & f_{K+1} \end{bmatrix} u \\ &= \begin{bmatrix} c_1 & -d_1 \\ -e_1 & f_1 \end{bmatrix} \begin{bmatrix} c_2 & -d_2 \\ -e_2 & f_2 \end{bmatrix} v'_2 \\ &= \begin{bmatrix} c_1 c_2 + d_1 e_2 & -c_1 d_2 - d_1 f_2 \\ -e_1 c_2 - f_1 e_2 & f_1 f_2 + e_1 d_2 \end{bmatrix} v'_2 \\ &\geq \begin{bmatrix} c_1 c_2 & -d_2 - d_1 \\ -e_1 - e_2 & f_1 f_2 \end{bmatrix} v'_2 \\ &= \begin{bmatrix} c'_1 & -d'_1 \\ -e'_1 & f'_1 \end{bmatrix} v'_2 \\ &= v'_1 \ > \ 0, \end{aligned}$$

i.e., (23) holds for $j = 1$.

To this end, we have proved that (23) holds for $j = 1, \ldots, K+1$, i.e., Lemma 4 also holds for $i = K+1$.

According to i) and ii), Lemma 4 holds for $i \geq 1$. □

Proof of Claim 7. Fix an arbitrary $l \in \mathcal{L}$, and assume $l_k = k$ for $k = 0, \ldots, n_l$ without loss of generality. Fix an arbitrary $t \in \{1, \ldots, n_l\}$, then it suffices to prove that $\underline{A}_s \cdots \underline{A}_{t-1} u_t > 0$ for $s = 1, \ldots, t$. Denote $r_k := r_{k,k-1}$ and $S_k := S_{k,k-1}$ for $k = 1, \ldots, t$ for brevity.

Substitute $(i,j) = (k, k-1)$ in (21) to obtain

$$\begin{bmatrix} \prod_{s=1}^{k-1} \left(1 - \frac{2 r_s \hat{P}_s^+}{\underline{v}_s}\right) & -\sum_{s=1}^{k-1} \frac{2 r_s \hat{Q}_s^+}{\underline{v}_s} \\ -\sum_{s=1}^{k-1} \frac{2 x_s \hat{P}_s^+}{\underline{v}_s} & \prod_{s=1}^{k-1} \left(1 - \frac{2 x_s \hat{Q}_s^+}{\underline{v}_s}\right) \end{bmatrix} \begin{bmatrix} r_k \\ x_k \end{bmatrix} > 0 \tag{24}$$

for $k = 1, \ldots, t$. Hence,

$$\prod_{s=1}^{k-1} \left(1 - \frac{2 r_s \hat{P}_s^+}{\underline{v}_s}\right) r_k > \sum_{s=1}^{k-1} \frac{2 r_s \hat{Q}_s^+(\overline{q})}{\underline{v}_s} x_k \geq 0$$

for $k = 1, \ldots, t$. It follows that $1 - 2 r_k \hat{P}_k^+ / \underline{v}_k > 0$ for $k = 1, \ldots, t-1$. Similarly, $1 - 2 x_k \hat{Q}_k^+ / \underline{v}_k > 0$ for $k = 1, \ldots, t-1$. Then, substitute $k = t$ in (24) and apply Lemma 4 to obtain

$$\begin{bmatrix} 1 - \frac{2 r_s \hat{P}_s^+}{\underline{v}_s} & -\frac{2 r_s \hat{Q}_s^+}{\underline{v}_s} \\ -\frac{2 x_s \hat{P}_s^+}{\underline{v}_s} & 1 - \frac{2 x_s \hat{Q}_s^+}{\underline{v}_s} \end{bmatrix} \cdots \begin{bmatrix} 1 - \frac{2 r_{t-1} \hat{P}_{t-1}^+(\overline{p})}{\underline{v}_{t-1}} & -\frac{2 r_{t-1} \hat{Q}_{t-1}^+(\overline{q})}{\underline{v}_{t-1}} \\ -\frac{2 x_{t-1} \hat{P}_{t-1}^+(\overline{p})}{\underline{v}_{t-1}} & 1 - \frac{2 x_{t-1} \hat{Q}_{t-1}^+(\overline{p})}{\underline{v}_{t-1}} \end{bmatrix} \begin{bmatrix} r_t \\ x_t \end{bmatrix} > 0$$

for $s = 1, \ldots, t$, i.e., $\underline{A}_s \cdots \underline{A}_{t-1} u_t > 0$ for $s = 1, \ldots, t$. This completes the proof of Claim 7. □