# Exactness of OPF Relaxation on Three-phase Radial Networks with Delta Connections

Fengyu Zhou, *Student Member, IEEE,* Ahmed S. Zamzam, *Member, IEEE,* Steven H. Low, *Fellow, IEEE,*
and Nicholas D. Sidiropoulos, *Fellow, IEEE*

*Abstract*—Simulations have shown that while semi-definite relaxations of AC optimal power flow (AC-OPF) on three-phase radial networks with only wye connections tend to be exact, the presence of delta connections seem to render them inexact. This paper shows that the such inexactness originates from the non-uniqueness of relaxation solutions and numerical errors amplified by the non-uniqueness. This finding motivates two algorithms to recover the exact solution of AC-OPF in the presence of delta connections. In simulations using IEEE 13, 37 and 123-bus systems, the proposed algorithms provide exact optimal solutions up to numerical precision.

*Index Terms*—Optimal power flow, semi-definite relaxation, distribution networks, delta connected devices.

## Nomenclature

| | |
|---|---|
| $\mathcal{G}(\mathcal{V}, \mathcal{E})$ | Graph with vertices $\mathcal{V}$ and edges $\mathcal{E}$. |
| $z_{jk}, y_{jk}$ | Impedance and admittance matrices for line $(j, k)$. |
| $\mathbf{V}_j, \mathbf{V}_j^\phi$ | Complex voltage at bus $j$ (phase $\phi$). |
| $\underline{\mathbf{V}}, \overline{\mathbf{V}}$ | Limits on voltage magnitude, which may be set as $\underline{\mathbf{V}} = \underline{V}\mathbf{1}$ and $\overline{\mathbf{V}} = \overline{V}\mathbf{1}$. |
| $\mathbf{s}_j, \mathbf{s}_j^\phi$ | Wye injection at bus $j$ (phase $\phi$). |
| $\mathbf{p}_j, \mathbf{q}_j$ | Active and reactive wye injection at bus $j$, i.e., $\mathbf{s}_j = \mathbf{p}_j + \mathrm{i}\mathbf{q}_j$. |
| $\mathbf{s}_{\Delta,j}, \mathbf{s}_{\Delta,j}^\phi$ | Delta injection at bus $j$ (phase $\phi$). |
| $\mathbf{p}_{\Delta,j}, \mathbf{q}_{\Delta,j}$ | Active and reactive delta injection at bus $j$, i.e., $\mathbf{s}_{\Delta,j} = \mathbf{p}_{\Delta,j} + \mathrm{i}\mathbf{q}_{\Delta,j}$. |
| $f(\mathbf{s}, \mathbf{s}_\Delta)$ | Cost function of AC-OPF. |
| $\mathcal{S}_j, \mathcal{S}_{\Delta,j}$ | Feasible sets for $\mathbf{s}_j$ and $\mathbf{s}_{\Delta,j}$. |
| $\mathbf{I}_{jk}$ | Sending-end currents from $j$ to $k$. |
| $\mathbf{I}_{\Delta,j}$ | Delta line currents at bus $j$. |
| $\mathbf{S}_{jk}$ | Sending-end branch power flow, i.e., $\mathbf{V}_j\mathbf{I}_{jk}^\mathsf{H}$. |
| $\mathbf{X}_j$ | Power flow on delta devices, i.e., $\mathbf{V}_j\mathbf{I}_{\Delta,j}^\mathsf{H}$. |
| $\mathbf{W}, \mathbf{v}_j$ | $\mathbf{W} = \mathbf{V}\mathbf{V}^\mathsf{H}$, $\mathbf{v}_j = \mathbf{V}_j\mathbf{V}_j^\mathsf{H}$. |
| $\ell_{jk}, \rho_j$ | $\ell_{jk} = \mathbf{I}_{jk}\mathbf{I}_{jk}^\mathsf{H}$, $\rho_j = \mathbf{I}_{\Delta,j}\mathbf{I}_{\Delta,j}^\mathsf{H}$. |
| $\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}$ | Block matrix defined in (4). |
| $\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}$ | Block matrix defined in (8a). |
| $\mathbf{M}^{\mathbf{v},\mathbf{X},\rho}$ | Block matrix defined in (8b). |
| $p_{\mathrm{loss}}$ | Total active power loss in the network. |
| $d_p, d_q$ | Penalty functions on active and reactive power deviation. |
| $\Phi_Y^j, \Phi_\Delta^j$ | Available wye and delta connected phases at bus $j$. |

Fengyu Zhou and Steven H. Low are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: {f.zhou, slow}@caltech.edu). Ahmed S. Zamzam is with the National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: ahmed.zamzam@nrel.gov). Nicholas D. Sidiropoulos is with the Electrical and Computer Engineering Department, University of Virginia, Charlottesville, VA 22901 USA (email: nikos@virginia.edu)

## I. Introduction

**O**PTIMAL power flow (OPF) is a mathematical program that finds an optimal operating point for a power grid subject to laws of physics and operational constraints [1]. OPF formulated under the AC model is known to be both nonconvex and NP-hard to solve [2], [3]. Commonly used methods to approximately solve this non-convex problem include, e.g., the Newton-Raphson method [4] and various linearizations [5], [6], [7]. Another approach that has emerged over the last decade or so is to relax the OPF problem to a convex program, such as semi-definite program (SDP) relaxations and second-order cone program (SOCP) relaxations [8], [9], [10]. For single-phase radial networks (i.e., networks with a tree topology) as well as the single-phase equivalent of a balanced three-phase radial network, simulations have shown that these relaxations often yield solutions that are also global optima of the original nonconvex problems [11], [12]. Sufficient conditions that guarantee exact relaxations for single-phase radial networks have subsequently been proved. There is now sizeable literature on OPF relaxations; see, e.g., comprehensive surveys in [13], [14] for pointers to various contributions including many earlier surveys on OPF. Many recent works further refine and improve the convexification of OPF problems and exhibit promising performance [15], [16].

Most of this literature focuses on single-phase models, but distribution systems have multiple phases that are increasingly unbalanced as distributed energy resources continue to grow. Mathematically, we can identify each bus-phase pair of a three-phase radial network with an equivalent single-phase bus and transform it into a single-phase equivalent circuit with a meshed topology [17], [18]. Most analytical results on exact relaxation for meshed networks are restricted to weakly-cyclic networks [19], but the single-phase equivalent of a three-phase radial network is beyond this class even for a 2-bus

toy example. SDP relaxation has been generalized in [20], [21] to three-phase radial networks with only wye connections and shown to be exact in the simulation of several test cases. A recent work [22] proves a sufficient condition for exact relaxation in this case. Besides SDP relaxation, [23] also provides an iterative algorithm for three-phase networks, also without delta connections.

Semi-definite relaxation is recently extended in [24] to networks with delta connected devices by introducing a new positive semi-definite matrix that represents the outer product of voltages and phase-to-phase currents in the delta connections (matrix $\mathbf{M}^{\mathbf{V},\mathbf{X},\rho}(j)$ in (8b) below). Simulation results in [24] showed that, surprisingly, this matrix was never rank-1 at an optimal solution of the relaxation. This seems to suggest that the SDP relaxation was inexact in these simulations. In this paper, we show that even though the matrix $\mathbf{M}^{\mathbf{V},\mathbf{X},\rho}(j)$ fails to attain rank 1, an exact solution can still be recovered under certain conditions; see Theorem 2 and Remark 1. The inexactness in previous works is due to two issues. First, optimal solutions to the SDP relaxation in these simulations are generally not unique, and the exact solution is only one of them which is not returned by the solver. Second, such non-uniqueness could significantly amplify the numerical error and make it computationally challenging to recover the exact solution. We propose two variants of the standard SDP relaxation that address both issues. The first algorithm applies post-processing to the relaxation solution and tends to provide lower cost but larger constraint violation, while the second algorithm adds a penalty term to the cost and tends to provide higher cost but smaller constraint violation. Simulations of both algorithms corroborate the theoretical results and show that they can recover exact solutions for three IEEE distribution feeders.

The remainder of the paper is organized as follows. In Section II, we define the network structure and formulate the three-phase OPF problem in both the bus injection model (BIM) and the branch flow model (BFM). We subsequently show that the two models are essentially equivalent to each other. In Section III, we prove that the global optimal solution to the nonconvex OPF problem can be recovered from its relaxation under certain conditions, and two algorithms are presented. Section IV shows the equivalence between BIM and BFM. Finally, in Section V, we apply our algorithms to IEEE 13-, 37- and 123-bus systems.

## II. SYSTEM MODEL

### A. Network Structure

We study the model proposed in [21], [24]. Let the directed graph representing the electrical network be $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{0, 1, \ldots, n\}$ denotes the set of buses, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges; and let $N := |\mathcal{V}| = n + 1$. In this paper, we focus on the case where $\mathcal{G}$ represents a radial network (i.e., a tree) because most distribution networks have a tree topology. 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 substation bus where the distribution feeder is connected to a transmission
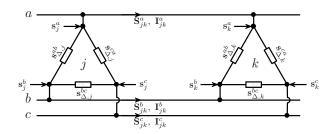


Fig. 1. Illustration of line $(j, k)$ with both wye and delta connections.

network. Suppose the substation also serves as the slack bus, so the voltages at the substation bus are fixed and specified. We use $j \to k$ to denote a directed edge from bus $j$ to $k$. In many situations, when we do not care about the direction of the edge, we simply use $(j, k)$ and $j \sim k$ interchangeably to denote an edge connecting bus $j$ and $k$. That means either $j \to k$ or $k \to j$ is in $\mathcal{E}$. Consider a three-phase line $(j, k)$ characterized by the series impedance matrix $z_{jk} \in \mathbb{C}^{3 \times 3}$. When line $(j, k)$ has three phases, the inverse of $z_{jk}$, denoted as $y_{jk}$, is the admittance of line $(j, k)$. If branch $(j, k)$ has less than three phases, then we fill the rows and columns of $z_{jk}$ corresponding to the missing phases with zeros, and we let the admittance matrix $y_{jk}$ be the pseudo-inverse of $z_{jk}$. Last, let $y_j \in \mathbb{C}^{3 \times 3}$ denote the admittance of a shunt device connected to bus $j$.[1]

For each bus $j$, let the voltages of all three phases at bus $j$ be collected in the vector $\mathbf{V}_j \in \mathbb{C}^3$. We use $\mathbf{V}_j^\phi$ for $\phi \in \{a, b, c\}$ to indicate the voltage of phase $\phi$. The voltage $\mathbf{V}_0$ at slack bus 0 is known and denoted by $\mathbf{V}_{\text{ref}}$. Let $\mathbf{V} = [\mathbf{V}_0^\mathsf{T}, \mathbf{V}_1^\mathsf{T}, \ldots, \mathbf{V}_n^\mathsf{T}]^\mathsf{T}$ collect the voltages for the entire network. Similarly, we use $\mathbf{s}_j^\phi$ to denote the bus injection for phase $\phi$ at bus $j$; and we denote $\mathbf{s}_j$ and $\mathbf{s}$ as the injections at bus $j$ and in the entire network, respectively.

For delta connected components, we use $\mathbf{I}_{\Delta,j} \in \mathbb{C}^3$ to collect the delta line currents for phases in $\{ab, bc, ca\}$. Define the matrix

$$\Gamma := \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Therefore, the complex power injections of the delta connected components at bus $j$ can be expressed as $\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{V}_j \mathbf{I}_{\Delta,j}^\mathsf{H})$. The net nodal injections contributed by delta connections at bus $j$ are given by $-\text{diag}(\mathbf{V}_j \mathbf{I}_{\Delta,j}^\mathsf{H} \Gamma)$ (see the illustration in Fig. 1). Assume that the operation regions for $\mathbf{s}_j$ and $\mathbf{s}_{\Delta,j}$ at bus $j$ are convex compact sets $\mathcal{S}_j$ and $\mathcal{S}_{\Delta,j}$, respectively.

The AC power flow equations are:

$$\mathbf{s}_j - \text{diag}(\mathbf{V}_j \mathbf{I}_{\Delta,j}^\mathsf{H} \Gamma) - \text{diag}(\mathbf{V}_j \mathbf{V}_j^\mathsf{H} y_j^\mathsf{H})$$
$$= \sum_{k:j \sim k} \text{diag}((\mathbf{V}_j \mathbf{V}_j^\mathsf{H} - \mathbf{V}_j \mathbf{V}_k^\mathsf{H}) y_{jk}^\mathsf{H}) \quad \text{(1a)}$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{V}_j \mathbf{I}_{\Delta,j}^\mathsf{H}). \quad \text{(1b)}$$

[1]The shunt here refers to a capacitive device at bus $j$ and not the line charging in the Π circuit model.

where (1a) is the power balance equation at bus $j$ and (1b) defines the power through delta connected components.

Similar to [24], we will adopt $\mathbf{X}_j, \rho_j \in \mathbb{C}^{3\times3}$ as auxiliary matrices in the following models as well as their corresponding relaxations. The outer products of voltages and currents are defined as follows:

$$\mathbf{X}_j = \mathbf{V}_j \mathbf{I}_{\Delta,j}^{\mathsf{H}} \tag{2a}$$

$$\rho_j = \mathbf{I}_{\Delta,j} \mathbf{I}_{\Delta,j}^{\mathsf{H}}. \tag{2b}$$

We consider an OPF problem that minimizes a cost function $f(\mathbf{s}, \mathbf{s}_\Delta)$ over variables $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{V}, \mathbf{I}_\Delta)$ subject to power flow equations (1) as well as voltage and injection limits:

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{V}, \mathbf{I}_\Delta}{\text{minimize}} \quad f(\mathbf{s}, \mathbf{s}_\Delta) \tag{3a}$$

$$\text{subject to} \quad (1) \tag{3b}$$

$$\mathbf{V}_0 = \mathbf{V}_{\text{ref}} \tag{3c}$$

$$\mathbf{s}_j \in \mathcal{S}_j, \ \mathbf{s}_{\Delta,j} \in \mathcal{S}_{\Delta,j}, \ \text{for } j \in \mathcal{V} \tag{3d}$$

$$\underline{\mathbf{V}} \le |\mathbf{V}| \le \overline{\mathbf{V}}. \tag{3e}$$

In (3e), $|\mathbf{V}|$ stands for the modulus of $\mathbf{V}$ elementwise, and $\underline{\mathbf{V}}, \overline{\mathbf{V}} \in \mathbb{R}^{3N}$ are the lower and upper limits for the voltages. If the limits are homogeneous across all buses and phases, we can denote them as $\underline{V}\mathbf{1}, \overline{V}\mathbf{1}$, where $\underline{V}, \overline{V}$ are scalars and $\mathbf{1}$ is the all-one vector.

## B. Bus Injection Model

The *bus injection model* (BIM) is defined in terms of $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho)$, where we use $\mathbf{W} \in \mathbb{C}^{3N\times3N}$ to replace $\mathbf{V}\mathbf{V}^{\mathsf{H}}$ in (1). The matrix $\mathbf{W}_{jk} \in \mathbb{C}^{3\times3}$ is the $(j,k)$ submatrix of $\mathbf{W}$. For notational simplicity, we let

$$\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}(j) := \begin{bmatrix} \mathbf{W}_{jj} & \mathbf{X}_j \\ \mathbf{X}_j^{\mathsf{H}} & \rho_j \end{bmatrix} \qquad \text{for } j \in \mathcal{V}. \tag{4}$$

The power flow model is represented as

$$\mathbf{s}_j - \text{diag}(\mathbf{X}_j \Gamma) - \text{diag}(\mathbf{W}_{jj} y_j^{\mathsf{H}})$$
$$= \sum_{k:j\sim k} \text{diag}((\mathbf{W}_{jj} - \mathbf{W}_{jk}) y_{jk}^{\mathsf{H}}) \tag{5a}$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{X}_j) \tag{5b}$$

$$\mathbf{W}_{00} = \mathbf{V}_{\text{ref}} \mathbf{V}_{\text{ref}}^{\mathsf{H}} \tag{5c}$$

$$\mathbf{W} \succeq 0 \tag{5d}$$

$$\text{rank}(\mathbf{W}) = 1 \tag{5e}$$

$$\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}(j) \succeq 0 \tag{5f}$$

$$\text{rank}\left(\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}(j)\right) = 1. \tag{5g}$$

Hence, the AC-OPF problem in BIM formulation is:

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho}{\text{minimize}} \quad f(\mathbf{s}, \mathbf{s}_\Delta) \tag{6a}$$

$$\text{subject to} \quad (5), (3d). \tag{6b}$$

$$\text{diag}(\underline{\mathbf{V}}\underline{\mathbf{V}}^{\mathsf{H}}) \le \text{diag}(\mathbf{W}) \le \text{diag}(\overline{\mathbf{V}}\overline{\mathbf{V}}^{\mathsf{H}}) \tag{6c}$$

## C. Branch Flow Model

In a *branch flow model* (BFM), we introduce $\mathbf{S}$, $\mathbf{v}$, and $\ell$ to model the branch power flow, squared voltages, and squared currents, respectively. The matrices $\mathbf{S}$, $\mathbf{v}$, and $\ell$ can be written as

$$\mathbf{S} = (\mathbf{S}_{jk} \in \mathbb{C}^{3\times3})_{(j\to k)\in\mathcal{E}}, \qquad \mathbf{S}_{jk} = \mathbf{V}_j \mathbf{I}_{jk}^{\mathsf{H}} \tag{7a}$$

$$\mathbf{v} = (\mathbf{v}_j \in \mathbb{C}^{3\times3})_{j\in\mathcal{V}}, \qquad \mathbf{v}_j = \mathbf{V}_j \mathbf{V}_j^{\mathsf{H}} \tag{7b}$$

$$\ell = (\ell_{jk} \in \mathbb{C}^{3\times3})_{(j\to k)\in\mathcal{E}}, \qquad \ell_{jk} = \mathbf{I}_{jk} \mathbf{I}_{jk}^{\mathsf{H}}. \tag{7c}$$

Here, $\mathbf{I}_{jk} := y_{jk}(\mathbf{V}_j - \mathbf{V}_k)$ is the sending-end current from bus $j$ to bus $k$, and $\mathbf{S}_{jk}$ is the sending-end branch power from $j$ to $k$. Let

$$\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}(j,k) := \begin{bmatrix} \mathbf{v}_j & \mathbf{S}_{jk} \\ \mathbf{S}_{jk}^{\mathsf{H}} & \ell_{jk} \end{bmatrix} \qquad \text{for } j \to k \tag{8a}$$

$$\mathbf{M}^{\mathbf{v},\mathbf{X},\rho}(j) := \begin{bmatrix} \mathbf{v}_j & \mathbf{X}_j \\ \mathbf{X}_j^{\mathsf{H}} & \rho_j \end{bmatrix} \qquad \text{for } j \in \mathcal{V}. \tag{8b}$$

The branch flow model is defined in terms of $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho)$, and it is expressed as

$$\mathbf{v}_k = \mathbf{v}_j - (\mathbf{S}_{jk} z_{jk}^{\mathsf{H}} + z_{jk} \mathbf{S}_{jk}^{\mathsf{H}}) + z_{jk} \ell_{jk} z_{jk}^{\mathsf{H}} \tag{9a}$$

$$\sum_{k:j\to k} \text{diag}(\mathbf{S}_{jk}) - \sum_{l:l\to j} \text{diag}(\mathbf{S}_{lj} - z_{lj}\ell_{lj}) \tag{9b}$$
$$= - \text{diag}(\mathbf{v}_j y_j^{\mathsf{H}} + \mathbf{X}_j \Gamma) + \mathbf{s}_j$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{X}_j) \tag{9c}$$

$$\mathbf{v}_0 = \mathbf{V}_{\text{ref}} \mathbf{V}_{\text{ref}}^{\mathsf{H}} \tag{9d}$$

$$\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}(j,k) \succeq 0 \tag{9e}$$

$$\text{rank}\left(\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}(j,k)\right) = 1 \tag{9f}$$

$$\mathbf{M}^{\mathbf{v},\mathbf{X},\rho}(j) \succeq 0 \tag{9g}$$

$$\text{rank}\left(\mathbf{M}^{\mathbf{v},\mathbf{X},\rho}(j)\right) = 1. \tag{9h}$$

The AC-OPF problem in the BFM form can be formulated as:

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho}{\text{minimize}} \quad f(\mathbf{s}, \mathbf{s}_\Delta) \tag{10a}$$

$$\text{subject to} \quad (9), (3d) \tag{10b}$$

$$\text{diag}(\underline{\mathbf{V}}_j \underline{\mathbf{V}}_j^{\mathsf{H}}) \le \text{diag}(\mathbf{v}_j) \le \text{diag}(\overline{\mathbf{V}}_j \overline{\mathbf{V}}_j^{\mathsf{H}}) \tag{10c}$$

## III. MAIN RESULTS

The main challenge to solving OPF problems (6) and (10) is the nonconvex rank constraints in (5e), (5g), (9f) and (9h). If we drop all the rank-1 constraints, then we obtain

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho}{\text{minimize}} \quad f(\mathbf{s}, \mathbf{s}_\Delta) \tag{11a}$$

$$\text{subject to} \quad (5a) - (5d), (5f), (3d), (6c). \tag{11b}$$

as the relaxation for the BIM and

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho}{\text{minimize}} \quad f(\mathbf{s}, \mathbf{s}_\Delta) \tag{12a}$$

$$\text{subject to} \quad (3d), (10c), (9a) - (9e), (9g) \tag{12b}$$

as the relaxation for the BFM. Solving the relaxed problems (11) and (12) could lead to solutions that are infeasible for the original nonconvex problems (6) and (10) respectively when the solutions do not satisfy the rank-1 constraints. In what follows, we will explore conditions under which optimal solutions of (6) and (10) can be recovered from their respective relaxations. First, the following lemma is presented, which is the main ingredient in the proofs of subsequent results.

*Lemma 1:* Consider a block Hermitian matrix

$$\mathbf{M} := \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathsf{H}} & \mathbf{C} \end{bmatrix} \tag{13}$$

where $\mathbf{A}$ and $\mathbf{C}$ are both square matrices. If $\mathbf{M} \succeq 0$ and $\mathbf{A} = \mathbf{x}\mathbf{x}^{\mathsf{H}}$ for some vector $\mathbf{x}$, then there must exist some vector $\mathbf{y}$ such that $\mathbf{B} = \mathbf{x}\mathbf{y}^{\mathsf{H}}$.

*Proof:* As $\mathbf{M} \succeq 0$, it can be decomposed as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{bmatrix} \begin{bmatrix} \mathbf{M}_1^{\mathsf{H}} & \mathbf{M}_2^{\mathsf{H}} \end{bmatrix} \tag{14}$$

and $\mathbf{A} = \mathbf{M}_1\mathbf{M}_1^{\mathsf{H}}$, $\mathbf{B} = \mathbf{M}_1\mathbf{M}_2^{\mathsf{H}}$, $\mathbf{C} = \mathbf{M}_2\mathbf{M}_2^{\mathsf{H}}$. Because $\mathbf{A} = \mathbf{x}\mathbf{x}^{\mathsf{H}}$ has rank-1, matrix $\mathbf{M}_1$ is in the column space of $\mathbf{x}$ and has rank-1 as well. There must exist vector $\mathbf{z}$ such that $\mathbf{M}_1 = \mathbf{x}\mathbf{z}^{\mathsf{H}}$. As a result, $\mathbf{B} = \mathbf{M}_1\mathbf{M}_2^{\mathsf{H}} = \mathbf{x}\mathbf{z}^{\mathsf{H}}\mathbf{M}_2^{\mathsf{H}} = \mathbf{x}(\mathbf{M}_2\mathbf{z})^{\mathsf{H}}$. ∎

One observation in Lemma 1 is when submatrices $\mathbf{A}$ and $\mathbf{B}$ are fixed and specified as $\mathbf{x}\mathbf{x}^{\mathsf{H}}$ and $\mathbf{x}\mathbf{y}^{\mathsf{H}}$, there are non-unique $\mathbf{C}$ to make $\mathbf{M}$ positive semi-definite. Similarly, in the relaxations (11) and (12), the optimal solutions are always non-unique. Taking (11) as an example, for any optimal solution $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$, one could add to $\rho^*$ an arbitrary PSD matrix to obtain a different optimal solution $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^* + \mathbf{K}\mathbf{K}^{\mathsf{H}})$. This non-uniqueness in the optimal $\rho$ explains why in existing literature such as [24], the relaxation (11) could compute rank-1 $\mathbf{W}$ (within numerical tolerance) but the resulting $\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}$ is always not rank-1. In fact, the next result shows in theory, if the optimal $\mathbf{W}$ is perfectly of rank 1 without any numerical error, then an feasible and optimal solution of (6) is recoverable.

*Theorem 1:* If $\mathbf{u}^* = (\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$ is an optimal solution to (11) that satisfies $\mathrm{rank}(\mathbf{W}^*) = 1$, then an optimal solution of (6) can be recovered from $\mathbf{u}^*$.

*Proof:* We decompose $\mathbf{W}_{jj}^*$ as $\mathbf{V}_j\mathbf{V}_j^{\mathsf{H}}$ for each $j$, where $\mathbf{V}_j$ is a vector. By Lemma 1, there exists vector $\mathbf{I}_{\Delta,j}$ such that $\mathbf{X}_j^* = \mathbf{V}_j\mathbf{I}_{\Delta,j}^{\mathsf{H}}$. One could construct $\tilde{\rho}$ such that $\tilde{\rho}_j = \mathbf{I}_{\Delta,j}\mathbf{I}_{\Delta,j}^{\mathsf{H}}$.

Since (11) is a relaxation of (6), for $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \tilde{\rho})$ to be optimal for (6), it is sufficient that it is feasible for (6). Clearly, constraints (3d), (6c), (5a)–(5d) are satisfied because they are also the constraints in (11) and they do not involve the decision variable $\rho$. Constraint (5e) also holds as $\mathrm{rank}(\mathbf{W}^*) = 1$. Further, by Lemma 1, we have

$$\begin{bmatrix} \mathbf{W}_{jj}^* & \mathbf{X}_j^* \\ (\mathbf{X}_j^*)^{\mathsf{H}} & \tilde{\rho}_j \end{bmatrix} = \begin{bmatrix} \mathbf{V}_j \\ \mathbf{I}_{\Delta,j} \end{bmatrix} \begin{bmatrix} \mathbf{V}_j \\ \mathbf{I}_{\Delta,j} \end{bmatrix}^{\mathsf{H}}$$

is both positive semi-definite and of rank-1. Hence, (5f) and (5g) are also satisfied. Hence, $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \tilde{\rho})$ is feasible for (6), and this completes the proof. ∎

*Theorem 2:* If $\mathbf{u}^* = (\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{S}^*, \mathbf{v}^*, \ell^*, \mathbf{X}^*, \rho^*)$ is an optimal solution to (12) and satisfies $\mathrm{rank}(\mathbf{M}^{\mathbf{v}^*, \mathbf{S}, *\ell^*}(j,k)) = 1$ for $j \sim k$ and $\mathrm{rank}(\mathbf{v}_j^*) = 1$ for $j \in \mathcal{V}$, then an optimal solution of (10) can be recovered from $\mathbf{u}^*$.

The proof of Theorem 2 is omitted because it is similar to the proof of Theorem 1.

Theorem 1 asserts that in theory, the only critical nonconvex constraint of (6) is (5e), in the sense that a solution satisfying (5g) could always be recovered whenever (5e) holds. However in practice, $\mathbf{W}^*$ is typically not exactly rank-1 due to numerical precision and therefore Theorem 2 could not be directly applied to recover the optimal solution as long as numerical error exists. This is because the recovery method in Theorem 1 relies on the rank-1 decomposition of $\mathbf{X}^*$. In practice even if $\mathbf{W}^*$ is close to being rank-1, the optimal $\mathbf{X}^*$ could still be very different from being rank-1, as we will explain below in Remark 1.

*Remark 1 (Spectrum Error):* More specifically, the matrix $\mathbf{A}$ in (14) being *approximately* rank-1 does not necessarily mean that $\mathbf{B}$ is also approximately rank-1.[2] For example, consider the case where $\mathbf{x}$, $\mathbf{e}_1$, and $\mathbf{e}_2$ are orthogonal vectors with norms 1, $10^{-4}$, and $10^{-5}$, respectively. Similarly, let $\mathbf{y}$, $\mathbf{z}_1$, and $\mathbf{z}_2$ be orthogonal vectors with norms 1, $10^4$, and $10^5$, respectively. Then, construct the matrix $\mathbf{M}$ as in (14) with $\mathbf{M}_1 = [\mathbf{x} \quad \mathbf{e}_1 \quad \mathbf{e}_2]$ and $\mathbf{M}_2 = [\mathbf{y} \quad \mathbf{z}_1 \quad \mathbf{z}_2]$. Clearly, the resulting matrix $\mathbf{M}$ has the upper left diagonal block that is approximately rank-1. In fact, the ratio between the two leading eigenvalues of the upper diagonal block is $10^{-8}$. On the other hand, the upper right block is of rank 3 with three singular values of 1. Consequently, even when $\mathbf{W}^*$ is close to rank-1 within a certain numerical tolerance, $\mathbf{X}^*$ could be far from being a rank-1 matrix, especially if $\rho^*$ already contains a large redundant PSD matrix $\mathbf{K}\mathbf{K}^{\mathsf{H}}$. Decomposing $\mathbf{X}^*$ as the product of two vectors, as in the proof of Theorem 1, could result in a large numerical error. In other words, the non-uniqueness of $\rho$ could significantly amplify the numerical error. In fact, as long as the second largest eigenvalue of $\mathbf{W}^*$ is not exactly 0, then no matter how small it is, such spectrum error could potentially be significant, especially when the trace of $\mathbf{K}\mathbf{K}^{\mathsf{H}}$ is large.

To summarize, there are two factors that prevent the relaxation output from being exact. The first is the non-uniqueness in the relaxation solution, and the second is that such non-uniqueness further greatly amplify the numerical error in computation. This finding motivates two algorithms for practical implementation.

## A. Relaxation with Post-Processing

Remark 1 shows even if $\mathbf{W}^*$ is approximately rank-1, $\mathbf{X}^*$ could be far away from being rank-1, and thus recovering the vector $\mathbf{I}_{\Delta,j}$ from $\mathbf{X}_j^*$ can lead to poor numerical performance. In the first algorithm, we instead recover $\mathbf{I}_{\Delta,j}^{\mathsf{H}}$ as $\mathrm{diag}^{-1}(\Gamma\mathbf{V}_j)\mathbf{s}_{\Delta,j}^*$ from (1b), and then we reconstruct $\tilde{\mathbf{X}}_j$ as $\mathbf{V}_j\mathbf{I}_{\Delta,j}^{\mathsf{H}}$. If there is no numerical error, $\mathbf{X}^*$ and $\tilde{\mathbf{X}}$ should be equal; however, in the presence of spectrum error, they could

---

[2]Here, being approximately rank-1 means that the second largest eigenvalue of the matrix is nonzero but smaller than the largest eigenvalue by several orders of magnitude.

be different, as discussed in Remark 1. The pseudo code is provided in Algorithm 1.

---

**Algorithm 1** Relaxation Algorithm with Post-Processing.

---

**Input:** $y$, $\mathcal{S}$, $\mathcal{S}_\Delta$
**Output:** Optimal solution $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho)$ to (6).
1: Solve (11) to obtain $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$.
2: **if** $(\mathrm{rank}(\mathbf{W}^*) > 1)$ **then**
3:     Output 'Failed!'
4:     Exit
5: **else**
6:     Decompose $\mathbf{W}_{jj}^* = \mathbf{V}_j \mathbf{V}_j^{\mathsf{H}}$
7:     $\mathbf{I}_{\Delta,j}^{\mathsf{H}} \leftarrow \mathrm{diag}^{-1}(\Gamma \mathbf{V}_j) \mathbf{s}_{\Delta,j}^*$
8:     $\tilde{\mathbf{X}} \leftarrow \mathbf{V}_j \mathbf{I}_{\Delta,j}^{\mathsf{H}}$, $\tilde{\rho}_j \leftarrow \mathbf{I}_{\Delta,j} \mathbf{I}_{\Delta,j}^{\mathsf{H}}$
9:     **return** $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \tilde{\mathbf{X}}, \tilde{\rho})$
10: **end if**

---

*Theorem 3:* If Algorithm 1 does not fail, then its output is an optimal solution of (6).

Theorem 3 is the direct consequence of Theorem 1. Algorithm 1 did not fail in any of our simulations in Section V of IEEE test cases.

Similarly for BFM, one could also apply post-processing to recover the solution of (10) from an optimal solution of (12). In the BFM, instead of checking the rank of $\mathbf{W}^*$, we check the rank of $\mathbf{M}^{\mathbf{v}^*, \mathbf{S}, {}^*\ell^*}(j, k)$ for each $j \sim k$ and $\mathbf{v}_j^*$ for $j \in \mathcal{V}$.

### B. Relaxation with Penalized Cost Function

Since the inexactness of relaxations (11) and (12) originates from two issues: the non-uniqueness in $\rho^*$ and the spectrum error, where the latter is essentially amplified by the former. The second algorithm we propose is to penalize and suppress the trace of $\rho_j$ in the cost function. With such penalty term, the value of $\rho^*$ will be unique for fixed $\mathbf{W}^*$ and $\mathbf{X}^*$ in the solution of (11) and the spectrum error can also be restricted. Similar penalization approaches were previously proposed in [25], [26] to promote low-rank solutions of relaxed problems.

The penalized relaxed problem formulation under the BIM becomes

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho}{\text{minimize}} \qquad f(\mathbf{s}, \mathbf{s}_\Delta) + \lambda \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j) \qquad (15a)$$

$$\text{subject to} \qquad (5a) - (5d), (5f), (3d), (6c) \qquad (15b)$$

Similarly, the penalized relaxed program under the BFM becomes

$$\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho}{\text{minimize}} \qquad f(\mathbf{s}, \mathbf{s}_\Delta) + \lambda \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j) \qquad (16a)$$

$$\text{subject to} \qquad (3d), (10c), (9a) - (9e), (9g). \qquad (16b)$$

Because $\mathrm{tr}(\rho_j)$ is linear and all constraints in (15b) and (16b) are convex, both (15) and (16) are convex optimization problems and can be efficiently solved in polynomial time.[3] Here, $\lambda > 0$ controls the weight of $\sum \mathrm{tr}(\rho_j)$ in the cost function. The pseudo code (based on BIM) is summarized in Algorithm 2. The algorithm for BFM is similar.

---

**Algorithm 2** Relaxation Algorithm with Penalized Cost Function.

---

**Input:** $y$, $\mathcal{S}$, $\mathcal{S}_\Delta$
**Output:** Optimal solution $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho)$ to (6).
1: Pick a sufficiently small $\lambda > 0$
2: Solve (15) and obtain $\mathbf{u}^* := (\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$
3: **if** $(\mathrm{rank}(\mathbf{W}^*) > 1)$ **then**
4:     Output 'Failed!'
5:     Exit
6: **else**
7:     **return** $\mathbf{u}^*$
8: **end if**

---

Because the cost function in the penalized program has been changed, the output of Algorithm 2 might not be the global optimal solution of (6). We next show that the output of Algorithm 2 serves as an approximation of the true optimal solution. We make the following assumption.

*Assumption 1:* The problem (11) has at least one finite optimal solution.

Now consider a sequence of positive and decreasing $\lambda_i$ for $i = 1, 2, \cdots$. Taking BIM as an example, let the optimal solution of (15) with respect to $\lambda_i$ be $\mathbf{u}^{(i)}$.[4] Then the following lemma implies the sequence $\mathbf{u}^{(i)}$ has a limit point.

*Lemma 2:* The sequence $(\mathbf{u}^{(i)})_{i=1}^{\infty}$ resides in a compact set, and hence has a limit point.

*Proof:* Because all the constraints in (15) are closed, we only need to prove boundedness. By assumption, $\mathbf{s}_j^{(i)}$ and $\mathbf{s}_{\Delta,j}^{(i)}$ at bus $j$ are in compact sets $\mathcal{S}_j$ and $\mathcal{S}_{\Delta,j}$ respectively. The positive semi-definite matrix $\mathbf{W}^{(i)}$ has upper bounds on its diagonal elements and is therefore bounded. We only need to show that $\sum_j \mathrm{tr}(\rho_j^{(i)})$ is also bounded because the boundedness of $\mathbf{X}^{(i)}$ is implied by the constraint (5f) as long as $\sum_j \mathrm{tr}(\rho_j^{(i)})$ is bounded.

To show $\sum_j \mathrm{tr}(\rho_j^{(i)})$ is bounded, let $\hat{\mathbf{u}} = (\hat{\mathbf{s}}, \hat{\mathbf{s}}_\Delta, \hat{\mathbf{W}}, \hat{\mathbf{X}}, \hat{\rho})$ be an optimal solution of (11). Then $\hat{\mathbf{u}}$ is feasible for (15) regardless of the value of $\lambda$. For any $i$, we must have $\sum_j \mathrm{tr}(\rho_j^{(i)}) \leq \sum_j \mathrm{tr}(\hat{\rho}_j)$; otherwise, $\hat{\mathbf{u}}$ will always give a strictly smaller cost value in (15) for $\lambda = \lambda_i$ and it would contradict the optimality of $\mathbf{u}^{(i)}$. ∎

Suppose $\tilde{\mathbf{u}} := (\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta, \tilde{\mathbf{W}}, \tilde{\mathbf{X}}, \tilde{\rho})$ is an arbitrary limit point of the sequence $\mathbf{u}^{(i)}$. We present sufficient conditions for $\tilde{\mathbf{u}}$ to be an optimal solution of (6). First, we introduce the following lemma, which will be used to prove the optimality conditions.

*Lemma 3:* Consider the positive semi-definite matrix $\mathbf{M}$ as in (13) where $\mathbf{A} = \mathbf{x}\mathbf{x}^{\mathsf{H}}$ for some vector $\mathbf{x}$ such that $\mathbf{x} \neq \mathbf{0}$, and $\mathbf{B} = \mathbf{x}\mathbf{y}^{\mathsf{H}}$. Then,

$$\mathrm{tr}(\mathbf{M}) \geq \mathbf{x}^{\mathsf{H}}\mathbf{x} + \mathbf{y}^{\mathsf{H}}\mathbf{y}$$

and equality holds if and only if $\mathbf{C} = \mathbf{y}\mathbf{y}^{\mathsf{H}}$.

*Proof:* It is sufficient to prove that $\mathbf{C} - \mathbf{y}\mathbf{y}^{\mathsf{H}} \succeq 0$. If not, then suppose there exists $\mathbf{z}$ such that $\mathbf{z}^{\mathsf{H}}(\mathbf{C} - \mathbf{y}\mathbf{y}^{\mathsf{H}})\mathbf{z} < 0$. Because $\mathbf{x} \neq \mathbf{0}$, we can always find $\mathbf{w}$ such that $\mathbf{w}^{\mathsf{H}}\mathbf{x} =$

---

[3] If the constraint (3d) has a linear or quadratic formulation, then both problems are SOCPs.

[4] If the program has multiple solutions, then pick any one of them.

$-\mathbf{z}^{\mathsf{H}}\mathbf{y}$. Consider

$$\begin{bmatrix} \mathbf{w} \\ \mathbf{z} \end{bmatrix}^{\mathsf{H}} \mathbf{M} \begin{bmatrix} \mathbf{w} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{w} \\ \mathbf{z} \end{bmatrix}^{\mathsf{H}} \begin{bmatrix} \mathbf{x}\mathbf{x}^{\mathsf{H}} & \mathbf{x}\mathbf{y}^{\mathsf{H}} \\ \mathbf{y}\mathbf{x}^{\mathsf{H}} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{z} \end{bmatrix}$$
$$= \mathbf{w}^{\mathsf{H}}\mathbf{x}\mathbf{x}^{\mathsf{H}}\mathbf{w} + \mathbf{z}^{\mathsf{H}}\mathbf{y}\mathbf{x}^{\mathsf{H}}\mathbf{w} + \mathbf{w}^{\mathsf{H}}\mathbf{x}\mathbf{y}^{\mathsf{H}}\mathbf{z} + \mathbf{z}^{\mathsf{H}}\mathbf{C}\mathbf{z}$$
$$= \mathbf{z}^{\mathsf{H}}\mathbf{y}\mathbf{y}^{\mathsf{H}}\mathbf{z} - \mathbf{z}^{\mathsf{H}}\mathbf{y}\mathbf{y}^{\mathsf{H}}\mathbf{z} - \mathbf{z}^{\mathsf{H}}\mathbf{y}\mathbf{y}^{\mathsf{H}}\mathbf{z} + \mathbf{z}^{\mathsf{H}}\mathbf{C}\mathbf{z}$$
$$= \mathbf{z}^{\mathsf{H}}(\mathbf{C} - \mathbf{y}\mathbf{y}^{\mathsf{H}})\mathbf{z} < 0.$$

This contradicts the positive semi-definiteness of $\mathbf{M}$. ∎

*Theorem 4:* If $\mathrm{rank}(\tilde{\mathbf{W}}) = 1$, then $\tilde{\mathbf{u}}$ is an optimal solution of (6).

*Proof:* We first show that $\tilde{\mathbf{u}}$ is the optimal solution of (11). Since (11) and (15) have the same feasible set, which is closed, $\tilde{\mathbf{u}}$ is also feasible for (11) and (15) for any $\lambda$. If $\tilde{\mathbf{u}}$ is not optimal for (11), then there must exist another point $\bar{\mathbf{u}}$ such that $f(\bar{\mathbf{s}}, \bar{\mathbf{s}}_\Delta) + \alpha = f(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta)$ and $\alpha > 0$. Then for some sufficiently large $i_0$, we have

$$\lambda_{i_0} \sum_{j \in \mathcal{V}} \mathrm{tr}(\bar{\rho}_j) < \frac{\alpha}{2}$$
$$|f(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta) - f(\mathbf{s}^{(i_0)}, \mathbf{s}_\Delta^{(i_0)})| < \frac{\alpha}{2}.$$

Therefore:

$$f(\bar{\mathbf{s}}, \bar{\mathbf{s}}_\Delta) + \lambda_{i_0} \sum_{j \in \mathcal{V}} \mathrm{tr}(\bar{\rho}_j) < f(\mathbf{s}^{(i_0)}, \mathbf{s}_\Delta^{(i_0)}) + \lambda_{i_0} \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j^{(i_0)})$$

which contradicts the optimality of $\mathbf{u}^{(i_0)}$.

Then for each $j$, we decompose $\tilde{\mathbf{W}}_{jj} = \tilde{\mathbf{x}}_j \tilde{\mathbf{x}}_j^{\mathsf{H}}$, $\tilde{\mathbf{X}} = \tilde{\mathbf{x}}_j \tilde{\mathbf{y}}_j^{\mathsf{H}}$, and construct $\rho_j^\dagger$ as $\tilde{\mathbf{y}}_j \tilde{\mathbf{y}}_j^{\mathsf{H}}$. Under the same argument as in the proof of Theorem 1, the solution $\mathbf{u}^\dagger := (\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta, \tilde{\mathbf{W}}, \tilde{\mathbf{X}}, \rho^\dagger)$ is an optimal solution for both (6) and (11). We want to prove $\tilde{\mathbf{u}} = \mathbf{u}^\dagger$ and conclude that $\tilde{\mathbf{u}}$ is also an optimal solution for (6). The proof is by contradiction.

Since $\mathbf{u}^\dagger$ is optimal for (6), $\mathbf{M}^{\tilde{\mathbf{W}}, \tilde{\mathbf{X}}, \rho^\dagger}(j)$ must be of rank-1 for all $j$. By Lemma 3, we have

$$\mathrm{tr}(\mathbf{M}^{\tilde{\mathbf{W}}, \tilde{\mathbf{X}}, \tilde{\rho}}(j)) \geq \mathrm{tr}(\mathbf{M}^{\tilde{\mathbf{W}}, \tilde{\mathbf{X}}, \rho^\dagger}(j))$$

and therefore $\mathrm{tr}(\tilde{\rho}_j) \geq \mathrm{tr}(\rho_j^\dagger)$ for all $j$. If $\tilde{\mathbf{u}} \neq \mathbf{u}^\dagger$, then some equalities cannot be achieved, and as a result, $\sum_j \mathrm{tr}(\tilde{\rho}_j) - \sum_j \mathrm{tr}(\rho_j^\dagger) = \beta$ for some $\beta > 0$.

As $\tilde{\mathbf{u}}$ is a limit point of $(\mathbf{u}^{(i)})_{i=1}^\infty$, there must be some sufficiently large $i_1$ such that

$$\left| \sum_{j \in \mathcal{V}} \mathrm{tr}(\tilde{\rho}_j) - \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j^{(i_1)}) \right| < \frac{\beta}{2},$$

Hence

$$\sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j^\dagger) < \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j^{(i_1)}).$$

On the other hand, (11) and (15) have the same feasible set, so the optimality of $\mathbf{u}^\dagger$ for (11) implies $f(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta) \leq f\left(\mathbf{s}^{(i_1)}, \mathbf{s}_\Delta^{(i_1)}\right)$. Therefore

$$f(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}_\Delta) + \lambda_{i_1} \sum_{j \in \mathcal{V}} \mathrm{tr}(\rho_j^\dagger) < f\left(\mathbf{s}^{(i_1)}, \mathbf{s}_\Delta^{(i_1)}\right) + \lambda_{i_1} \sum_{j \in \mathcal{V}} \mathrm{tr}\left(\rho_j^{(i_1)}\right)$$

which contradicts the fact that $\mathbf{u}^{(i_1)}$ is the optimal solution for

(15) with respect to $\lambda_{i_1}$. ∎

Theorem 4 shows that when we solve the penalized program with a sequence of decreasing $\lambda_i$, any limit point would be a global optimal for (6) as long as the $\mathbf{W}$ matrix associated with the limit point is of rank-1. In our simulations, we apply Algorithm 2 to solve (15) with a fixed but sufficiently small $\lambda$, which usually results in rank-1 solutions.

*Remark 2:* Further, if all optimal solutions of (11) have the same value for $\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}$, then Algorithm 1 succeeds if and only if $\mathrm{rank}(\tilde{\mathbf{W}}) = 1$ holds in Theorem 4. If Algorithm 1 succeeds, its output will also be the same as $\tilde{\mathbf{u}}$.

## IV. MODEL EQUIVALENCE

In the previous sections, our results for the BIM and BFM always come in pairs and are analogous. A natural question is whether there exist instances where one model produces an exact solution while the other does not. In single-phase networks and multi-phase systems with only wye connections, [27] and [21] have shown that the two models are equivalent in the sense that one will produce an exact solution if and only if the other will. We show in this subsection that a similar result holds in the presence of delta connections. We first define the equivalence between two optimization problems as follows.

*Definition 1:* Consider two optimization problems

$$\underset{x}{\mathrm{minimize}} \qquad f_A(x) \qquad (17a)$$
$$\mathrm{subject\ to} \qquad x \in \mathcal{X} \qquad (17b)$$

and

$$\underset{y}{\mathrm{minimize}} \qquad f_B(y) \qquad (18a)$$
$$\mathrm{subject\ to} \qquad y \in \mathcal{Y}. \qquad (18b)$$

We say (17) and (18) are *equivalent* if there exist mappings $g_1 : \mathcal{X} \to \mathcal{Y}$ and $g_2 : \mathcal{Y} \to \mathcal{X}$ such that

$$x \in \mathcal{X} \Rightarrow g_1(x) \in \mathcal{Y}, \ f_A(x) = f_B(g_1(x)),$$
$$y \in \mathcal{Y} \Rightarrow g_2(y) \in \mathcal{X}, \ f_B(y) = f_A(g_2(y)).$$

We do not require $g_1$ and $g_2$ to be bijections, but if one of the mappings is a bijection, then we can always set the other as its inverse. We denote the decision variables for the BIM as

$$\mathbf{u}^{\mathrm{BIM}} = (\mathbf{s}^{\mathrm{BIM}}, \mathbf{s}_\Delta^{\mathrm{BIM}}, \mathbf{W}^{\mathrm{BIM}}, \mathbf{X}^{\mathrm{BIM}}, \rho^{\mathrm{BIM}})$$

and the decision variables for the BFM as

$$\mathbf{u}^{\mathrm{BFM}} = (\mathbf{s}^{\mathrm{BFM}}, \mathbf{s}_\Delta^{\mathrm{BFM}}, \mathbf{S}^{\mathrm{BFM}}, \mathbf{v}^{\mathrm{BFM}}, \ell^{\mathrm{BFM}}, \mathbf{X}^{\mathrm{BFM}}, \rho^{\mathrm{BFM}}).$$

The superscripts here are to distinguish the same variable for different models.

*Proposition 1:* Problems (11) and (12) are equivalent. Moreover, for the pairs $g_1$ and $g_2$ in Definition 1, if $\mathbf{u}^{\mathrm{BIM}}$ satisfies (5e), then $g_1(\mathbf{u}^{\mathrm{BIM}})$ satisfies (9f). If $\mathbf{u}^{\mathrm{BFM}}$ satisfies (9f), then $g_2(\mathbf{u}^{\mathrm{BFM}})$ satisfies (5e).

We only sketch a proof here by providing the mappings $g_1$

and $g_2$, where $g_1$ can be written as

$$\mathbf{s}^{\mathrm{BFM}} = \mathbf{s}^{\mathrm{BIM}}, \ \mathbf{s}_\Delta^{\mathrm{BFM}} = \mathbf{s}_\Delta^{\mathrm{BIM}} \tag{19a}$$

$$\mathbf{S}_{jk}^{\mathrm{BFM}} = (\mathbf{W}_{jj}^{\mathrm{BIM}} - \mathbf{W}_{jk}^{\mathrm{BIM}})y_{jk}^{\mathsf{H}} \tag{19b}$$

$$\mathbf{v}_j^{\mathrm{BFM}} = \mathbf{W}_{jj}^{\mathrm{BIM}} \tag{19c}$$

$$\ell_{jk}^{\mathrm{BFM}} = y_{jk}(\mathbf{W}_{jj}^{\mathrm{BIM}} + \mathbf{W}_{kk}^{\mathrm{BIM}} - \mathbf{W}_{jk}^{\mathrm{BIM}} - \mathbf{W}_{kj}^{\mathrm{BIM}})y_{jk}^{\mathsf{H}} \tag{19d}$$

$$\mathbf{X}_j^{\mathrm{BFM}} = \mathbf{X}_j^{\mathrm{BIM}}, \ \rho_j^{\mathrm{BFM}} = \rho_j^{\mathrm{BIM}} \tag{19e}$$

and $g_2$ as

$$\mathbf{s}^{\mathrm{BIM}} = \mathbf{s}^{\mathrm{BFM}}, \ \mathbf{s}_\Delta^{\mathrm{BIM}} = \mathbf{s}_\Delta^{\mathrm{BFM}} \tag{20a}$$

$$\mathbf{W}_{jj}^{\mathrm{BIM}} = \mathbf{v}_j^{\mathrm{BFM}} \tag{20b}$$

$$\mathbf{W}_{jk}^{\mathrm{BIM}} = \begin{cases} \mathbf{v}_j^{\mathrm{BFM}} - \mathbf{S}_{jk}^{\mathrm{BFM}} z_{jk}^{\mathsf{H}}, & \text{if } j \to k \\ (\mathbf{W}_{kj}^{\mathrm{BIM}})^{\mathsf{H}}, & \text{if } k \to j \end{cases} \tag{20c}$$

$$\mathbf{X}_j^{\mathrm{BIM}} = \mathbf{X}_j^{\mathrm{BFM}}, \ \rho_j^{\mathrm{BIM}} = \rho_j^{\mathrm{BFM}}. \tag{20d}$$

For $g_2$, the value of $\mathbf{W}_{jk}^{\mathrm{BIM}}$ where $j \neq k$ and $(j,k) \notin \mathcal{E}$ can be determined arbitrarily as long as $\mathbf{W} \succeq 0$. As $\mathcal{G}$ is a tree, the values in (20) can guarantee that we can always complete the matrix $\mathbf{W}^{\mathrm{BIM}} \succeq 0$, but not necessarily in a unique way.

Proposition 1 shows that to apply Algorithm 1, if an optimal solution of (11) can produce an exact solution of (6), then there must also be an optimal solution of (12) that can produce an exact solution of (10), even though both (6) and (10) may have multiple solutions. The converse is also true. Informally, for Algorithm 1, both the BIM and BFM have the same capability of producing exact solutions.

The same holds for the penalized program.

*Proposition 2:* Problems (15) in the BIM and (16) in the BFM are equivalent when $\lambda$ takes the same value for both problems.

The proposition can be easily proved using the same mappings $g_1$ and $g_2$ in (19) and (20), respectively.

Beyond OPF problems, mappings $g_1$ and $g_2$ also provide the correspondence between feasible points under the two models. Thereby, a solution of power flow equations under one model can also be translated into a solution with the same physical meaning under the other model by applying $g_1$ or $g_2$. Note that power flow equations may have multiple solutions.

Though BIM and BFM are mathematically equivalent, the two models may behave differently in practice and shed lights on different properties. Some analysis may rely on the structure of one model but not the other, which is indeed the case for single-phase networks [28]. The equivalence implies that one could freely choose a model that is more convenient for a specific problem. For instance, a recent work [22] on exact relaxation for three-phase networks was derived based on BIM.

## V. NUMERICAL RESULTS

In this section, we show the ability of the proposed relaxation algorithms to recover the optimal solution to (6) and (10). We use the IEEE 13-, 37-, and 123-node distribution feeders [29] to assess the exactness of both algorithms for both the BIM and BFM models. Note that the IEEE 123-bus feeder does not include delta-connected components. Hence, we artificially added 4 delta-connected loads to the feeder

to assess the performance of the proposed approaches. In our experiments, we check how close the output matrices $\mathbf{W}, \mathbf{M}^{\mathbf{W},\mathbf{X},\rho}, \mathbf{M}^{\mathbf{v},\mathbf{S},\ell}, \mathbf{M}^{\mathbf{v},\mathbf{X},\rho}$ are to being rank-1, and we evaluate the maximum violation of the constraints when the decision variables are produced from the two proposed algorithms. For all the experiments in this section, we show that both algorithms succeed up to numerical precision, and each has its own advantages and disadvantages.

### A. Experimental Setup

The load transformer in the IEEE test feeders are modeled as lines with equivalent impedance, whereas the substation transformers and regulators are removed. The switches are assumed to be open or short according to their default status. The capacitor banks are modeled as controllable reactive power sources with continuous control space.

The voltage at the substation is assumed to be $\mathbf{V}_{\mathrm{ref}} = \overline{V}[1, e^{-\mathrm{i}2\pi/3}, e^{\mathrm{i}2\pi/3}]^{\mathsf{T}}$, where $\overline{V}$ is the maximum allowed voltage magnitude. The operational constraints for controllable loads are set as in [24]. The AC-OPF problem is solved with the cost function $f(\mathbf{s}, \mathbf{s}_\Delta)$ comprising three parts. The first part minimizes the total power losses in the network, and it can be written as

$$p_{\mathrm{loss}} = \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \mathbf{p}_j^\phi + \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_\Delta^j} \mathbf{p}_{\Delta,j}^\phi.$$

The second part penalizes deviations of the active and reactive injection profile from nominal profiles, and it is given by $d_p(\mathbf{p}, \mathbf{p}_\Delta)$ and $d_q(\mathbf{q}, \mathbf{q}_\Delta)$ as follows.

$$\begin{aligned} d_p(\mathbf{p}, \mathbf{p}_\Delta) = &\sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \frac{1}{2\overline{\mathbf{p}}_j^\phi}(\mathbf{p}_j^\phi - \overline{\mathbf{p}}_j^\phi)^2 \\ &+ \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_\Delta^j} \frac{1}{2\overline{\mathbf{P}}_{\Delta,j}^\phi}(\mathbf{p}_{\Delta,j}^\phi - \overline{\mathbf{p}}_{\Delta,j}^\phi)^2, \end{aligned}$$

$$\begin{aligned} d_q(\mathbf{q}, \mathbf{q}_\Delta) = &\sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \frac{1}{2\overline{\mathbf{q}}_j^\phi}(\mathbf{q}_j^\phi - \overline{\mathbf{q}}_j^\phi)^2 \\ &+ \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_\Delta^j} \frac{1}{2\overline{\mathbf{q}}_{\Delta,j}^\phi}(\mathbf{q}_{\Delta,j}^\phi - \overline{\mathbf{q}}_{\Delta,j}^\phi)^2. \end{aligned}$$

The values $\overline{\mathbf{p}}_j^\phi, \overline{\mathbf{q}}_j^\phi, \overline{\mathbf{P}}_{\Delta,j}^\phi, \overline{\mathbf{q}}_{\Delta,j}^\phi$ represent the nominal active and reactive injection values for phase $\phi$ at bus $j$. All the tracking errors are normalized by their nominal values to have the same order of magnitude for all quantities. In addition, $\Phi_Y^j \subseteq \{a,b,c\}$ and $\Phi_\Delta^j \subseteq \{ab, bc, ca\}$ denote the available wye and delta connections at bus $j \in \mathcal{V}$, respectively. Penalizing the deviation of power injection can characterize either the operational cost of controllable loads, the curtailment of photovoltaic systems, or the charging/discharging cost of batteries. The same cost expression was also used in [24], [30].

The last part minimizes the deviation of the power injections at the substation from the reference injections $\overline{p}_0, \overline{q}_0 \in \mathbb{R}$

provided by the transmission system operator. Therefore, the system operational cost function can be written as

$$f(\mathbf{s}, \mathbf{s}_\Delta) = \mu_\ell \, p_{\text{loss}} + w_p \, d_p(\mathbf{p}, \mathbf{p}_\Delta) + w_q \, d_q(\mathbf{q}, \mathbf{q}_\Delta)$$
$$+ \mu_p \, \frac{(\mathbf{1}^{\mathsf{T}}\mathbf{p}_0 - \overline{p}_0)^2}{\overline{p}_0} + \mu_q \, \frac{(\mathbf{1}^{\mathsf{T}}\mathbf{q}_0 - \overline{q}_0)^2}{\overline{q}_0}.$$

The nonnegative weights $w_p$, $w_q$, $\mu_\ell$, $\mu_p$ and $\mu_q$ are used to reflect the relative importance of the components of the cost function and are set as follows

$$w_p = w_q = \mu_\ell = 1, \qquad \mu_p = \mu_q = 4. \tag{21}$$

### B. Exactness Results for Algorithm 1

In this subsection, we assess the quality of the solutions recovered using Algorithm 1. We solve (11) for the BIM as well as (12) for the BFM with different values of voltage limits for the three considered feeders. We invoke the Mosek conic solver using CVX, a MATLAB-based convex optimization toolbox.

The left-hand side of Table I provides the result of Algorithm 1 based on the BIM. The voltage column represents the maximum and minimum voltage deviation allowed, i.e., 3% means that the value of $\overline{V}$ and $\underline{V}$ are set to 1.03 and 0.97, respectively. We assess the rank of matrices $\mathbf{W}_{jj}$, for all $j \in \mathcal{V}$, in terms of the ratio between the top two largest eigenvalues of these matrices. The maximum ratio among all $j \in \mathcal{V}$ is listed in the table. In the solution of (11) (before post-processing), the ratio between the two maximum eigenvalues of the matrices $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}$ is on the order of $10^{-1}$, and after the post-processing in Algorithm 1, the final $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}$-ratio will be dominated by $\mathbf{W}$-ratio and is thus not informative to be displayed in the table. Because of the the spectrum error, the output $\tilde{\mathbf{X}}$ could be different from $\mathbf{X}^*$, and thus having a very small $\mathbf{W}$-ratio is not enough to guarantee the feasibility of the final output of Algorithm 1. Therefore, we also assess the infeasibilty of the power flow equations by measuring the maximum violation in (5a) for the solutions returned by Algorithm 1. Here, the violation is defined as the difference between the left- and right-hand sides of (5a) (in kW) when $\mathbf{s}, \mathbf{W}, \mathbf{X}$ are evaluated as the output of Algorithm 1.

As Algorithm 1 circumvents the numerical error to a certain extent by recovering $\mathbf{I}_\Delta$ from (1b), the infeasibility is on the order of $10^{-3}$ to $10^{-1}$ kW, which reflects the effect of the spectrum error. As a benchmark, the load injections for those feeders are on the order of $10^1$ to $10^2$ kW, and are typically two orders of magnitude higher than the infeasibility.

On the right-hand side of Table I, the rank of $\mathbf{M}^{\mathbf{v}, \mathbf{S}, \ell}(j, k)$ for all lines $(j, k) \in \mathcal{E}$ is examined for the same algorithm under the BFM. Again, we present the maximum ratio between the two largest eigenvalues. Similar to the BIM, the infeasibility, i.e., the violation of (1a), is shown in the table.

### C. Exactness Results for Algorithm 2

In our setting, the penalized formulations (15) and (16) are solved with the parameter $\lambda = 10$ in all experiments. We will later show how the value of $\lambda$ affects the solution quality.

Table II presents the maximum ratio between the top two largest eigenvalues of matrices $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}(j)$ for the BIM and $\mathbf{M}^{\mathbf{v}, \mathbf{X}, \rho}(j)$ for the BFM returned by the solvers. Comparing the infeasibility of the solutions obtained using Algorithm 1, shown in Table I, and Algorithm 2, shown in Table II, it is clear that adding a penalty helps reduce the effect of the spectrum error and leads to global optimal solutions with much lower infeasibility.

To assess the effect of the penalization approach on the quality of the solutions in terms of cost and feasibility, Table III shows the effect of increasing the penalty parameter in the cost function as well as the maximum infeasibility of the power equations (in kVA) for the IEEE 37-bus network with 3% voltage limits. Note that the case $\lambda = 0$ corresponds to the output of Algorithm 1. Although the solution feasibility is enhanced by increasing the penalty parameter, the cost associated with the solution obtained also increases. Note that the cost obtained with no penalty, i.e., from Algorithm 1, represents a lower bound for the optimal cost of the original AC-OPF problem. In addition, increasing the penalty parameter beyond the values considered in Table III leads to uninteresting solutions because the cost function becomes dominated by the penalty term.

### D. Algorithm Summary and Comparison

Algorithms 1 and 2 can be useful for different applications. Algorithm 1 solves the un-penalized problem and therefore prioritizes cost minimization at the cost of larger constraint violation. This is because the recovered $\tilde{\mathbf{X}}$ could be different from $\mathbf{X}^*$ and therefore may not precisely satisfy the constraints such as (5a). That is the reason we evaluate the infeasibility as a main metric in Tables I and II. The simulation shows that the infeasibility is typically two orders of magnitude smaller than the load injections in the network and should be acceptable. Algorithm 2, on the other hand, can recover an OPF solution with much smaller constraint violation, but the optimal cost is higher because of the penalty term.

We also benchmark the computational time of the proposed algorithms for solving the AC-OPF problems in our case studies. Since both algorithms require solving similar optimization problems with different cost functions, the computational time of both algorithms is similar. Hence, we only present the computational time of Algorithm 1 to solve the AC-OPF problem for all networks using both the BIM and BFM formulations. The algorithm was implemented using Mosek as a conic solver on a laptop with Intel Core i9 CPU (2.40 GHz), 16 GB RAM, macOS Catalina OS, and MATLAB R2019b. The results show that the proposed algorithms take less than 10 seconds to solve the AC-OPF problem for the IEEE 123-bus network on a standard laptop, which demonstrates the computational efficiency of the proposed algorithms. Note that sparse semidefinite programming solvers, e.g., [31], are expected to be very efficient in solving problems with thousands of buses.

## VI. Discussion and Future Work

We now discuss the limitations and future directions of our work. First, our result guarantees the exactness when

TABLE I
RANK AND INFEASIBILITY FOR THE OUTPUTS OF ALGORITHM 1 (WITH POST-PROCESSING).

| Network | Voltage | BIM | | BFM | |
| --- | --- | --- | --- | --- | --- |
| | | $\mathbf{W}$-ratio | Infeas. (kW) | $\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}$-ratio | Infeas. (kW) |
| *IEEE-13* | 3% | $1.91 \times 10^{-8}$ | $3.20 \times 10^{-1}$ | $1.74 \times 10^{-5}$ | $4.29 \times 10^{-2}$ |
| | 5% | $2.73 \times 10^{-9}$ | $3.20 \times 10^{-1}$ | $1.58 \times 10^{-5}$ | $4.10 \times 10^{-2}$ |
| *IEEE-37* | 3% | $4.81 \times 10^{-10}$ | $9.84 \times 10^{-2}$ | $7.81 \times 10^{-5}$ | $9.67 \times 10^{-2}$ |
| | 5% | $2.77 \times 10^{-8}$ | $9.83 \times 10^{-2}$ | $2.70 \times 10^{-5}$ | $9.72 \times 10^{-2}$ |
| *IEEE-123* | 3% | $7.75 \times 10^{-8}$ | $1.54 \times 10^{-3}$ | $1.07 \times 10^{-4}$ | $1.03 \times 10^{-2}$ |
| | 5% | $7.67 \times 10^{-8}$ | $1.54 \times 10^{-3}$ | $9.64 \times 10^{-5}$ | $1.02 \times 10^{-2}$ |

TABLE II
RANK AND INFEASIBILITY FOR THE OUTPUTS OF ALGORITHM 2 (WITH PENALIZED COST FUNCTION).

| Network | Voltage | BIM | | | BFM | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | $\mathbf{W}$-ratio | $\mathbf{M}^{\mathbf{W},\mathbf{X},\rho}$-ratio | Infeas. (kW) | $\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}$-ratio | $\mathbf{M}^{\mathbf{v},\mathbf{X},\rho}$-ratio | Infeas. (kW) |
| *IEEE-13* | 3% | $1.34 \times 10^{-10}$ | $2.36 \times 10^{-9}$ | $8.85 \times 10^{-2}$ | $1.44 \times 10^{-10}$ | $1.97 \times 10^{-10}$ | $4.43 \times 10^{-5}$ |
| | 5% | $1.31 \times 10^{-10}$ | $1.96 \times 10^{-9}$ | $8.84 \times 10^{-2}$ | $1.36 \times 10^{-10}$ | $1.57 \times 10^{-10}$ | $1.46 \times 10^{-5}$ |
| *IEEE-37* | 3% | $3.04 \times 10^{-8}$ | $6.22 \times 10^{-8}$ | $5.75 \times 10^{-6}$ | $8.85 \times 10^{-8}$ | $3.38 \times 10^{-5}$ | $1.45 \times 10^{-6}$ |
| | 5% | $2.94 \times 10^{-8}$ | $1.05 \times 10^{-8}$ | $1.06 \times 10^{-6}$ | $2.12 \times 10^{-8}$ | $3.18 \times 10^{-5}$ | $1.00 \times 10^{-6}$ |
| *IEEE-123* | 3% | $1.45 \times 10^{-9}$ | $7.03 \times 10^{-9}$ | $9.13 \times 10^{-7}$ | $1.05 \times 10^{-8}$ | $8.99 \times 10^{-9}$ | $1.34 \times 10^{-6}$ |
| | 5% | $1.94 \times 10^{-8}$ | $9.31 \times 10^{-8}$ | $7.84 \times 10^{-6}$ | $7.98 \times 10^{-9}$ | $6.59 \times 10^{-9}$ | $1.40 \times 10^{-6}$ |

TABLE III
EFFECT OF THE PENALTY PARAMETER ON THE COST AND INFEASIBILITY.

| $\lambda$ | BIM | | BFM | |
| --- | --- | --- | --- | --- |
| | Cost | Infeas. (kW) | Cost | Infeas. (kW) |
| 0 | 100.0036 | $9.84 \times 10^{-2}$ | 100.0194 | $9.67 \times 10^{-2}$ |
| 0.1 | 103.9504 | $1.15 \times 10^{-2}$ | 104.7141 | $5.99 \times 10^{-4}$ |
| 1 | 104.7846 | $6.00 \times 10^{-5}$ | 104.7840 | $1.80 \times 10^{-5}$ |
| 10 | 104.7886 | $5.75 \times 10^{-6}$ | 104.7982 | $1.45 \times 10^{-6}$ |
| 100 | 104.9431 | $3.17 \times 10^{-6}$ | 105.0332 | $1.23 \times 10^{-7}$ |

TABLE IV
COMPUTATIONAL TIME FOR BOTH BIM AND BFM FORMULATIONS (IN SECONDS).

| Model | IEEE-13 | IEEE-37 | IEEE-123 |
| --- | --- | --- | --- |
| BIM | 2.46 | 5.56 | 9.03 |
| BFM | 2.77 | 4.93 | 9.87 |

$\mathbf{W}$ or $\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}$ is of rank 1. When $\mathbf{W}$ or $\mathbf{M}^{\mathbf{v},\mathbf{S},\ell}$ is not of rank 1, our algorithms will fail to output feasible solutions but can still provide a lower bound for the optimal value. Reference [22] provides a sufficient condition to guarantee that $\mathbf{W}$ is of rank 1 for unbalanced systems without delta connections. It would be interesting to extend the result to networks with delta connections. For single-phase networks, it is also known that when SDP or SOCP relaxation fails to provide a feasible solution, a heuristic post-processing can be applied to encourage feasibility. Another direction is whether the same techniques can be applied when the system is unbalanced and contains delta connections.

Second, the main focus of this paper is on relaxing the non-convexity due to quadratic power flow equations. There are other sources of non-convexity in OPF problems. For instance, for systems with controllable switches and voltage regulators, there are additional decision variables which are discrete, and the underlying non-convexity is beyond the scope of this paper. The modelling of those devices has been studied in [32], [33]. Including switches and voltage regulators in the formulation of three-phase networks with delta connections is an important future work for more applications in distribution networks.

Finally, there is a trade-off between relaxation-based algorithms and local search algorithms for general non-convex programs. On one hand, local search algorithms are faster and more scalable for large systems compared to SDP relaxation. On the other hand, there is no guarantee that a local search algorithm will converge, and when it does the output is in general a local optimum. There are also many active lines of work on improving the computational efficiency of convex relaxations, and combining the relaxation methods with local algorithms. More discussion could be found in the survey [13] and references therein.

## VII. CONCLUSION

This paper studied the SDP relaxation of the AC-OPF problem for an unbalanced three-phase radial network with delta connections, formulated under both the BIM and BFM. we showed the equivalence between the BIM and BFM formulations and presented sufficient conditions for recovering exact solutions of the nonconvex AC-OPF formulations from their respective relaxations. The paper also showed why conventional relaxation (by directly dropping rank-1 constraints) always fail even if the sufficient conditions we proposed are satisfied. It is due to the non-uniqueness in the relaxation solution and the spectrum error in computation. Inspired by this finding, we then proposed two algorithms which are guaranteed to produce exact solutions whenever our sufficient conditions are satisfied. One applies post-processing and produces lower cost but larger constraint violation. The other adds a penalty term and produces higher cost but smaller constraint violation. In simulations, we demonstrated that for three IEEE standard test cases, both algorithms are able to recover near global optimal solutions with tolerable constraint violation and cost sub-optimality.

## REFERENCES

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

[2] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.

[3] 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.

[4] 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.

[5] D. Wells, "Method for economic secure loading of a power system," in *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8. IET, 1968, pp. 1190–1194.

[6] R. Burchett, H. Happ, and K. Wirgau, "Large scale optimal power flow," *IEEE Transactions on Power Apparatus and Systems*, no. 10, pp. 3722–3732, 1982.

[7] A. Bernstein, C. Wang, E. DallAnese, J.-Y. Le Boudec, and C. Zhao, "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5832–5843, 2018.

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

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

[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] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *IEEE Power and Energy Society General Meeting*, 2012, pp. 1–7.

[12] V. Kekatos, G. Wang, A. J. Conejo, and G. B. Giannakis, "Stochastic reactive power management in microgrids with renewables," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3386–3395, Dec. 2014.

[13] D. K. Molzahn and I. A. Hiskens, "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.

[14] S. H. Low, "Convex relaxation of optimal power flow–part II: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, 2014.

[15] 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.

[16] A. S. Zamzam, C. Zhao, E. Dall'Anese, and N. D. Sidiropoulos, "A QCQP approach for OPF in multiphase radial networks with wye and delta connections," in *Proceedings of the 10th Bulk Power Systems Dynamics and Control Symposium*, Espinho, Portugal, 2017.

[17] R. Berg, E. Hawkins, and W. Pleines, "Mechanized calculation of unbalanced load flow on radial distribution circuits," *IEEE Transactions on power apparatus and systems*, no. 4, pp. 415–421, 1967.

[18] 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.

[19] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, 2014.

[20] 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.

[21] 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.

[22] F. Zhou, Y. Chen, and S. H. Low, "Sufficient conditions for exact semidefinite relaxation of optimal power flow in unbalanced multiphase radial networks," *arXiv preprint arXiv:1910.08258*, 2019.

[23] W. Wang and N. Yu, "Chordal conversion based convex iteration algorithm for three-phase optimal power flow problems," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1603–1613, 2017.

[24] C. Zhao, E. Dall'Anese, and S. H. Low, "Optimal power flow in multiphase radial networks with delta connections," National Renewable Energy Lab. (NREL), Golden, CO (United States), Tech. Rep., 2017.

[25] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1297–1307, 2015.

[26] D. K. Molzahn, C. Josz, I. A. Hiskens, and P. Panciatici, "A Laplacian-based approach for finding near globally optimal solutions to OPF problems," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 305–315, 2017.

[27] S. Bose, S. H. Low, T. Teeraratkul, and B. Hassibi, "Equivalent relaxations of optimal power flow," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 729–742, 2014.

[28] S. H. Low, "Convex relaxation of optimal power flow-part I: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

[29] K. Schneider, B. Mather, B. Pal, C.-W. Ten, G. Shirek, H. Zhu, J. Fuller, J. Pereira, L. Ochoa, L. De Araujo *et al.*, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Transactions on power systems*, vol. 33, no. 3, pp. 3181–3188, 2017.

[30] E. DallAnese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1868–1881, 2017.

[31] R. Y. Zhang and J. Lavaei, "Sparse semidefinite programs with near-linear time complexity," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 1624–1631.

[32] M. Bazrafshan, N. Gatsis, and H. Zhu, "Optimal power flow with step-voltage regulators in multi-phase distribution networks," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4228–4239, 2019.

[33] H.-D. Chiang and R. Jean-Jumeau, "Optimal network reconfigurations in distribution systems: Part 1: A new formulation and a solution methodology," *IEEE Transactions on Power Delivery*, vol. 5, no. 4, pp. 1902–1909, 1990.