Sufficient Conditions for Exact Semidefinite Relaxation of Optimal Power Flow in Unbalanced Multiphase Radial Networks

Fengyu Zhou, Yue Chen, and Steven H. Low

Abstract— This paper proves that in an unbalanced multiphase network with a tree topology, the semidefinite programming relaxation of optimal power flow problems is exact when critical buses are not adjacent to each other. Here a critical bus either contributes directly to the cost function or is where an injection constraint is tight at optimality. Our result generalizes a sufficient condition for exact relaxation in single-phase tree networks to tree networks with arbitrary number of phases.

I. INTRODUCTION

Optimal power flow (OPF) is a mathematical program that minimizes disutility subject to physical laws and other constraints [1]. It was first proposed in [2] and there is a vast literatures on a large number of different solution methods. In general, the OPF problem under alternating current (AC) model is both non-convex and NP-hard [3], [4]. There is thus a strong interest in studying its convexification or approximation; see e.g. a recent survey in [5] on relaxations and approximations of OPF. Using semidefinite programming (SDP) to relax the non-convex constraints was first proposed in [6], [7], and turns out to have good performance in many testcases [8], [9]. Many papers have proposed sufficient conditions under which the SDP relaxation is exact for a single-phase radial network (i.e. network with a tree topology) or its single-phase equivalent of a balanced three-phase network, e.g. [10], [11], [12], [13], [14].

Most radial networks are however unbalanced multiphase, e.g., [15], [16]. SDP relaxation has recently been applied to unbalanced multiphase radial networks [17], [18], [19], [20]. Simulation results in these papers suggest that SDP relaxation is often exact even though no sufficient condition for exact relaxation is known to the best of our knowledge. Indeed, it has been observed in [21], [22], [23] that a multiphase unbalanced network has an equivalent single-phase circuit model where each bus-phase pair in the multiphase network is identified with a single bus in the equivalent model. The single-phase equivalent model is then a meshed network and therefore existing guarantees on exact SDP relaxation are not applicable. Most distribution systems are unbalanced multiphase networks [24] and hence the performance of SDP relaxation of OPF on these networks is important.

Fengyu Zhou is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, 91125. Email: `f.zhou@caltech.edu`

Yue Chen is with the State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, 100084 Beijing, China. Email: `yue-chen15@mails.tsinghua.edu.cn`

Steven H. Low is with the Department of Electrical Engineering and the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, 91125. Email: `slow@caltech.edu`

In this paper, we generalize the sufficient condition for single-phase network proposed in [12] to the multiphase setting. It is shown that when the critical buses or bottleneck buses in a network are non-adjacent, then the SDP relaxation is exact. We prove in this paper the exactness results when the SDP has a unique solution, and state the result for the case of multiple solutions without proof.

II. SYSTEM MODEL

A. Network Structure

We use a similar model as in [18], [19].We assume that all buses have the same number of phases and all generations and loads are Wye-connected. Let the underlying simple undirected graph be $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{0, 1, \ldots, n-1\}$ denotes the set of buses and $\mathcal{E}$ the set of edges. Throughout the paper, we will use (graph, vertex, edge) and (power network, bus, line) interchangeably. Without loss of generality, we let bus 0 be the slack bus where the voltage is specified. Assume all buses have m phases for $m \in \mathbb{Z}^+$. We will use (j, k) and $j \sim k$ interchangeably to denote an edge connecting bus j and k. Consider an m-phase line (j, k) characterized by the admittance matrix $y_{jk} \in \mathbb{C}^{m \times m}$, we assume y_{jk} is invertible. The admittance matrix $\mathbf{Y} \in \mathbb{C}^{mn \times mn}$ for the entire network can be divided into $n \times n$ number of $m \times m$ block matrices. Let $\mathbf{Y}_{jk} \in \mathbb{C}^{m \times m}$ denote the block matrix corresponding to the admittance between bus j and k, then we have

$$\mathbf{Y}_{jj} = \sum_{k:j \sim k} y_{jk}, \ j \in \mathcal{V}$$

$$\mathbf{Y}_{jk} = \begin{cases} -y_{jk} & , j \sim k \\ 0 & , j \not\sim k \end{cases}.$$

For each bus j, let the voltages of all m phases at bus j be the vector $\mathbf{V}_j \in \mathbb{C}^m$. We use $\mathbf{V}_j^\phi$ for $\phi \in \mathcal{M} := \{1, 2, \ldots, m\}$ to indicate the voltage for phase ϕ. Let $\mathbf{V} = [\mathbf{V}_0^\mathsf{T}, \mathbf{V}_1^\mathsf{T}, \ldots, \mathbf{V}_{n-1}^\mathsf{T}]^\mathsf{T}$ be the voltage vector for the entire network. Similarly, we use s_j^ϕ to denote the bus injection for phase ϕ at bus j. Let $\mathbf{e}_j^\phi \in \mathbb{R}^{mn}$ be the base vector which has 1 at the $(jm + \phi)^{\text{th}}$ entry and 0 elsewhere. Let $\mathbf{E}_j^\phi = \mathbf{e}_j^\phi (\mathbf{e}_j^\phi)^\mathsf{T}$, then we define

$$Y_j^\phi := \mathbf{E}_j^\phi \mathbf{Y} \in \mathbb{C}^{mn \times mn}$$

and

$$\mathbf{\Phi}_j^\phi := \frac{1}{2} \big((Y_j^\phi)^\mathsf{H} + Y_j^\phi \big)$$

$$\mathbf{\Psi}_j^\phi := \frac{1}{2\mathrm{i}} \big((Y_j^\phi)^\mathsf{H} - Y_j^\phi \big).$$

Both $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are Hermitian matrices. The relationship between bus voltages and injections can be expressed as

$$\mathrm{Re}(s_j^\phi) = \mathbf{V}^{\mathsf{H}}\mathbf{\Phi}_j^\phi\mathbf{V},$$
$$\mathrm{Im}(s_j^\phi) = \mathbf{V}^{\mathsf{H}}\mathbf{\Psi}_j^\phi\mathbf{V}. \tag{1}$$

B. Optimal Power Flow

Optimal power flow problems minimize certain cost functions subject to constraints involving voltages and injections. Here we consider problems that take the linear combination of bus injections as the cost function and are subject to operational constraints for both voltage magnitudes and real/reactive injections. For problems with nonlinear cost functions, see Section VI. Suppose the bounds $\underline{V}$ and $\overline{V}$ for the voltage magnitudes are always positive and finite, but the bounds for real/reactive injections can be $\pm\infty$ if there are no such constraints.

$$\underset{\mathbf{V},s}{\text{minimize}} \quad \sum_{j,\phi} c_{j,\mathrm{re}}^\phi \mathrm{Re}(s_j^\phi) + c_{j,\mathrm{im}}^\phi \mathrm{Im}(s_j^\phi) \tag{2a}$$

$$\text{subject to} \quad (1) \tag{2b}$$
$$\underline{V}_j^\phi \le |\mathbf{V}_j^\phi| \le \overline{V}_j^\phi, \quad \forall j, \phi \tag{2c}$$
$$\underline{p}_j^\phi \le \mathrm{Re}(s_j^\phi) \le \overline{p}_j^\phi, \quad \forall j, \phi \tag{2d}$$
$$\underline{q}_j^\phi \le \mathrm{Im}(s_j^\phi) \le \overline{q}_j^\phi, \quad \forall j, \phi \tag{2e}$$
$$\mathbf{V}_0 = \mathbf{V}_{\mathrm{ref}} \tag{2f}$$

Here, $\mathbf{V}_{\mathrm{ref}} \in \mathbb{C}^m$ denotes the reference voltage for m phases at the slack bus. Substituting the decision variables s and $\mathbf{V}$ with $\mathbf{W} := \mathbf{V}\mathbf{V}^{\mathsf{H}}$, the following equivalent formulation of (2) is obtained.

$$\underset{\mathbf{W} \ge 0}{\text{minimize}} \quad \mathrm{tr}(\mathbf{C}_0\mathbf{W}) \tag{3a}$$

$$\text{subject to} \quad \underline{v}_j^\phi \le \mathrm{tr}(\mathbf{E}_j^\phi\mathbf{W}) \le \overline{v}_j^\phi, \quad \forall j, \phi \tag{3b}$$
$$\underline{p}_j^\phi \le \mathrm{tr}(\mathbf{\Phi}_j^\phi\mathbf{W}) \le \overline{p}_j^\phi, \quad \forall j, \phi \tag{3c}$$
$$\underline{q}_j^\phi \le \mathrm{tr}(\mathbf{\Psi}_j^\phi\mathbf{W}) \le \overline{q}_j^\phi, \quad \forall j, \phi \tag{3d}$$
$$[\mathbf{W}]_{00} = \mathbf{v}_{\mathrm{ref}} \tag{3e}$$
$$\mathrm{rank}(\mathbf{W}) = 1. \tag{3f}$$

Here, $\underline{v}_j^\phi = |\underline{V}_j^\phi|^2$, $\overline{v}_j^\phi = |\overline{V}_j^\phi|^2$, $\mathbf{v}_{\mathrm{ref}} = \mathbf{V}_{\mathrm{ref}}\mathbf{V}_{\mathrm{ref}}^{\mathsf{H}}$, and $[\mathbf{W}]_{00}$ stands for the upper left $m \times m$ submatrix of $\mathbf{W}$. The cost matrix $\mathbf{C}_0 = \sum_{j,\phi} c_{j,\mathrm{re}}^\phi \mathbf{\Phi}_j^\phi + c_{j,\mathrm{im}}^\phi \mathbf{\Psi}_j^\phi$. Dropping the rank-1 constraint in (3f) yields the semidefinite relaxation.

$$\underset{\mathbf{W} \ge 0}{\text{minimize}} \quad \mathrm{tr}(\mathbf{C}_0\mathbf{W}) \tag{4a}$$
$$\text{subject to} \quad (3b) - (3e). \tag{4b}$$

We use the following exactness definition.

Definition 1: A relaxation problem (4) is *exact* if at least one of its optimal solutions $\mathbf{W}^*$ is of rank 1.
Given a rank-1 solution $\mathbf{W}^*$ of (4), a $\mathbf{V}^*$ can be uniquely determined, which is feasible, and hence optimal, for (3).

We first make the assumption that (4) has a unique optimal solution. In Section VI, we discuss the case when multiple optimal solutions exist.

III. PERTURBATION ANALYSIS

We first study a perturbed version of (4).

A. Perturbed Problem

Fix a nonzero Hermitian matrix $\mathbf{C}_1$, and consider the following perturbed problem for $\varepsilon \ge 0$.

$$\underset{\mathbf{W} \ge 0}{\text{minimize}} \quad \mathrm{tr}((\mathbf{C}_0 + \varepsilon\mathbf{C}_1)\mathbf{W}) \tag{5a}$$
$$\text{subject to} \quad (3b) - (3e). \tag{5b}$$

We say that (5) is *exact* if one of its optimal solution is of rank 1.

Lemma 1: For any nonzero $\mathbf{C}_1$, if there exists a sequence $\{\varepsilon_l\}_{l=1}^\infty$ with $\lim_{l \to \infty} \varepsilon_l = 0$ such that (5) is exact for all ε_l, then (4) is exact.

Proof: Suppose the rank-1 optimal solution to (5) for ε_l is $\mathbf{W}_l$. If the rank-1 optimal solution is non-unique, then pick any one as $\mathbf{W}_l$. As all the $\overline{v}_j^\phi$ are finite, we assume they are upper bounded by a constant α. Hence the constraint (3b) implies all the diagonal elements of $\mathbf{W}$ are upper bounded by α. Since $\mathbf{W}$ is positive semidefinite, the norms of all their entries can be upper bounded by α as well. Consider the set

$$\mathcal{S} = \{\mathbf{W} \ge 0 : (3b) - (3f)\}. \tag{6}$$

The set $\{\mathbf{W} : \mathrm{rank}(\mathbf{W}) = 1\}$ is closed [25] and all other constraints (3b)-(3e) also prescribe closed sets. Further, we have shown that for any $\mathbf{W} \in \mathcal{S}$, its max norm must be upper bounded by α, so $\mathcal{S}$ is compact. The infinite set $\{\mathbf{W}_l\}_{l=1}^\infty$ is a subset in $\mathcal{S}$ and hence has a limit point $\mathbf{W}_{\mathrm{lim}} \in \mathcal{S}$ [26]. For any ε_l, (5) has the same feasible set as (4), and hence the rank-1 matrix $\mathbf{W}_{\mathrm{lim}}$ is also feasible for (4). Next we show that $\mathbf{W}_{\mathrm{lim}}$ is also an optimal point for (4).

If there exists another feasible $\mathbf{W}_{\mathrm{opt}} \ne \mathbf{W}_{\mathrm{lim}}$ such that $\mathrm{tr}(\mathbf{C}_0\mathbf{W}_{\mathrm{lim}}) - \mathrm{tr}(\mathbf{C}_0\mathbf{W}_{\mathrm{opt}}) = \nu > 0$. Clearly $\forall\mathbf{W}$ feasible for (4), $|\mathrm{tr}(\mathbf{C}_1\mathbf{W})| \le m^2n^2\|\mathbf{C}_1\|_\infty\|\mathbf{W}\|_\infty \le m^2n^2\alpha\|\mathbf{C}_1\|_\infty$. For sufficiently large l such that

$$\varepsilon_l < \frac{\nu}{4m^2n^2\alpha\|\mathbf{C}_1\|_\infty}$$
$$\|\mathbf{W}_l - \mathbf{W}_{\mathrm{lim}}\|_\infty < \frac{\nu}{4m^2n^2\|\mathbf{C}_0\|_\infty},$$

we have

$$\mathrm{tr}(\mathbf{C}_0(\mathbf{W}_l - \mathbf{W}_{\mathrm{lim}})) \ge -\frac{\nu}{4} \tag{7a}$$
$$\mathrm{tr}(\varepsilon_l\mathbf{C}_1\mathbf{W}_l) \ge -\frac{\nu}{4} \tag{7b}$$
$$\mathrm{tr}(\mathbf{C}_0\mathbf{W}_{\mathrm{lim}}) = \mathrm{tr}(\mathbf{C}_0\mathbf{W}_{\mathrm{opt}}) + \nu \tag{7c}$$
$$\frac{\nu}{4} \ge \mathrm{tr}(\varepsilon_l\mathbf{C}_1\mathbf{W}_{\mathrm{opt}}) \tag{7d}$$

Summing up (7a)-(7d) gives

$$\mathrm{tr}((\mathbf{C}_0 + \varepsilon_l\mathbf{C}_1)\mathbf{W}_l) > \mathrm{tr}((\mathbf{C}_0 + \varepsilon_l\mathbf{C}_1)\mathbf{W}_{\mathrm{opt}}).$$

contradicting the optimality of $\mathbf{W}_l$ for ε_l. ∎

B. Duality

The dual problem of (5) is as follows.

$$\begin{array}{ll} \underset{\overline{\lambda}_j^\phi, \underline{\lambda}_j^\phi, \overline{\mu}_j^\phi, \underline{\mu}_j^\phi,}{\text{maximize}} & -\sum_{j,\phi}(\overline{\lambda}_j^\phi \overline{v}_j^\phi - \underline{\lambda}_j^\phi \underline{v}_j^\phi + \overline{\mu}_j^\phi \overline{p}_j^\phi - \underline{\mu}_j^\phi \underline{p}_j^\phi + \\ {}_{\overline{\eta}_j^\phi, \underline{\eta}_j^\phi, \kappa} & \qquad \overline{\eta}_j^\phi \overline{q}_j^\phi - \underline{\eta}_j^\phi \underline{q}_j^\phi + \operatorname{tr}(\kappa \mathbf{v}_{\text{ref}})) \end{array} \quad (8a)$$

$$\text{subject to} \quad \overline{\lambda}_j^\phi, \underline{\lambda}_j^\phi, \overline{\mu}_j^\phi, \underline{\mu}_j^\phi, \overline{\eta}_j^\phi, \underline{\eta}_j^\phi \geq 0 \quad (8b)$$

$$\mathbf{A}(\varepsilon) \geq 0. \quad (8c)$$

Dual variables $(\overline{\lambda}_j^\phi, \underline{\lambda}_j^\phi), (\overline{\mu}_j^\phi, \underline{\mu}_j^\phi), (\overline{\eta}_j^\phi, \underline{\eta}_j^\phi)$ and κ correspond to (3b)-(3e) in (5b), respectively. Specifically $\kappa \in \mathbb{C}^{m \times m}$ is Hermitian but not necessarily semidefinite positive. Matrix $\mathbf{A}(\varepsilon)$ denotes

$$\mathbf{A}(\varepsilon) := \sum_{j,\phi}(\overline{\lambda}_j^\phi - \underline{\lambda}_j^\phi)\mathbf{E}_j^\phi + (\overline{\mu}_j^\phi - \underline{\mu}_j^\phi)\mathbf{\Phi}_j^\phi + (\overline{\eta}_j^\phi - \underline{\eta}_j^\phi)\mathbf{\Psi}_j^\phi$$

$$+ \mathbf{C}_0 + \varepsilon\mathbf{C}_1 + \Pi(\kappa) \quad (9)$$

and $\Pi(\kappa)$ is an $mn \times mn$ matrix whose upper left $m \times m$ block is κ and other elements are 0. Note that the upper and lower bounds in (3c) and (3d) could take values of $\pm\infty$. However, since the feasible set prescribed by (5b) is compact, the actual values of $\mathbf{\Phi}_j^\phi \mathbf{W}$ and $\mathbf{\Psi}_j^\phi \mathbf{W}$ are always finite and hence the dual variables associated with such constraints will be 0. These constraints can be removed from (5) and (8). We will use $\overline{\lambda}_j^\phi(\varepsilon)$, $\underline{\lambda}_j^\phi(\varepsilon)$ and so on to denote the Lagrange multipliers for ε. Clearly, when $\varepsilon = 0$, (8) is the dual problem of (4) with $\overline{\lambda}_j^\phi(0)$, $\underline{\lambda}_j^\phi(0)$ and so on as the Lagrange multipliers. If the value of ε is clear in the context, we might denote them simply as $\overline{\lambda}_j^\phi$, $\underline{\lambda}_j^\phi$ and so on for convenience. Let $\mathbf{A}^*(\varepsilon)$ be the dual matrix when dual variables are evaluated at a KKT point.

IV. SUFFICIENT CONDITIONS

The first condition we assume is:

A1: Problem (4) is strictly feasible, i.e., there exists a feasible point such that strict inequality holds in *all* inequality constraints in (3b)-(3e).

Then the Slater's condition is satisfied for both (4) and (5) as they share the same feasible set, and the strong duality between (5) and (8) holds. The KKT condition is necessary and sufficient optimality condition for the primal (5) and the dual (8) problem. In this section, $\mathbf{W}^*$ refers to the unique solution of (4).

A. Notations

The following notations and definitions will be used throughout the rest of the paper.

For each bus-phase pair (j, ϕ), we define

$$f_p(j, \phi) := \left\{ \begin{array}{ll} 0, & \operatorname{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*) \notin \{\overline{p}_j^\phi, \underline{p}_j^\phi\} \\ 1, & \operatorname{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*) = \overline{p}_j^\phi \\ -1, & \operatorname{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*) = \underline{p}_j^\phi \end{array} \right. .$$

The strict feasibility in A1 guarantees that $\overline{p}_j^\phi$ and $\underline{p}_j^\phi$ cannot be attained simultaneously, so the definition above is fully

specified. Similarly we define

$$f_q(j, \phi) := \left\{ \begin{array}{ll} 0, & \operatorname{tr}(\mathbf{\Psi}_j^\phi \mathbf{W}^*) \notin \{\overline{q}_j^\phi, \underline{q}_j^\phi\} \\ 1, & \operatorname{tr}(\mathbf{\Psi}_j^\phi \mathbf{W}^*) = \overline{q}_j^\phi \\ -1, & \operatorname{tr}(\mathbf{\Psi}_j^\phi \mathbf{W}^*) = \underline{q}_j^\phi \end{array} \right. .$$

Definition 2: The *critical objective* bus set is

$$\mathcal{S}_{\text{o}} := \{j \in \mathcal{V} : \exists \phi \text{ s.t. } c_{j,\text{re}}^\phi \neq 0 \text{ or } c_{j,\text{im}}^\phi \neq 0\}.$$

Definition 3: The *critical constraint* bus set is

$$\mathcal{S}_{\text{c}} := \{j \in \mathcal{V} : \exists \phi \text{ s.t. } f_p(j, \phi) \neq 0 \text{ or } f_q(j, \phi) \neq 0\}.$$

For any $mn \times mn$ matrix $\mathbf{X}$, we use $[\mathbf{X}]_{j,k}$ to denote the $m \times m$ block of $\mathbf{X}$ from rows $jm + 1$ to $jm + m$ and from columns $km + 1$ to $km + m$. Further, for $\phi \in \mathcal{M}$, we denote $[\mathbf{X}]_{j,k}^{\phi,:}$ and $[\mathbf{X}]_{j,k}^{:,\phi}$ as the ϕth row and column of $[\mathbf{X}]_{j,k}$, respectively. Similarly, for an mn dimensional vector $\mathbf{x}$, we use $[\mathbf{x}]_j$ to denote the subvector of $\mathbf{x}$ from the $(jm+1)^{\text{th}}$ to $(jm + m)^{\text{th}}$ entry. Denote

$$\Omega(\mathbf{x}) := \{j \in \mathcal{V}, [\mathbf{x}]_j \neq \mathbf{0}\}$$

and we use $|\Omega|$ to denote its cardinality.

We say $\mathcal{V}_1 \subseteq \mathcal{V}$ is *connected* in $\mathcal{G}$ if $\mathcal{G}$ has a connected subgraph whose vertex set is $\mathcal{V}_1$. For any node $j \in \mathcal{V}$, we denote the set of its neighbors in $\mathcal{G}$ as $\mathcal{N}(j)$. For $\mathcal{K} \subseteq \mathcal{V}$, we reload $\mathcal{N}(\mathcal{K}) := \cup_{j \in \mathcal{K}} \mathcal{N}(j)$.

We say a set of real numbers are *sign-semidefinite* if all the non-zero numbers are of the same sign.

B. Main Results

Consider the following conditions.

A2: The underlying graph $\mathcal{G}$ is a tree.

A3: $(\mathcal{S}_{\text{o}} \cup \mathcal{S}_{\text{c}}) \cap \mathcal{N}(\mathcal{S}_{\text{o}} \cup \mathcal{S}_{\text{c}}) = \emptyset$.

A4: $\mathcal{S}_{\text{o}} \cap \mathcal{S}_{\text{c}} = \emptyset$.

A5: For any $j \in \mathcal{S}_{\text{o}} \cap \mathcal{S}_{\text{c}}$ and $\phi \in \mathcal{M}$, $c_{j,\text{re}}^\phi f_p(j, \phi) \geq 0$ and $c_{j,\text{im}}^\phi f_q(j, \phi) \geq 0$.

Informally, A3 means all the critical buses are not adjacent to each other. A5 means if a bus is both critical in objective function and constraints, then for all m phases, $\{c_{j,\text{re}}^\phi, f_p(j, \phi)\}$ and $\{c_{j,\text{im}}^\phi, f_q(j, \phi)\}$ are sign-semidefinite, respectively. The following two theorems provide two sets of sufficient conditions for exact SDP relaxation.

Theorem 1: If conditions A1, A2, A3 and A4 hold, then (4) is exact.

Theorem 2: If conditions A1, A2, A3 and A5 hold, then (4) is exact.

Both theorems rely on strict feasibility, tree structure and critical buses not be adjacent. Theorem 1 needs $\mathcal{S}_{\text{o}}$ and $\mathcal{S}_{\text{c}}$ to be also disjoint. On the other hand, Theorem 2 allows them to intersect, but says for each (j, ϕ) in the intersection, the objective and constraints should encourage its injection to move in the same direction. [1] Since A4 implies A5, Theorem 2 is stronger than Theorem 1. In the next section, we will only provide a proof of Theorem 2.

---

[1] For example, if $\operatorname{Re}(s_j^\phi)$ is minimized in the objective function, then the lower bound of $\operatorname{Re}(s_j^\phi)$ should not be active in the constraints.

One drawback of Theorems 1 and 2 is that the sufficient conditions are given in terms of the optimal solution $\mathbf{W}^*$. The next result provides a sufficient condition that depends only on the primal parameters in (2). Let

$$\tilde{\mathcal{S}}_{\mathrm{c}} := \{j \in \mathcal{V} : \exists \phi \text{ s.t. } \{\pm\infty\} \nsubseteq \{\underline{p}_j^\phi, \overline{p}_j^\phi, \underline{q}_j^\phi, \overline{p}_j^\phi\}\}.$$

Corollary 1: Suppose A1 and A2 hold, If $(\mathcal{S}_{\mathrm{o}} \cup \tilde{\mathcal{S}}_{\mathrm{c}}) \cap \mathcal{N}(\mathcal{S}_{\mathrm{o}} \cup \tilde{\mathcal{S}}_{\mathrm{c}}) = \emptyset$ and $\mathcal{S}_{\mathrm{o}} \cap \tilde{\mathcal{S}}_{\mathrm{c}} = \emptyset$, then (4) is exact.

Proof: As $\mathcal{S}_{\mathrm{c}} \subseteq \tilde{\mathcal{S}}_{\mathrm{c}}$, the conditions in the corollary imply A1–A4 and thus exactness holds. ∎

Informally, Corollary 1 shows that if all the buses involved in the objective function and constraints are not adjacent to each other, then the SDP relaxation is exact.

V. PROOF OF SUFFICIENT CONDITIONS

A. Review

The existing works [12], [13] prove that the optimal solution of SDP relaxation is of rank 1 in single phase networks. A crucial step in their proof uses the strong duality to show that the product of the primal optimal solution $\mathbf{W}^*$ and the dual matrix $\mathbf{A}^*$ is a zero matrix, and hence the rank of $\mathbf{W}^*$ cannot exceed the dimension of $\mathbf{A}^*$'s null space. Under certain conditions [12], [13] prove that $\mathbf{A}^*$'s null space is of dimension at most 1. Hence the optimal primal solution $\mathbf{W}^*$ must be of rank at most 1.

This argument however breaks down in a multiphase network for the following two reasons. First, although the underlying graph for m phase network is still a tree, each bus now has m different phases and might have m unbalanced voltages in general. If we extend each phase to a separate vertex in the new graph and connect every phase pair between every two neighboring buses, then the m phase network will be transformed into an (mn)-node meshed network with multiple cycles [21], [22], [23]. Hence the theory for single-phase radial network is not applicable. Second, in an m phase network, it is unknown wether the null space of $\mathbf{A}^*$ at the optimal point is still of dimension 1. It is therefore not clear how to prove $\mathrm{rank}(\mathbf{W}^*) = 1$ via analyzing the dimension of $\mathrm{null}(\mathbf{A}^*)$.

In the following argument, we use a similar proof framework to that in [12], but the proof will be based on the eigenvectors of $\mathbf{W}^*$ instead of the dimension of $\mathrm{null}(\mathbf{A}^*)$. From now on, we suppose A1, A2, A3 and A5 hold.

B. Preliminaries

Our strategy is to prove the exactness of the perturbed OPF problem and then use Lemma 1 to show (4) is also exact. It is important to make sure that all the non-active constraints will remain non-active in the perturbation neighborhood.

Lemma 2: For any nonzero $\mathbf{C}_1$, there exists a positive sequence $\varepsilon \downarrow 0$ such that for each ε in the sequence, one can collect $(\overline{\mu}_j^\phi(\varepsilon), \underline{\mu}_j^\phi(\varepsilon), \overline{\eta}_j^\phi(\varepsilon), \underline{\eta}_j^\phi(\varepsilon))$ from at least one of

its KKT multiplier tuples satisfying

$$f_p(j, \phi) = 0 \implies \overline{\mu}_j^\phi(\varepsilon) = \underline{\mu}_j^\phi(\varepsilon) = 0 \quad (10a)$$

$$f_p(j, \phi) \neq 0 \implies f_p(j, \phi) \cdot (\overline{\mu}_j^\phi(\varepsilon) - \underline{\mu}_j^\phi(\varepsilon)) \geq 0 \quad (10b)$$

$$f_q(j, \phi) = 0 \implies \overline{\eta}_j^\phi(\varepsilon) = \underline{\eta}_j^\phi(\varepsilon) = 0 \quad (10c)$$

$$f_q(j, \phi) \neq 0 \implies f_q(j, \phi) \cdot (\overline{\eta}_j^\phi(\varepsilon) - \underline{\eta}_j^\phi(\varepsilon)) \geq 0. \quad (10d)$$

Proof: First consider any positive sequence $\{\varepsilon_l\}_{l=1}^\infty$ such that $\lim_{l \to \infty} \varepsilon_l = 0$. Suppose the optimal solution to (5) under ε_l is $\mathbf{W}_l$ (if there are multiple solutions then select one of them). As (5b) prescribes a compact set, using a similar argument as in the proof of Lemma 1 we know there must be a subsequence of $\{\varepsilon_l\}_{l=1}^\infty$, denoted by $\{\varepsilon_{z_t}\}_{t=1}^\infty$, such that $\mathbf{W}_{z_t}$ converges to $\mathbf{W}^*$ in the max norm. The difference $\|\mathbf{W}_{z_t} - \mathbf{W}^*\|_\infty$ can be arbitrarily small for sufficiently large t. When t is large enough, the non-active constraints in (5b) under $\mathbf{W}^*$ will remain non-active under $\mathbf{W}_{z_t}$, and the corresponding KKT multipliers will remain 0. As a result,

$$f_p(j, \phi) = 0 \implies \underline{p}_j^\phi < \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*) < \overline{p}_j^\phi$$
$$\implies \underline{p}_j^\phi < \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}_{z_t}) < \overline{p}_j^\phi \implies \overline{\mu}_j^\phi(\varepsilon_{z_t}) = \underline{\mu}_j^\phi(\varepsilon_{z_t}) = 0,$$

$$f_p(j, \phi) = +1 \implies \underline{p}_j^\phi < \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*)$$
$$\implies \underline{p}_j^\phi < \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}_{z_t}) \implies \underline{\mu}_j^\phi(\varepsilon_{z_t}) = 0$$
$$\implies f_p(j, \phi) \cdot (\overline{\mu}_j^\phi(\varepsilon_{z_t}) - \underline{\mu}_j^\phi(\varepsilon_{z_t})) \geq 0,$$

$$f_p(j, \phi) = -1 \implies \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}^*) < \overline{p}_j^\phi$$
$$\implies \mathrm{tr}(\mathbf{\Phi}_j^\phi \mathbf{W}_{z_t}) < \overline{p}_j^\phi \implies \overline{\mu}_j^\phi(\varepsilon_{z_t}) = 0$$
$$\implies f_p(j, \phi) \cdot (\overline{\mu}_j^\phi(\varepsilon_{z_t}) - \underline{\mu}_j^\phi(\varepsilon_{z_t})) \geq 0$$

all hold. A similar argument can also be applied to prove (10c) and (10d). ∎

C. Properties of Dual Matrix $\mathbf{A}^*(\varepsilon)$

In order to apply Lemma 1, we construct $\mathbf{C}_1 \in \mathbb{C}^{mn \times mn}$ in the following manner.

$$[\mathbf{C}_1]_{jj} = \mathbf{0} \in \mathbb{C}^{m \times m}, \quad \text{for } j \in \mathcal{V}$$
$$[\mathbf{C}_1]_{jk} = \mathbf{0} \in \mathbb{C}^{m \times m}, \quad \text{for } (j, k) \notin \mathcal{E}$$

When $(j, k) \in \mathcal{E}$, we assume $j < k$. If neither j nor k is in $\mathcal{S}_{\mathrm{o}} \cup \mathcal{S}_{\mathrm{c}}$, then we construct $[\mathbf{C}_1]_{jk} = \mathbf{Y}_{jk}$.

If $j \in \mathcal{S}_{\mathrm{o}} \cup \mathcal{S}_{\mathrm{c}}$, then A3 guarantees $k \notin \mathcal{S}_{\mathrm{o}} \cup \mathcal{S}_{\mathrm{c}}$. $\forall \phi \in \mathcal{M}$, we set $[\mathbf{C}_1]_{jk}^{\phi,:}$ to $\mathbf{Y}_{jk}^{\phi,:}$ if $c_{j,\mathrm{re}}^\phi = c_{j,\mathrm{im}}^\phi = f_p(j, \phi) = f_q(j, \phi) = 0$, and to $(f_p(j, \phi) + f_q(j, \phi)\mathrm{i})\mathbf{Y}_{jk}^{\phi,:}$ otherwise.

If $k \in \mathcal{S}_{\mathrm{o}} \cup \mathcal{S}_{\mathrm{c}}$, then A3 guarantees $j \notin \mathcal{S}_{\mathrm{o}} \cup \mathcal{S}_{\mathrm{c}}$. $\forall \phi \in \mathcal{M}$, we similarly set $[\mathbf{C}_1]_{jk}^{:,\phi}$ to $(\mathbf{Y}_{kj}^{\phi,:})^{\mathsf{H}}$ if $c_{k,\mathrm{re}}^\phi = c_{k,\mathrm{im}}^\phi = f_p(k, \phi) = f_q(k, \phi) = 0$, and to $(f_p(k, \phi) - f_q(k, \phi)\mathrm{i})(\mathbf{Y}_{kj}^{\phi,:})^{\mathsf{H}}$ otherwise.

Finally, we set $[\mathbf{C}_1]_{kj} := [\mathbf{C}_1]_{jk}^{\mathsf{H}}$ for all $j < k$ to make $\mathbf{C}_1$ Hermitian.

Definition 4: An $mn \times mn$ positive semidefinite matrix $\mathbf{X}$ is $\mathcal{G}$-invertible for some graph $\mathcal{G}$ if the following two

conditions hold:

1) $\forall (a,b) \in \mathcal{E}$, $[\mathbf{X}]_{ab}$ is invertible.
2) $\forall a,b \in \mathcal{V}$ such that $a \neq b$ and $(a,b) \notin \mathcal{E}$, $[\mathbf{X}]_{ab}$ is all zero.

The next theorem provides a key intermediate result to prove Theorem 2. Suppose under such $\mathbf{C}_1$, the sequence guaranteed by Lemma 2 is $\{\varepsilon_l\}_{l=1}^{\infty}$.

Theorem 3: Under A1, A2, A3 and A5, for each ε_l, the dual matrix $\mathbf{A}^*(\varepsilon_l)$ is $\mathcal{G}$-invertible. [2]

Proof: The value of $\mathbf{A}^*(\varepsilon_l)$ is the same as the right hand side of (9) when all dual variables take values at their corresponding KKT multipliers (with respect to ε_l). If not otherwise specified, all the $(\overline{\mu}_j^\phi, \underline{\mu}_j^\phi, \overline{\eta}_j^\phi, \underline{\eta}_j^\phi)$ in this proof refer to the tuple in Lemma 2 with respect to ε_l. Since for all $a \neq b$, $[\mathbf{E}_j^\phi]_{ab}$ and $[\Pi(\kappa)]_{ab}$ are always zero matrices, it is sufficient to show

$$\mathbf{Q} := \sum_{j,\phi} \left((\overline{\mu}_j^\phi - \underline{\mu}_j^\phi)\mathbf{\Phi}_j^\phi + (\overline{\eta}_j^\phi - \underline{\eta}_j^\phi)\mathbf{\Psi}_j^\phi \right) + \mathbf{C}_0 + \varepsilon_l \mathbf{C}_1$$

satisfies the two conditions in Definition 4.[3]

For $a \neq b$ and $(a,b) \notin \mathcal{E}$, recall that $\mathbf{C}_0$ is the linear combination of $\mathbf{\Phi}_j^\phi$ and $\mathbf{\Psi}_j^\phi$. When $(a,b) \notin \mathcal{E}$, $\mathbf{Y}_{ab}$ is a zero matrix and so are all $[\mathbf{\Phi}_j^\phi]_{ab}$ and $[\mathbf{\Psi}_j^\phi]_{ab}$. The construction of $\mathbf{C}_1$ also guarantees $[\mathbf{C}_1]_{ab}$ is all zero. Hence $[\mathbf{Q}]_{ab}$ is all zero as well.

Now assume $a < b$. If $(a,b) \in \mathcal{E}$, we have

$$[\mathbf{Q}]_{ab}$$
$$= \sum_{\phi} \left((\overline{\mu}_a^\phi - \underline{\mu}_a^\phi + c_{a,\text{re}}^\phi)[\mathbf{\Phi}_a^\phi]_{ab} + (\overline{\eta}_a^\phi - \underline{\eta}_a^\phi + c_{a,\text{im}}^\phi)[\mathbf{\Psi}_a^\phi]_{ab} \right)$$
$$+ \sum_{\phi} \left((\overline{\mu}_b^\phi - \underline{\mu}_b^\phi + c_{b,\text{re}}^\phi)[\mathbf{\Phi}_b^\phi]_{ab} + (\overline{\eta}_b^\phi - \underline{\eta}_b^\phi + c_{b,\text{im}}^\phi)[\mathbf{\Psi}_b^\phi]_{ab} \right)$$
$$+ \varepsilon_l[\mathbf{C}_1]_{ab}. \tag{11}$$

If neither a nor b is in $\mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$, then by definition, for all $\phi \in \mathcal{M}$ there must be

$$c_{a,\text{re}}^\phi = c_{a,\text{im}}^\phi = f_p(a,\phi) = f_q(a,\phi) = 0, \tag{12a}$$
$$c_{b,\text{re}}^\phi = c_{b,\text{im}}^\phi = f_p(b,\phi) = f_q(b,\phi) = 0. \tag{12b}$$

Equation (11) and Lemma 2 imply $[\mathbf{Q}]_{ab} = \varepsilon_l[\mathbf{C}_1]_{ab}$. By construction, $[\mathbf{C}_1]_{ab} = \mathbf{Y}_{ab}$ is invertible, and so is $[\mathbf{Q}]_{ab}$.

If $a \in \mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$, then A3 guarantees $b \notin \mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$. Thus (12b) holds for all $\phi \in \mathcal{M}$. For a given $\phi \in \mathcal{M}$, if (12a) holds, then by construction, we have $[\mathbf{Q}]_{ab}^{\phi,:} = \varepsilon_l[\mathbf{C}_1]_{ab}^{\phi,:} = \varepsilon_l \mathbf{Y}_{ab}^{\phi,:}$. If (12a) does not hold for the given ϕ, then we have

$$[\mathbf{Q}]_{ab}^{\phi,:} = (\overline{\mu}_a^\phi - \underline{\mu}_a^\phi + c_{a,\text{re}}^\phi + 2\varepsilon_l f_p(a,\phi))\frac{\mathbf{Y}_{ab}^{\phi,:}}{2}$$
$$+ (\overline{\eta}_a^\phi - \underline{\eta}_a^\phi + c_{a,\text{im}}^\phi + 2\varepsilon_l f_q(a,\phi))\frac{\mathbf{Y}_{ab}^{\phi,:}}{2}\mathrm{i}.$$

Note that Condition A5 and Lemma 2 imply both $\{\overline{\mu}_a^\phi - \underline{\mu}_a^\phi, f_p(a,\phi), c_{a,\text{re}}^\phi\}$ and $\{\overline{\eta}_a^\phi - \underline{\eta}_a^\phi, f_q(a,\phi), c_{a,\text{im}}^\phi\}$ are sign-

semidefinite sets, respectively. When (12a) does not hold, at least one of $\{c_{a,\text{re}}^\phi, c_{a,\text{im}}^\phi, f_p(a,\phi), f_q(a,\phi)\}$ is non-zero. As a result, there exists some non-zero $\sigma_{ab}^{\phi,:} \in \mathbb{C}$ such that $[\mathbf{Q}]_{ab}^{\phi,:} = \sigma_{ab}^{\phi,:}\mathbf{Y}_{ab}^{\phi,:}$. In short, in the case $a \in \mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$, $[\mathbf{Q}]_{ab}^{\phi,:}$ is always a non-zero multiple of $\mathbf{Y}_{ab}^{\phi,:}$. The invertibility of $\mathbf{Y}_{ab}$ indicates all the $\mathbf{Y}_{ab}^{\phi,:}$ are independent for $\phi \in \mathcal{M}$, so $[\mathbf{Q}]_{ab}$ is also invertible.

If $b \in \mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$, then A3 guarantees $a \notin \mathcal{S}_\text{o} \cup \mathcal{S}_\text{c}$. Then (12a) holds for all $\phi \in \mathcal{M}$. For a given $\phi \in \mathcal{M}$, if (12b) holds, then by construction, we have $[\mathbf{Q}]_{ab}^{:,\phi} = \varepsilon_l[\mathbf{C}_1]_{ab}^{:,\phi} = \varepsilon_l(\mathbf{Y}_{ba}^{\phi,:})^\mathsf{H}$. If (12b) does not hold, then similar to the previous case, there exists some non-zero $\sigma_{ab}^{:,\phi} \in \mathbb{C}$ such that $[\mathbf{Q}]_{ab}^{:,\phi} = \sigma_{ab}^{:,\phi}(\mathbf{Y}_{ba}^{\phi,:})^\mathsf{H}$. Hence $[\mathbf{Q}]_{ab}^{:,\phi}$ is always a non-zero multiple of $(\mathbf{Y}_{ba}^{\phi,:})^\mathsf{H}$. The invertibility of $\mathbf{Y}_{ba}$ indicates all the $\mathbf{Y}_{ba}^{\phi,:}$ are independent for $\phi \in \mathcal{M}$, so $[\mathbf{Q}]_{ab}$ is also invertible. $\blacksquare$

The next theorem is a generalization of Theorem 3.3 in [27]. While [27] studies the matrices whose non-zero off-diagonal entries correspond to an edge in $\mathcal{G}$, we extend the results to $\mathcal{G}$-invertible matrices.

Theorem 4: Let $\mathbf{y} \in \mathbb{C}^{mn}$ be a non-zero vector with the smallest $|\Omega(\mathbf{y})|$ satisfying $\mathbf{X}\mathbf{y} = \mathbf{0}$, where $\mathbf{X}$ is $\mathcal{G}$-invertible. Then $\Omega(\mathbf{y})$ is connected in $\mathcal{G}$.

Proof: If not, then assume $\Omega(\mathbf{y}) = \Omega_1 \cup \Omega_2$ where non-empty sets Ω_1 and Ω_2 are not connected in $\mathcal{G}$. Construct $\tilde{\mathbf{y}}$ in the following manner:

$$[\tilde{\mathbf{y}}]_k = \begin{cases} [\mathbf{y}]_k & , k \notin \Omega_2 \\ \mathbf{0} & , k \in \Omega_2 \end{cases}.$$

Then for each $j \in \Omega_1$,

$$[\mathbf{X}\tilde{\mathbf{y}}]_j = \sum_{k \in \mathcal{V}} [\mathbf{X}]_{jk}[\tilde{\mathbf{y}}]_k = [\mathbf{X}]_{jj}[\tilde{\mathbf{y}}]_j + \sum_{k:k\sim j} [\mathbf{X}]_{jk}[\tilde{\mathbf{y}}]_k$$
$$= [\mathbf{X}]_{jj}[\mathbf{y}]_j + \sum_{k:k\sim j} [\mathbf{X}]_{jk}[\mathbf{y}]_k = [\mathbf{X}\mathbf{y}]_j = \mathbf{0}.$$

The third equality above is due to the fact that $j \in \Omega_1$ is not connected to any nodes in Ω_2. Therefore,

$$\tilde{\mathbf{y}}^\mathsf{H}\mathbf{X}\tilde{\mathbf{y}} = \sum_{j \in \mathcal{V}} [\tilde{\mathbf{y}}]_j^\mathsf{H}[\mathbf{X}\tilde{\mathbf{y}}]_j = \sum_{j \in \Omega_1} [\tilde{\mathbf{y}}]_j^\mathsf{H}[\mathbf{X}\tilde{\mathbf{y}}]_j + \sum_{j \notin \Omega_1} [\tilde{\mathbf{y}}]_j^\mathsf{H}[\mathbf{X}\tilde{\mathbf{y}}]_j$$
$$= \sum_{j \in \Omega_1} [\tilde{\mathbf{y}}]_j^\mathsf{H}\mathbf{0} + \sum_{j \notin \Omega_1} \mathbf{0}^\mathsf{H}[\mathbf{X}\tilde{\mathbf{y}}]_j = \mathbf{0}.$$

Since $\mathcal{G}$-invertibility implies $\mathbf{X} \geq 0$, there must be $\mathbf{X}\tilde{\mathbf{y}} = \mathbf{0}$ as well. As $|\Omega(\tilde{\mathbf{y}})| = |\Omega_1| < |\Omega(\mathbf{y})|$ and $\tilde{\mathbf{y}}$ is non-zero by construction, it contradicts the minimality of $|\Omega(\mathbf{y})|$. $\blacksquare$

D. Proof of Theorem 2

We now prove that (4) is exact under conditions A1, A2, A3 and A5. According to Lemma 1, we only need to show (5) is exact for any ε_l in the sequence $\{\varepsilon_l\}_{l=1}^{\infty}$ used in Theorem 3. If (5) is not exact, then there exists an optimal solution $\mathbf{W}^*$ such that $\text{rank}(\mathbf{W}^*) \geq 2$. [4] Note that in this subsection, $\mathbf{W}^*$ stands for the optimal solution to (5).

---

[2]If the KKT multiplier tuple at ε_l is non-unique, then $\mathbf{A}^*(\varepsilon_l)$ is evaluated at the multiplier tuple in Lemma 2 satisfying (10).

[3]The matrix $\mathbf{Q}$ itself might not be $\mathcal{G}$-invertible as $\mathbf{Q}$ might not be positive semidefinite, but $\mathbf{A}^* \geq 0$ always hold.

[4]Note that $\text{rank}(\mathbf{W}^*)$ cannot be 0 as the constraint $[\mathbf{W}^*]_{00} = \mathbf{v}_{\text{ref}}$ requires $\mathbf{W}^*$ to be a non-zero matrix.

Suppose the eigen-decomposition of $\mathbf{W}^*$ is

$$\mathbf{W}^* = \sum_{l=1}^{mn} \varrho_l \mathbf{u}_l \mathbf{u}_l^{\mathsf{H}}$$

where $\varrho_1 \geq \varrho_2 \geq \ldots \varrho_{mn} \geq 0$ are $\mathbf{W}^*$'s eigenvalues in decreasing order and $\mathbf{u}_l$ is the eigenvector associated with ϱ_l. All the $\mathbf{u}_l$ are non-zero and orthogonal. As $\mathrm{rank}(\mathbf{W}^*) \geq 2$, we have $\varrho_2 > 0$. Now let $2 \leq L \leq mn$ be the largest number such that $\varrho_L > 0$, then we have

$$\mathbf{V}_{\mathrm{ref}}\mathbf{V}_{\mathrm{ref}}^{\mathsf{H}} = [\mathbf{W}^*]_{00} = \sum_{l=1}^{L} \varrho_l [\mathbf{u}_l]_0 [\mathbf{u}_l]_0^{\mathsf{H}} =: \mathbf{U}\mathbf{U}^{\mathsf{H}},$$

$$\mathbf{U} := \left[\sqrt{\varrho_1}[\mathbf{u}_1]_0, \sqrt{\varrho_2}[\mathbf{u}_2]_0, \ldots, \sqrt{\varrho_L}[\mathbf{u}_L]_0 \right].$$

If the rank of $\mathbf{U}$ is strictly greater than 1, then we can find $\mathbf{z} \in \mathrm{span}(\mathbf{U})$ such that $\mathbf{z}^{\mathsf{H}}\mathbf{V}_{\mathrm{ref}} = 0$. Then $\mathbf{U}^{\mathsf{H}}\mathbf{z} \neq 0$ implies

$$0 = \mathbf{z}^{\mathsf{H}}\mathbf{V}_{\mathrm{ref}}\mathbf{V}_{\mathrm{ref}}^{\mathsf{H}}\mathbf{z} = \mathbf{z}^{\mathsf{H}}\mathbf{U}\mathbf{U}^{\mathsf{H}}\mathbf{z} > 0.$$

The contradiction means $\mathrm{rank}(\mathbf{U}) \leq 1$, and therefore $[\mathbf{u}_1]_0$ and $[\mathbf{u}_2]_0$ are linearly dependent. If $[\mathbf{u}_1]_0 = r[\mathbf{u}_2]_0$ for some $r \in \mathbb{C}$, then we construct $\tilde{\mathbf{u}} = \mathbf{u}_1 - r\mathbf{u}_2$. Otherwise $[\mathbf{u}_2]_0$ must be zero and we construct $\tilde{\mathbf{u}} = \mathbf{u}_2$. Clearly we have

$$\tilde{\mathbf{u}} \neq 0, [\tilde{\mathbf{u}}]_0 = \mathbf{0}. \tag{13}$$

On the other hand, KKT conditions give $\mathrm{tr}(\mathbf{A}^*\mathbf{W}^*) = 0$. As both $\mathbf{A}^*$ and $\mathbf{W}^*$ are positive semidefinite, we have

$$0 = \mathrm{tr}(\mathbf{A}^*\mathbf{W}^*) = \mathrm{tr}\left(\mathbf{A}^* \sum_{l=1}^{L} \varrho_l \mathbf{u}_l \mathbf{u}_l^{\mathsf{H}}\right)$$
$$= \sum_{l=1}^{L} \mathrm{tr}\left(\varrho_l \mathbf{A}^* \mathbf{u}_l \mathbf{u}_l^{\mathsf{H}}\right) = \sum_{l=1}^{L} \mathrm{tr}\left(\varrho_l \mathbf{u}_l^{\mathsf{H}} \mathbf{A}^* \mathbf{u}_l\right) \geq 0.$$

The equality holds only when $\mathbf{A}^*\mathbf{u}_l = \mathbf{0}$ for all $l \leq L$. Hence

$$\mathbf{A}^*\tilde{\mathbf{u}} = \mathbf{0}. \tag{14}$$

As (13) has shown $1 \leq |\Omega(\tilde{\mathbf{u}})| \leq n-1$, putting together Theorem 3, Theorem 4 and (14) implies that there exists $\hat{\mathbf{u}}$ such that $\Omega(\hat{\mathbf{u}})$ is non-empty, connected in $\mathcal{G}$, $1 \leq |\Omega(\hat{\mathbf{u}})| \leq n-1$, and $\mathbf{A}^*\hat{\mathbf{u}} = \mathbf{0}$. Let j be a node not in $\Omega(\hat{\mathbf{u}})$ but is connected to some node $k \in \Omega(\hat{\mathbf{u}})$. Since A2 requires $\mathcal{G}$ to be a tree and $\Omega(\hat{\mathbf{u}})$ is connected in $\mathcal{G}$, k must be the only node in $\Omega(\hat{\mathbf{u}})$ which is connected to $j \notin \Omega(\hat{\mathbf{u}})$. Otherwise there is a cycle. Then

$$[\mathbf{A}^*\hat{\mathbf{u}}]_j = \sum_{l \in \mathcal{V}} [\mathbf{A}^*]_{jl}[\hat{\mathbf{u}}]_l = [\mathbf{A}^*]_{jj}[\hat{\mathbf{u}}]_j + \sum_{l:l\sim j}[\mathbf{A}^*]_{jl}[\hat{\mathbf{u}}]_l$$
$$= [\mathbf{A}^*]_{jj}\mathbf{0} + [\mathbf{A}^*]_{jk}[\hat{\mathbf{u}}]_k + \sum_{l:l\sim j, l\notin\Omega(\hat{\mathbf{u}})}[\mathbf{A}^*]_{jl}[\hat{\mathbf{u}}]_l.$$

As $[\hat{\mathbf{u}}]_l = \mathbf{0}$ for $l \notin \Omega(\hat{\mathbf{u}})$, we have $[\mathbf{A}^*\hat{\mathbf{u}}]_j = [\mathbf{A}^*]_{jk}[\hat{\mathbf{u}}]_k$. Further, $(j,k) \in \mathcal{E}$ and the $\mathcal{G}$-invertibility of $\mathbf{A}^*$ implies $[\mathbf{A}^*]_{jk}$ is invertible. Node k is in $\Omega(\hat{\mathbf{u}})$ implies $[\hat{\mathbf{u}}]_k \neq \mathbf{0}$. As a result, $[\mathbf{A}^*\hat{\mathbf{u}}]_j = [\mathbf{A}^*]_{jk}[\hat{\mathbf{u}}]_k$ must be non-zero, contracting $\mathbf{A}^*\hat{\mathbf{u}} = \mathbf{0}$. This implies that (5) is exact. Theorem 2 is proved. ∎
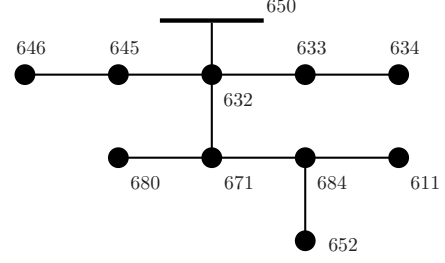

Fig. 1. An 11 bus network revised from IEEE 13 node test feeder. The switch in the original system is assumed to be open so 2 buses are removed.

VI. DISCUSSION AND EXAMPLE

A. Discussion

The main results in this paper are Theorem 1, Theorem 2 and Corollary 1. They provide three sets of sufficient conditions under which the SDP relaxation for unbalanced multiphase network is exact. These results have different interpretations and implications.

Sufficient conditions in Corollary 1 do not rely on the optimal solution of SDP relaxation, and can be checked *a priori*. Though these conditions are still restrictive in practice, we hope this result can stimulate more work on unbalanced multiphase networks.

Conditions in Theorems 1 and 2 rely on knowing the active constraints at the optimal point, which cannot be checked a priori. Nevertheless, the actual value of the optimal point is not involved as long as one knows where the bottlenecks are. These conditions also suggest that relaxation is more likely to be exact if critical buses turn out to be spread over the network rather than concentrated in some neighborhood

So far we have assumed that (4) has a unique optimal solution so that inactive constraints at the optimal solution of (4) remain inactive under a small perturbation. If (4) has multiple solutions, A4 and A5 in Theorems 1 and 2 and condition $\mathcal{S}_{\mathrm{o}} \cap \tilde{\mathcal{S}}_{\mathrm{c}} = \emptyset$ in Corollary 1 need to be replaced by the linear separability condition proposed in [12]. The proof will be similar to that in this paper.

To generalize the result here to nonlinear cost functions, note that the proposed conditions involving the cost function only rely on the signs of $c_{j,\mathrm{re}}^{\phi}$ and $c_{j,\mathrm{im}}^{\phi}$. The same argument in this paper can be extended to the nonlinear case when the cost function is convex, monotonic and additively separable in injections.

B. Illustrative Example

We use an 11 bus radial network shown in Fig. 1, adapted from IEEE 13 node test feeder, to illustrate our theoretical result. The line configuration is reassigned and noise is added to the admittance matrix, so all the buses have three complete phases and each y_{jk} is invertible. For illustrative purpose, all the real/reactive injections are bounded from at most one direction. Table I summarizes our setup. The '+' and '-' refer to the sign of $c_{j,\mathrm{re}}^{\phi}$ or $c_{j,\mathrm{im}}^{\phi}$ in the cost function. For constraints, 'u' (or 'l') means the upper (or lower) bound for

TABLE I

Buses	650			632			633			634			645			646			671			684			611			652			680		
Phases	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
Objective (real)	+	+	+				-	-	-				+	+	+				-	-	-				+	+	+	+	+	+			
Objective (reactive)	+	+	+				-	-	-				-	-	-				+	+	+				+	+	+						
Constraints (real)	u	u	u				1	1	1				u	u	u				1	1	1				u	u	u						
Constraints (reactive)	u	u	u				1	1	1				1	1	1				u	u	u				u	u	u						

the corresponding injection is finite. It is easy to check that no matter which constraints are active at the optimal point, conditions A1, A2, A3 and A5 must hold, so Theorem 2 implies the optimal solution is of rank 1.

After solving the problem, there are actually nine active constraints, highlighted in light red in Table I. The largest two eigenvalues of the resulting optimal solution $\mathbf{W}^*$ are 36.90 and 1.44×10^{-10}, respectively. It confirms that $\mathbf{W}^*$ is indeed rank 1 up to numerical precision.

Finally, we refer to [18] for more simulation results, which show that semidefinite relaxation is also exact for IEEE 13, 37, 123-bus networks and a real-world 2065-bus network. In the simulation of [18], our sufficient conditions are actually violated since the cost function is set as

$$\sum_{j \in \mathcal{V}} \sum_{\phi \in \mathcal{M}} \mathrm{Re}(s_j^\phi).$$

It means even when all the buses are critical, the semidefinite relaxation can still be exact.

VII. CONCLUSION

We have proposed sufficient conditions for exact SDP relaxation in unbalanced multiphase radial networks. These conditions suggest that having critical buses not adjacent to each other encourages exact relaxation.

REFERENCES

[1] J. A. Momoh, R. Adapa, and M. El-Hawary, "A review of selected optimal power flow literature to 1993. I. nonlinear and quadratic programming approaches," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 96–104, 1999.

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

[3] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," *arXiv:1512.07315*, Dec. 2015.

[4] K. Lehmann, A. Grastien, and P. Van Hentenryck, "AC-feasibility on tree networks is NP-hard," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, 2016.

[5] D. K. Molzahn, I. A. Hiskens *et al.*, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

[6] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, 2006.

[7] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int'l J. of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 383–392, 2008.

[8] 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.

[9] B. C. Lesieutre, D. K. Molzahn, A. R. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *2011 49th annual Allerton conference on communication, control, and computing (Allerton)*. IEEE, 2011, pp. 1492–1499.

[10] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification–part II," *IEEE Transactions on Power Systems*, vol. 3, no. 28, pp. 2565–2572, 2013.

[11] B. Zhang and D. Tse, "Geometry of injection regions of power networks," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 788–797, 2013.

[12] S. Bose, D. F. Gayme, K. M. Chandy, and S. H. Low, "Quadratically constrained quadratic programs on acyclic graphs with application to power flow," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 278–287, 2015.

[13] S. Sojoudi and J. Lavaei, "Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure," *SIAM Journal on Optimization*, vol. 24, no. 4, pp. 1746–1778, 2014.

[14] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Transactions on Automatic Control*, vol. 60, no. 1, pp. 72–87, 2015.

[15] M. Srinivas, "Distribution load flows: a brief review," in *2000 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 00CH37077)*, vol. 2. IEEE, 2000, pp. 942–945.

[16] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2006.

[17] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.

[18] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *2014 Power Systems Computation Conference*. IEEE, 2014, pp. 1–9.

[19] C. Zhao, E. Dall'Anese, and S. H. Low, "Convex relaxation of OPF in multiphase radial networks with delta connections," in *Proceedings of the 10th Bulk Power Systems Dynamics and Control Symposium*, 2017.

[20] A. S. Zamzam, N. D. Sidiropoulos, and E. Dall'Anese, "Beyond relaxation and Newton–Raphson: Solving AC OPF for multi-phase systems with renewables," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 3966–3975, 2018.

[21] J. R. Berg, E. S. Hawkins, and W. W. Pleines, "Mechanized calculation of unbalanced load flow on radial distribution circuits," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 4, pp. 451–421, April 1967.

[22] M. Laughton, "Analysis of unbalanced polyphase networks by the method of phase co-ordinates. Part 1: System representation in phase frame of reference," in *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8. IET, 1968, pp. 1163–1172.

[23] T.-H. Chen, M.-S. Chen, K.-J. Hwang, P. Kotas, and E. A. Chebli, "Distribution system power flow analysis – a rigid approach," *IEEE Transactions on Power Delivery*, vol. 6, no. 3, July 1991.

[24] K. Christakou, J.-Y. LeBoudec, M. Paolone, and D.-C. Tomozei, "Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 741–750, 2013.

[25] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.

[26] W. Rudin *et al.*, *Principles of mathematical analysis*. McGraw-hill New York, 1964, vol. 3.

[27] H. van der Holst, "Graphs whose positive semi-definite matrices have nullity at most two," *Linear Algebra and its Applications*, vol. 375, pp. 1–11, 2003.