Zero Duality Gap in Optimal Power Flow Problem

Javad Lavaei and Steven H. Low

Abstract— **The optimal power flow (OPF) problem is nonconvex and generally hard to solve. We provide a sufficient condition under which the OPF problem is equivalent to a convex problem and therefore is efficiently solvable. Specifically, we prove that the dual of OPF is a semidefinite program and our sufficient condition guarantees that the duality gap is zero and a globally optimal solution of OPF is recoverable from a dual optimal solution. This sufficient condition is satisfied by standard IEEE benchmark systems with 14, 30, 57, 118 and 300 buses after small resistance (10^{-5} per unit) is added to every transformer that originally assumes zero resistance. We justify why the condition might hold widely in practice from algebraic and geometric perspectives. The main underlying reason is that physical quantities such as resistance, capacitance and inductance, are all positive.**

Index Terms— **Power System, Optimal Power Flow, Convex Optimization, Linear Matrix Inequality, Polynomial-Time Algorithm.**

I. Introduction

Optimal power flow (OPF) problem deals with finding an optimal operating point of a power system that minimizes an appropriate cost function such as generation cost or transmission loss subject to certain constraints on power and voltage variables [1]. Started by the work [2] in 1962, the OPF problem has been extensively studied in the literature and numerous algorithms have been proposed for solving this highly nonconvex problem [3], [4], [5], including linear programming, Newton Raphson, quadratic programming, nonlinear programming, Lagrange relaxation, interior point methods, artificial intelligence, artificial neural network, fuzzy logic, genetic algorithm, evolutionary programming and particle swarm optimization [1], [6], [7], [8]. A good number of these methods are based on the Karush-Kuhn-Tucker (KKT) necessary conditions, which can only guarantee a locally optimal solution due to nonconvexity of the OPF problem [9]. This nonconvexity is partially due to the cross products of voltage variables corresponding to disparate buses. In the past decade, much attention has been paid to devising efficient algorithms with guaranteed performance for the OPF problem. For instance, the recent papers [10] and [11] propose nonlinear interior-point algorithms for an equivalent current injection model of the problem. An improved implementation of the automatic differentiation technique for the OPF problem is studied in the recent work [12]. In an effort to convexity the OPF problem, it is shown in [13] that the load flow problem of a radial distribution system can be modeled as a convex optimization problem in the form of a conic program. Nonetheless, the results fail to hold for a meshed network, due to the presence of arctangent equality constraints [14]. Nonconvexity appears in more sophisticated power problems such as the stability constrained OPF problem where the stability at the operating point is an extra constraint [15], [16] or the dynamic OPF problem where the dynamics of the generators are also taken into account [17], [18].

The OPF problem is in general NP-hard [19]. We also showed in our recent work that a closely related problem of finding an optimal operating point of a radiating antenna circuit is an NP-complete problem, by reducing the number partitioning problem to the antenna problem [20]. Using duality theory and semidefinite programming, however, we will show in this paper that a power system has special structure (Condition C0(ii) below) that often renders the OPF problems efficiently solvable.

Specifically, instead of solving the OPF problem directly, we propose solving its Lagrangian dual problem, and recover a primal solution from a dual optimal solution. We prove that the dual problem is a convex semidefinite program and therefore can be solved efficiently. However, the optimal objective value of the dual problem is only a lower bound on the optimal value of the original OPF problem and the lower bound may not be tight (nonzero duality gap) [21]. If the primal solution computed from an optimal dual solution indeed satisfies all the constraints of the OPF problem and the resulting objective value equals the optimal dual objective value (zero duality gap), then strong duality holds and the primal solution is indeed optimal for the original OPF problem. This approach has allowed us to solve exactly (globally optimal) and efficiently all the five IEEE benchmark systems archived at [22] with 14, 30, 57, 118 and 300 buses. Our main result (Theorem 1) provides a *sufficient* condition (C1 below) that guarantees zero duality gap and optimality of the resulting OPF solution. This is explained in Section II and proved in Section III through clarifying the duality structure of the OPF problem.

Therefore, even though the OPF problem is NP-hard in general, a subset of the problem instances that satisfy condition C1 are equivalent to its convex dual. Although the sufficient condition C1 is not satisfied by the IEEE benchmark systems, the duality gap is zero for all of them. Moreover, C1 is violated in a "trivial" manner in these systems: when it is violated, it is due to the simplifying assumption that transformers have zero resistance. When even a small resistance (10^{-5} per unit) is added to each transformer that originally assumes zero resistance, condition C1 is satisfied for all the IEEE benchmark systems. In Section IV, we provide informal justification on why condition C1 might hold widely, both from an algebraic and a geometric perspective. The geometrical interpretation is that the feasibility region of the dual of the OPF problem must be smooth on its boundary around the optimal point. The algebraic argument relies on the Perron-Frobenius theorem in graph theory and implies that the zero duality gap is

Javad Lavaei is with the Department of Control and Dynamical Systems, California Institute of Technology (email: lavaei@cds.caltech.edu).

Steven H. Low is with Computer Science and Electrical Engineering Departments, California Institute of Technology (email: slow@caltech.edu).

due to physical quantities, namely resistance, capacitance and inductance, being positive (condition C0(ii)). In other words, it is a consequence of the physical constraints that nature imposes that many OPF problems have zero duality gap. This suggests hope that the OPF problem for practical networks may be efficiently solvable using the algorithm prescribed in Section II. Generalizations to the basic OPF formulation are discussed in Section V. The various results are illustrated in Section VI through IEEE benchmark systems and smaller examples. Concluding remarks are drawn in Section VII. Some background on semidefinite programming is provided in Appendix A and, finally, a few proofs are collected in Appendix B.

Notations: We introduce the following notations:

- i : The imaginary unit.
- $\mathbf{R}$: The set of real numbers.
- Re$\{\cdot\}$ and Im$\{\cdot\}$: The operators returning the real and imaginary parts of a complex matrix.
- $*$: The conjugate transpose operator.
- T : The transpose operator.
- $\succeq$ and $\preceq$: The matrix inequality signs in the positive semidefinite sense (i.e. given two symmetric matrices A and B, $A \succeq B$ implies $A - B$ is a positive semidefinite matrix, meaning that its eigenvalues are all nonnegative).
- Trace: The matrix trace operator.
- $|\cdot|$: The absolute value operator.
- For any vector x, x_i generally denotes the ith component.

II. Problem Formulation and Main Result

A. Problem formulation

Consider a power network with n buses, labeled $1, ..., n$, where all buses are possibly directly connected to loads, but only the first m buses are directly connected to generators. For $k \in \{1, ..., n\}$ and $l \in \{1, ..., m\}$, define the following quantities:

- P_k^d and Q_k^d: Active and reactive loads at buses k, respectively. They are given demands that must be met.
- P_l^g and Q_l^g: Active and reactive powers generated at buses l, respectively. They are optimization variables.
- V_k: Complex voltages at buses k. They are optimization variables.
- $f_l(P_l^g) = c_{l2}(P_l^g)^2 + c_{l1}P_l^g + c_{l0}$: Cost functions associated with generators l, where c_{l2}, c_{l1}, c_{l0} are nonnegative numbers.

Derive the circuit model of the power network by replacing every transmission line and transformer with their equivalent Π models [1]. In this circuit model, let y_{kl} be the mutual admittance between buses k and l, and y_{kk} be the admittance-to-ground at bus k, for every $l, k \in \{1, ..., n\}$. Denote the admittance matrix of this equivalent circuit model with Y, which is an $n \times n$ complex-valued matrix whose (l, k) entry is equal to $-y_{lk}$ if $l \neq k$ and $y_{ll} + \sum_{p \in \mathcal{N}(l)} y_{lp}$ otherwise, where $\mathcal{N}(l)$ is the set of buses that are directly connected to bus l. Denote by the column vector $V := (V_k, k = 1, \ldots, n)$ the complex voltages. Define the current vector $I := YV = (I_k, k = 1, ..., n)$. Let $P^g := (P_1^g, l = 1, \ldots, m)$ and $Q^g := (Q_1^g, l = 1, \ldots, m)$.

The classical optimal power flow (OPF) problem is:

OPF:

$$\min_{V, P^g, Q^g} \quad \sum_{l=1}^{m} f_l(P_l^g) \tag{1}$$

subject to

$$P_l^{\min} \leq P_l^g \leq P_l^{\max}, \quad l = 1, 2, ..., m \tag{2a}$$

$$Q_l^{\min} \leq Q_l^g \leq Q_l^{\max}, \quad l = 1, 2, ..., m \tag{2b}$$

$$V_k^{\min} \leq |V_k| \leq V_k^{\max}, \quad k = 1, 2, ..., n \tag{2c}$$

$$V_l I_l^* = (P_l^g - P_l^d) + (Q_l^g - Q_l^d)\mathrm{i}, \quad l = 1, 2, ..., m \tag{2d}$$

$$V_k I_k^* = -P_k^d - Q_k^d \mathrm{i}, \quad k = m+1, ..., n \tag{2e}$$

The inequalities (2a), (2b) and (2c) limit the power and voltage variables to within the given bounds $P_l^{\min}, P_l^{\max}, Q_l^{\min}, Q_l^{\max}, V_k^{\min}, V_k^{\max}$, whereas the last two equations (2d) and (2e) express the physical constraints imposed by the network. There could be more constraints in the OPF problem, e.g. line flow limits, which we will discuss in Section V below.

Though not stated explicitly in the results that follow, we assume the following condition to hold throughout the paper:

C0: (i) OPF (1)–(2) is feasible. Moreover, $V = 0$ is not a feasible point of OPF.

(ii) The admittance matrix Y is symmetric ($Y_{ij} = Y_{ji}$) and has two important properties: the off-diagonal entries of the matrix Re$\{Y\}$ are all nonpositive, and the off-diagonal entries of the matrix Im$\{Y\}$ are all nonnegative.

Assumption C0(i) is to avoid triviality. Assumption C0(ii) always holds in standard power systems where the resistance, capacitance and inductance in the Π model of transmission lines are positive.

B. Main result

The voltage constraints (2c) and the network constraints (2d)–(2e) are the sources of nonconvexity that makes OPF generally hard. Our goal is to derive a sufficient condition under which the OPF problem is equivalent to a convex problem, and hence can be solved efficiently. Moreover, we will demonstrate in later sections that this sufficient condition is (essentially) satisfied by all the IEEE benchmark systems archived at [22] and provide informal justifications on why the condition is likely to hold in practice. To state our main result, we need the following notations.

Eliminating the variables $P_l^g = \mathrm{Re}\{Y_l I_l^*\} + P_l^d$ and $Q_l^g = \mathrm{Im}\{Y_l I_l^*\} + Q_l^d$ using the network constraints (2d) and (2e), we can write the OPF problem in terms only of the complex voltages V (noting $I = YV$). Extend the definition of $P_k^{\min}, P_k^{\max}, Q_k^{\min}, Q_k^{\max}$ to $k \in \{m+1, ..., n\}$, with $P_k^{\min} = P_k^{\max} = Q_k^{\min} = Q_k^{\max} = 0$ if $k \in \{m+1, ..., n\}$. Let $e_1, e_2, ..., e_n$ denote the standard basis vectors in $\mathbf{R}^n$. For every $k = 1, 2, ..., n$, define $M_k \in \mathbf{R}^{2n \times 2n}$ as a diagonal matrix whose entries are all equal to zero, except for its (k, k)

and $(n+k, n+k)$ entries that are equal to 1. Define also

$$\begin{aligned} Y_k &:= e_k e_k^* Y \\ \mathbf{Y}_k &:= \frac{1}{2}\begin{bmatrix} \mathrm{Re}\{Y_k + Y_k^T\} & \mathrm{Im}\{Y_k^T - Y_k\} \\ \mathrm{Im}\{Y_k - Y_k^T\} & \mathrm{Re}\{Y_k + Y_k^T\} \end{bmatrix} \\ \bar{\mathbf{Y}}_k &:= \frac{-1}{2}\begin{bmatrix} \mathrm{Im}\{Y_k + Y_k^T\} & \mathrm{Re}\{Y_k - Y_k^T\} \\ \mathrm{Re}\{Y_k^T - Y_k\} & \mathrm{Im}\{Y_k + Y_k^T\} \end{bmatrix} \end{aligned}$$

Define the variables for the dual problem as a $6n$-dimensional real vector:

$$x := (\lambda_k^{\min}, \lambda_k^{\max}, \bar{\lambda}_k^{\min}, \bar{\lambda}_k^{\max}, \mu_k^{\min}, \mu_k^{\max},\ k = 1, ..., n)$$

and a $2m$-dimensional real vector

$$r := (r_{l1}, r_{l2},\ l = 1, ..., m)$$

Define the affine function

$$\begin{aligned} h(x, r) := \sum_{k=1}^{n} \Big\{ & \lambda_k^{\min} P_k^{\min} - \lambda_k^{\max} P_k^{\max} + \boldsymbol{\lambda}_k P_k^d \\ & + \bar{\lambda}_k^{\min} Q_k^{\min} - \bar{\lambda}_k^{\max} Q_k^{\max} + \bar{\boldsymbol{\lambda}}_k Q_k^d + \mu_k^{\min}\left(V_k^{\min}\right)^2 \\ & - \mu_k^{\max}\left(V_k^{\max}\right)^2 \Big\} + \sum_{l=1}^{m} (c_{l0} - r_{l2}) \end{aligned}$$

where the bold variables are defined in terms of (x, r) as: for $k = 1, ..., n$

$$\begin{aligned} \boldsymbol{\lambda}_k &:= \begin{cases} -\lambda_k^{\min} + \lambda_k^{\max} + c_{k1} + 2\sqrt{c_{k2}}\, r_{k1} & \text{if } k = 1, ..., m \\ -\lambda_k^{\min} + \lambda_k^{\max} & \text{otherwise} \end{cases} \\ \bar{\boldsymbol{\lambda}}_k &:= -\bar{\lambda}_k^{\min} + \bar{\lambda}_k^{\max} \\ \boldsymbol{\mu}_k &:= -\mu_k^{\min} + \mu_k^{\max} \end{aligned}$$

Instead of the nonconvex OPF problem, we propose solving the following convex problem.

Dual OPF:

$$\max_{x \geq 0, r}\ h(x, r) \tag{3}$$

subject to

$$\sum_{k=1}^{n} \left(\boldsymbol{\lambda}_k \mathbf{Y}_k + \bar{\boldsymbol{\lambda}}_k \bar{\mathbf{Y}}_k + \boldsymbol{\mu}_k M_k\right) \succeq 0 \tag{4a}$$

$$\begin{bmatrix} 1 & r_{l1} \\ r_{l1} & r_{l2} \end{bmatrix} \succeq 0, \quad l = 1, 2, ..., m \tag{4b}$$

This semidefinite program is the dual of an equivalent form of OPF (see Section III-A for more details and Appendix A for a brief overview of semidefinite programming). It is therefore convex and can be solved efficiently. This motivates the following approach to solving OPF.

Algorithm for Solving OPF:

1) Compute a solution $(x^{\text{opt}}, r^{\text{opf}})$ of Dual OPF (3)–(4).
2) If the optimal value of Dual OPF is $+\infty$, then OPF is infeasible.
3) Compute any nonzero vector $\begin{bmatrix} U_1^T & U_2^T \end{bmatrix}^T$ in the null space of the $2n \times 2n$ positive semidefinite matrix

$$A^{\text{opt}} := \sum_{k=1}^{n} \left(\boldsymbol{\lambda}_k^{\text{opt}} \mathbf{Y}_k + \bar{\boldsymbol{\lambda}}_k^{\text{opt}} \bar{\mathbf{Y}}_k + \boldsymbol{\mu}_k^{\text{opt}} M_k\right) \tag{5}$$

4) Compute an optimal solution V^{opt} of OPF as

$$V^{\text{opt}} = (\zeta_1 + \zeta_2 \mathrm{i})(U_1 + U_2 \mathrm{i}) \tag{6}$$

by solving for ζ_1 and ζ_2 from optimality conditions.
5) Verify that V^{opt} satisfies all the constraints of OPF (1)–(2) and that the resulting objective value of OPF equals the optimal value of Dual OPF (zero duality gap).

We make several remarks. First, provided OPF is feasible, the null space of A^{opt} has an even dimension of at least 2 (see proof of Theorem 1 below). Hence Step 3 of the Algorithm will always yield a nonzero vector $\begin{bmatrix} U_1^T & U_2^T \end{bmatrix}^T$. Second, having found U_1 and U_2, the scalars ζ_1 and ζ_2 can be identified from the first order optimality (KKT) condition for Dual OPF or the feasibility condition for OPF. For instance, the voltage angle at the swing bus being zero introduces an equation in terms of ζ_1 and ζ_2. If, in addition, $(\mu_k^{\min})^{\text{opt}}$ (respectively, $(\mu_k^{\max})^{\text{opt}}$) turns out to be nonzero for some $k \in \{1, 2, ...n\}$, then the relation $|V_k^{\text{opt}}| = V_k^{\min}$ (respectively, $|V_k^{\text{opt}}| = V_k^{\max}$) must hold by complementary slackness, which provides another equation relating ζ_1 to ζ_2. Third, the weak duality theorem implies that the optimal value of OPF is greater than or equal to that of its dual. Hence, Step 2 detects when OPF is infeasible. Even when OPF is feasible, there is generally a nonzero duality gap and an optimal solution to OPF may not be recoverable from an optimal dual solution. However, if V^{opt} computed in Step 4 indeed is primal feasible as verified in Step 5, then duality gap is zero and V^{opt} is indeed optimal for OPF. This is the case with all the IEEE benchmark examples described in Section VI, and hence all of them can be solved efficiently by the above Algorithm.

Indeed, the following *sufficient* condition guarantees that the Algorithm finds an optimal solution of OPF:

C1: There exists a dual optimal solution $(x^{\text{opt}}, r^{\text{opt}})$ such that the $2n \times 2n$ positive semidefinite matrix A^{opt} in (5) has a zero eigenvalue of multiplicity 2.

In this case, the null space of A^{opt} has dimension 2.

Theorem 1: If condition C1 holds, then

1) There is no duality gap between OPF and Dual OPF.
2) Given any vector $\begin{bmatrix} U_1^T & U_2^T \end{bmatrix}^T$ in the null space of A^{opt}, the voltages V^{opt} calculated in (6) is indeed optimal for OPF.

III. Convex Relaxations of OPF

In this section, we provide further insights on the structure of OPF (1)–(2) and prove our main result (Theorem 1). We will do this by defining four optimization problems, clarifying their relationship as summarized in Figure 1, and showing how they imply Theorem 1.

A. Duality structure of OPF

As alluded to above, we can eliminate the variables $P_l^g = \mathrm{Re}\{Y_l I_l^*\} + P_l^d$ and $Q_l^g = \mathrm{Im}\{Y_l I_l^*\} + Q_l^d$ using the network constraints (2d) and (2e) to write OPF in the following equivalent form.

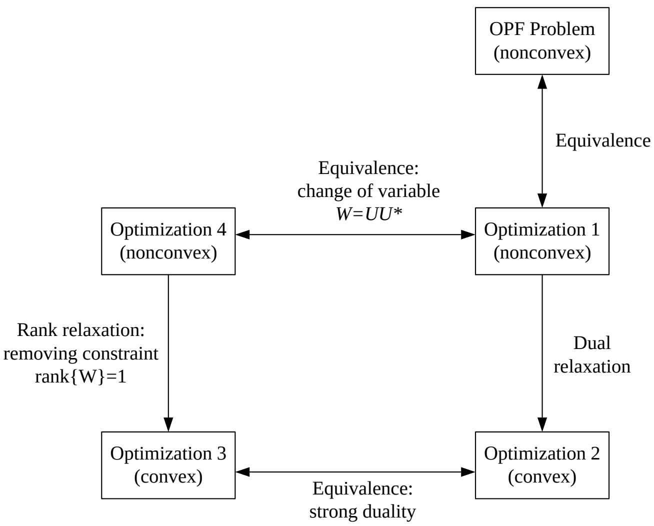


Fig. 1. The relationship among OPF and Optimizations 1-4.

Optimization 1:

$$\min_{\alpha, V} \qquad \sum_{l=1}^{m} \alpha_l \tag{7}$$

subject to

$$P_k^{\min} - P_k^d \le \text{Re}\{V_k I_k^*\} \le P_k^{\max} - P_k^d \tag{8a}$$
$$Q_k^{\min} - Q_k^d \le \text{Im}\{V_k I_k^*\} \le Q_k^{\max} - Q_k^d \tag{8b}$$
$$\left(V_k^{\min}\right)^2 \le |V_k|^2 \le (V_k^{\max})^2 \tag{8c}$$
$$\begin{bmatrix} c_{l1}\text{Re}\{V_l I_l^*\} - \alpha_l + a_l & \sqrt{c_{l2}}\text{Re}\{V_l I_l^*\} + b_l \\ \sqrt{c_{l2}}\text{Re}\{V_l I_l^*\} + b_l & -1 \end{bmatrix} \preceq 0 \tag{8d}$$

for $k = 1, 2, ..., n$ and $l = 1, 2, ..., m$. Here $I = YV$ and the problem parameters are $a_l := c_{l0} + c_{l1}P_l^d$, $b_l := \sqrt{c_{l2}}P_l^d$.

Optimization 2: This is Dual OPF (3)–(4).

We first establish that Dual OPF is a dual relaxation of OPF. Its proof is relegated to the Appendix.

Theorem 2: We have

1) Optimization 1 is equivalent to OPF (1)–(2).
2) Optimization 2 is the dual of Optimization 1 with (x, r) as the Lagrange multipliers.

By weak duality, Dual OPF (3)–(4) provides a lower bound on OPF. To prove that the bound is exact (zero duality gap) under condition C1 and to recover an optimal solution to OPF, we need another pair of optimization problems. To motivate, observe that for $k = 1, \ldots, n$:

$$\begin{aligned} \text{Re}\{V_k I_k^*\} &= \text{Re}\{V^* e_k e_k^* I\} = \text{Re}\{V^* Y_k V\} \\ &= U^T \begin{bmatrix} \text{Re}\{Y_k\} & -\text{Im}\{Y_k\} \\ \text{Im}\{Y_k\} & \text{Re}\{Y_k\} \end{bmatrix} U \\ &= \frac{1}{2} U^T \begin{bmatrix} \text{Re}\{Y_k + Y_k^T\} & \text{Im}\{Y_k^T - Y_k\} \\ \text{Im}\{Y_k - Y_k^T\} & \text{Re}\{Y_k + Y_k^T\} \end{bmatrix} U \\ &= U^T \mathbf{Y}_k U = \text{trace}\left\{\mathbf{Y}_k U U^T\right\} \end{aligned} \tag{9}$$

where

$$U := \begin{bmatrix} \text{Re}\{V\}^T & \text{Im}\{V\}^T \end{bmatrix}^T \tag{10}$$

Likewise,

$$\text{Im}\{V_k I_k^*\} = U^T \bar{\mathbf{Y}}_k U = \text{trace}\left\{\bar{\mathbf{Y}}_k U U^T\right\} \tag{11}$$

Furthermore,

$$|V_k|^2 = U^T M_k U = \text{trace}\left\{M_k U U^T\right\} \tag{12}$$

Substituting these expressions into Optimization 1 and identifying $W = UU^T$ motivate Optimizations 3 and 4.

Optimization 3:

$$\min_{\alpha, W} \qquad \sum_{l=1}^{m} \alpha_l$$

where $W \in \mathbf{R}^{2n \times 2n}$ denotes symmetric matrices, subject to

$$\begin{aligned} & P_k^{\min} - P_k^d \le \text{trace}\{\mathbf{Y}_k W\} \le P_k^{\max} - P_k^d \\ & Q_k^{\min} - Q_k^d \le \text{trace}\{\bar{\mathbf{Y}}_k W\} \le Q_k^{\max} - Q_k^d \\ & \left(V_k^{\min}\right)^2 \le \text{trace}\{M_k W\} \le (V_k^{\max})^2 \\ & \begin{bmatrix} c_{l1}\text{trace}\{\mathbf{Y}_k W\} - \alpha_l + a_l & \sqrt{c_{l2}}\text{trace}\{\mathbf{Y}_k W\} + b_l \\ \sqrt{c_{l2}}\text{trace}\{\mathbf{Y}_k W\} + b_l & -1 \end{bmatrix} \preceq 0 \\ & W \succeq 0 \end{aligned}$$

for $k = 1, 2, ..., n$ and $l = 1, 2, ..., m$.

Optimization 4: This is Optimization 3 with an additional constraint $\text{rank}\{W\} = 1$.

The next result, proved in the Appendix, completes the relationship depicted in Figure 1 among these optimization problems.

Theorem 3: We have

1) Optimization 4 is equivalent to Optimization 1 via the change of variable $W = UU^T$, i.e.,

$$W = \begin{bmatrix} \text{Re}\{V\}^T & \text{Im}\{V\}^T \end{bmatrix}^T \begin{bmatrix} \text{Re}\{V\}^T & \text{Im}\{V\}^T \end{bmatrix}$$

2) Optimization 3 is the Lagrangian dual of Optimization 2. Moreover, strong duality holds between them.

It is hard to prove directly our main result that condition C1 implies zero duality gap between Optimizations 1 (hence OPF) and 2. Figure 1 suggests a different proof approach, as follows. The rank of a symmetric matrix is the number of its nonzero eigenvalues. Optimization 4 is not convex because of the additional constraint $\text{rank}\{W\} = 1$. Optimization 3 is therefore a rank (convex) relaxation of Optimization 4. But Optimization 3 is equivalent to Optimization 2 through strong duality (Theorem 3(2)), and hence, when the rank relaxation turns out to be exact (i.e., Optimizations 3 and 4 are equivalent), Optimization 2 will be equivalent to Optimization 4 and therefore to OPF (Theorem 3(1) and Theorem 2(1)).

We now prove Theorem 1 by showing that condition C1 closes the gap between Optimizations 3 and 4, i.e., it guarantees that any solution of Optimization 3 always satisfies the constraint $\text{rank}\{W\} = 1$ and hence is also a solution of Optimization 4.

B. Proof of Main Result Theorem 1

A^{opt} has a simple structure, which becomes more transparent with an alternative representation, as follows. Define the following matrices (that depend implicitly on a feasible solution (x, r) of Optimization 2): [1]

$$\begin{aligned}\mathbf{\Lambda} &:= \text{diag}\{\boldsymbol{\lambda}_1, ..., \boldsymbol{\lambda}_n\} \\ \bar{\mathbf{\Lambda}} &:= \text{diag}\{\bar{\boldsymbol{\lambda}}_1, ..., \bar{\boldsymbol{\lambda}}_n\} \\ \mathbf{\Gamma} &:= \text{diag}\{\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_n\}\end{aligned}$$

where diag$\{\cdot\}$ maps its vector argument to a diagonal matrix. Denote by $A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ the matrix

$$A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) = \sum_{k=1}^{n} \left(\boldsymbol{\lambda}_k \mathbf{Y}_k + \bar{\boldsymbol{\lambda}}_k \bar{\mathbf{Y}}_k + \boldsymbol{\mu}_k M_k \right)$$

Then $A^{\text{opt}} = A(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}})$. An alternative representation is:

$$A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) = \frac{1}{2} \begin{bmatrix} H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) & H_2(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) \\ -H_2(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) & H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) \end{bmatrix} \tag{13}$$

where

$$\begin{aligned} H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) &:= \mathbf{\Lambda} \times \text{Re}\{Y\} + \text{Re}\{Y\} \times \mathbf{\Lambda} \\ &\quad - \bar{\mathbf{\Lambda}} \times \text{Im}\{Y\} - \text{Im}\{Y\} \times \bar{\mathbf{\Lambda}} + 2\mathbf{\Gamma} \\ H_2(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma}) &:= -\bar{\mathbf{\Lambda}} \times \text{Re}\{Y\} + \text{Re}\{Y\} \times \bar{\mathbf{\Lambda}} \\ &\quad - \mathbf{\Lambda} \times \text{Im}\{Y\} + \text{Im}\{Y\} \times \mathbf{\Lambda}. \end{aligned}$$

We will use these expressions in the proof.

1) As stated in the proof of Theorem 3, Optimization 3 is the dual of Optimization 2 with its variable W playing the role of a Lagrange multiplier for the matrix constraint (4a) in Optimization 2 (see the Appendix). One can write the KKT conditions for Optimization 2 to obtain

$$\text{trace}\left\{A^{\text{opt}} W^{\text{opt}}\right\} = 0 \tag{14}$$

where A^{opt} is the matrix in condition C1. Denote the nonzero eigenvalues of W^{opt} as $a_1, ..., a_f$ and their associated eigenvectors as $E_1, ..., E_f$ for some nonnegative integer f. Writing $W^{\text{opt}} = \sum_{l=1}^{f} a_l E_l E_l^T$, it follows from (14) that

$$\sum_{l=1}^{f} a_l E_l^T A^{\text{opt}} E_l = 0 \tag{15}$$

Furthermore, the constraints

$$W^{\text{opt}} \succeq 0, \quad A^{\text{opt}} \succeq 0$$

in Optimizations 2 and 3 imply that $a_1, ..., a_f$ are all positive and $E_l^T A^{\text{opt}} E_l$ are all nonnegative for $l = 1, \ldots, f$. Therefore, the equality (15) holds if and only if $E_l^T A^{\text{opt}} E_l = 0$ for $l = 1, \ldots, f$. Since A^{opt} is positive semidefinite, this is equivalent to

$$A^{\text{opt}} E_l \quad = \quad 0, \quad l = 1, ..., f$$

This implies that the orthogonal eigenvectors $E_1, ..., E_f$ all belong to the null space of A^{opt}, which, under condition C1, has dimension 2. Hence $f \leq 2$.

Since OPF is feasible by condition C0(i), $f > 0$ as W must be nonzero. If $f = 1$, then rank$\{W\} = 1$ and hence the solution W of Optimization 3 must also be a solution of Optimization 4.

If $f = 2$, let E_1 and E_2 be two orthogonal eigenvectors of A^{opt} associated with its zero eigenvalue. Decompose E_1 as $\begin{bmatrix} E_{11}^T & E_{12}^T \end{bmatrix}^T$ for some vectors $E_{11}, E_{12} \in \mathbf{R}^n$. ¿From the expression (13) for A^{opt}, the only vector (up to a constant factor) in the null space of A^{opt} that is orthogonal to E_1 is $E_2 = \begin{bmatrix} -E_{12}^T & E_{11}^T \end{bmatrix}^T$. Therefore

$$\begin{aligned} W^{\text{opt}} = a_1 &\begin{bmatrix} E_{11} \\ E_{12} \end{bmatrix} \begin{bmatrix} E_{11}^T & E_{12}^T \end{bmatrix} \\ &+ a_2 \begin{bmatrix} -E_{12} \\ E_{11} \end{bmatrix} \begin{bmatrix} -E_{12}^T & E_{11}^T \end{bmatrix} \end{aligned} \tag{16}$$

Consider now the rank-1 matrix

$$(a_1 + a_2) \begin{bmatrix} E_{11} \\ E_{12} \end{bmatrix} \begin{bmatrix} E_{11}^T & E_{12}^T \end{bmatrix} \tag{17}$$

Since W^{opt} given in (16) satisfies the constraints of Optimization 3 and also maximizes its objective function, it is easy to verify that the rank-1 matrix in (17) is also a solution of Optimization 3. In other words, Optimization 3 has a rank-1 solution, which must also be a solution of Optimization 4.

Hence, we have shown that condition C1 closes the gap between Optimizations 3 and 4. The optimal value of Optimization 2 is equal to that of Optimization 3 (Theorem 3(2)), which is equal to that of Optimization 4, and hence equal to that of Optimization 1 and OPF (Theorem 3(1) and Theorem 2(1)). This completes the proof of zero duality gap between OPF and Dual OPF.

2) One can solve the convex problem of Optimization 2 efficiently to find optimal values $\boldsymbol{\lambda}_k^{\text{opt}}, \bar{\boldsymbol{\lambda}}_k^{\text{opt}}, \boldsymbol{\mu}_k^{\text{opt}}$, $k = 1, 2, ..., n$. ¿From the expression (13) for A^{opt}, it can be easily verify that if $\begin{bmatrix} U_1^T & U_2^T \end{bmatrix}^T$ is an eigenvector associated with any eigenvalue, then $\begin{bmatrix} U_2^T & -U_1^T \end{bmatrix}^T$ is another (orthogonal) eigenvector associated with the same eigenvalue. Hence, A^{opt} either has no zero eigenvalue or its zero eigenvalue has multiplicity 2. Theorem 3(1) implies that if A^{opt} has no zero eigenvalue, then $W^{\text{opt}} = 0$ and $V^{\text{opt}} = 0$, contradicting condition C0(i). Thus, A^{opt} has a zero eigenvalue with multiplicity 2. Hence, there exist two orthogonal vectors $\begin{bmatrix} U_1^T & U_2^T \end{bmatrix}^T$ and $\begin{bmatrix} U_2^T & -U_1^T \end{bmatrix}^T$ in the null space of A^{opt} and two scalars ζ_1 and ζ_2 such that

$$\begin{bmatrix} \text{Re}\{V^{\text{opt}}\} \\ \text{Im}\{V^{\text{opt}}\} \end{bmatrix} = \zeta_1 \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} + \zeta_2 \begin{bmatrix} -U_2 \\ U_1 \end{bmatrix}$$

or equivalently

$$V^{\text{opt}} = (\zeta_1 + \zeta_2 \text{i})(U_1 + U_2 \text{i})$$

This completes the proof of Theorem 1. ■

IV. Discussion: Condition C1 for Zero Duality Gap

As we elaborate in Section VI, all the five IEEE benchmark systems archived at [22] can be solved exactly and quickly

[1]For any feasible solution (x, r) of Optimization 2, we will often write $\mathbf{\Lambda}$, $\bar{\mathbf{\Lambda}}$, $\mathbf{\Gamma}$ instead of $\mathbf{\Lambda}(x, r)$, $\bar{\mathbf{\Lambda}}(x, r)$, $\mathbf{\Gamma}(x, r)$ when the underlying variable (x, r) of Optimization 2 is understood.

following the Algorithm prescribed in Section II. Moreover all of them satisfy condition C1 after a small resistance (10^{-5}) has been added to each transformer that originally has zero resistance. This suggests hope that practical systems are likely to satisfy condition C1 and hence solvable efficiently. In this section, we discuss why this might be the case. We exhibit special power system scenarios in which condition C1 can be proved to hold and explain why we believe the optimal solutions of general scenarios are likely to be close to the optimal of the special scenarios and hence, by continuity, also have zero duality gap.

A graph is called *strongly connected* if there is a path between any two nodes. Consider the graph induced by the matrix $\text{Re}\{Y\}$.[2] We will need the following condition:

C2: The graph induced by $\text{Re}\{Y\}$ (the resistive part of the power system) is strongly connected.

Condition C2 can be checked by examining its $n\times n$ Laplacian matrix L where $L_{ij} = -1$ if $\text{Re}\{Y\}_{ij} \neq 0$ and $L_{ii} = -\sum_{j\neq i} L_{ij}$: C2 holds if and only if L has a zero eigenvalue of multiplicity 1 [32].

A. Algebraic structure

The next theorem exploits a result in algebraic graph theory to study condition C1.

Theorem 4: Suppose condition C2 holds. Consider a feasible point (x,r) of Optimization 2 so that the corresponding $\mathbf{\Lambda}$ and $\bar{\mathbf{\Lambda}}$ satisfy $\boldsymbol{\lambda}_i + \boldsymbol{\lambda}_j > 0$ and $\bar{\boldsymbol{\lambda}}_i + \bar{\boldsymbol{\lambda}}_j \geq 0$ whenever $Y_{ij} \neq 0$, $i \neq j$.

1) The smallest eigenvalue of the matrix $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ is simple (not repeated).
2) If $H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) = 0$ at optimality, then A^{opt} has a zero eigenvalue of multiplicity 2. Moreover, $(V^{\text{opt}})_k \neq 0$, $k = 1, \ldots, n$.

Proof:

1) For $i \neq j$ with $Y_{ij} \neq 0$

$$\begin{aligned}\big(H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})\big)_{ij} &= (\boldsymbol{\lambda}_i + \boldsymbol{\lambda}_j)\text{Re}\{Y_{ij}\} \\ &\quad - (\bar{\boldsymbol{\lambda}}_i + \bar{\boldsymbol{\lambda}}_j)\text{Im}\{Y_{ij}\}\end{aligned} \tag{18}$$

Recall from condition C0(ii) that $\text{Re}\{Y_{ij}\}$ are nonpositive and $\text{Im}\{Y_{ij}\}$ are nonnegative. Therefore, all off-diagonal entries of $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ are non-positive. Moreover, the condition in the theorem guarantees that $\big(H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})\big)_{ij} < 0$ if $Y_{ij} \neq 0$, and hence the graph induced by $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ is strongly connected (due to condition C2). The matrix $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ is a *generalized Laplacian* of some graph (see, e.g., [32, p. 296]). By Lemma 13.9.1 of [32, p. 297], its smallest eigenvalue is simple. Moreover, the corresponding eigenvector can be taken to have only positive entries.

2) If $H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) = 0$, A^{opt} in (13) is block diagonal with $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ on the diagonal. Since (x,r) is feasible, $A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ is positive semidefinite, and so is $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$. Hence, the eigenvalues of $A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ and $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ are all nonnegative. As argued in the proof of Theorem 1, feasibility of OPF (condition C0(i)) implies their smallest eigenvalue is zero. If U is an eigenvector of $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ associated with its zero eigenvalue, then $[U^T \;\; U^T]^T$ and $[U^T \;\; -U^T]^T$ are two orthogonal eigenvectors of A^{opt} associated with a zero eigenvalue. Since the zero eigenvalue of $H_1(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ is simple, that of A^{opt} must have multiplicity 2. Therefore, condition C1 is satisfied. Since U can be taken as strictly positive, Theorem 1(2) implies that every component of the optimal voltages V_k^{opt} are nonzero.

This completes the proof. ■

Hence, provided the resistive part of the power system is strongly connected, condition C1 holds and the duality gap of OPF is zero if the off diagonal entries of H_1 are nonpositive and $H_2 = 0$. Since eigenvalues are continuous in the entries of their matrix, we expect C1 to hold even if H_2 is nonzero but small enough (relative to H_1). We make two remarks on why this might be the case in practice, one on H_1 and the other on H_2.

First, we can interpret $\boldsymbol{\lambda}_k^{\text{opt}}$ and $\bar{\boldsymbol{\lambda}}_k^{\text{opt}}$, $k = 1, ..., n$, as the prices for active and reactive powers. We therefore expect $\boldsymbol{\lambda}_k^{\text{opt}}$ to be all positive and $\bar{\boldsymbol{\lambda}}_k^{\text{opt}}$, $\boldsymbol{\mu}_k^{\text{opt}}$ to be small. For example, consider a load bus $k \in \{m+1, ..., n\}$. That $\boldsymbol{\lambda}_k^{\text{opt}}$ being positive means that although the generators must provide at least P_k^{min} amount of active power at bus k, the optimal strategy is to provide exactly this minimum amount (complementary slackness). Even though $\bar{\boldsymbol{\lambda}}_k^{\text{opt}}$ might sometimes be negative, they are far smaller than $\boldsymbol{\lambda}_k^{\text{opt}}$ in all the examples we have tried (see Section VI). The off-diagonal entries of H_1 are likely negative if $\boldsymbol{\lambda}_k^{\text{opt}}$ are positive and dominant over $\bar{\boldsymbol{\lambda}}_k^{\text{opt}}$. Second,

$$\big(H_2(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})\big)_{ij} = (\bar{\boldsymbol{\lambda}}_j - \bar{\boldsymbol{\lambda}}_i)\text{Re}\{Y\} + (\boldsymbol{\lambda}_j - \boldsymbol{\lambda}_i)\text{Im}\{Y\}$$

Hence, we expect H_2 to be small relative to H_1; c.f. (18). This is indeed the case with the IEEE benchmark systems discussed in Section VI. Note that H_2 becomes zero when the power system is purely resistive.

B. Geometric structure

We now study condition C1 from an geometric perspective. Recall the definition of Dual OPF (3)–(4) with variables (x,r). To simplify, we assume that $f_l(P_l^g) = P_l^g$ for $l = 1, ..., m$, which specializes OPF to loss minimization. It can be shown that this implies $r_{l1}^{\text{opt}} = r_{l2}^{\text{opt}} = 0$, $l = 1, \ldots, m$, at optimality. Consider the remaining $6n$-dimensional variable x and let $\mathcal{D} \in \mathbf{R}^{6n}$ denote the set of x that makes $A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ in (13) positive semidefinite. We will write the $3n$-dimensional reduced vector $x_r := (\boldsymbol{\lambda}_k, \bar{\boldsymbol{\lambda}}_k, \boldsymbol{\mu}_k, k = 1, ..., n)$, without making explicit their dependence on $(x, r = 0)$. Define $\mathcal{D}_r \in \mathbf{R}^{3n}$ as the set of all vectors x_r that make $A(\mathbf{\Lambda}, \bar{\mathbf{\Lambda}}, \mathbf{\Gamma})$ positive semidefinite. With no loss of generality, assume also that the real part of Y has at least one eigenvalue at the origin meaning that the constant impedance-to-ground loads (if any) have no resistive parts (as satisfied for IEEE benchmark systems).

Optimization 2 minimizes a linear function $\rho^T x$, for some constant vector $\rho \in \mathbf{R}^{6n}$, over the convex feasible set $\mathcal{D}$. This finds the farthest point on the boundary of $\mathcal{D}$ (including infinity) in the direction of the negative gradient $-\boldsymbol{\rho}$. Denote

[2]Given an $n \times n$ symmetric matrix Q, a *graph induced by* Q is a graph that has n vertices labeled by $1, \ldots, n$ and an edge (i,j), $i \neq j$, if $Q_{ij} \neq 0$.

this farthest point by x^{opt}. Condition C1 given in Theorem 1 is closely related to the smoothness of the boundary of $\mathcal{D}_r$ around the point x_r^{opt}. Specifically, using results on the geometrical shape of the set of all semidefinite matrices [28] and the fact that every eigenvalue of the matrix A^{opt} is repeated twice, one can infer that:

- The boundary of $\mathcal{D}_r$ is composed of different minimal faces, where the zero eigenvalue of A^{opt} has a constant multiplicity at all points of every such face.
- The multiplicity of the zero eigenvalues of A^{opt} over each face is a positive multiple of 2.
- The boundary of $\mathcal{D}_r$ at a point x_r^{opt} is smooth (differentiable) if and only if x_r^{opt} belongs to the face over which the multiplicity of the zero eigenvalue of A^{opt} is 2.

Therefore, condition C1 means that *the optimal point x_r^{opt} is finite and the boundary of $\mathcal{D}_r$ is smooth at this point.* To appreciate why this condition is likely to hold, we first exhibit such a point.

Theorem 5: Suppose the cost functions of OPF are $f_l(P_l^g) = P_l^g$, $l = 1, \dots, m$. Suppose condition C2 holds. If the active power losses in the transmission lines are zero at optimality, then the following point

$$\boldsymbol{\lambda}_k^{\text{opt}} = 1, \quad \bar{\boldsymbol{\lambda}}_k^{\text{opt}} = 0, \quad \boldsymbol{\mu}_k^{\text{opt}} = 0, \quad k = 1, ..., n$$

is an optimal solution of Optimization 2 and satisfies condition C1.

Proof: Consider the point (x, r) defined by:

$$\lambda_k^{\min} = \bar{\lambda}_k^{\min} = \bar{\lambda}_k^{\max} = \mu_k^{\min} = \mu_k^{\max} = r_{l1} = r_{l2} = 0,$$

$$\lambda_k^{\max} := \begin{cases} 0 & \text{if } k = 1, ..., m \\ 1 & \quad \text{otherwise} \end{cases}$$

where $k = 1, ..., n$ and $l = 1, ..., m$. It is straightforward to verify that the given point (x, r) is feasible and $h(x, r) = \sum_{k=1}^{n} P_k^d$. Moreover, the corresponding x_r is given by:

$$\boldsymbol{\lambda}_k = 1, \quad \bar{\boldsymbol{\lambda}}_k = 0, \quad \boldsymbol{\mu}_k = 0, \quad k = 1, ..., n$$

On the other hand

$$\sum_{k=1}^{n} \left(1 \times \mathbf{Y}_k + 0 \times \bar{\mathbf{Y}}_k + 0 \times M_k\right) = \begin{bmatrix} \text{Re}\{Y\} & 0 \\ 0 & \text{Re}\{Y\} \end{bmatrix}$$

which has a zero eigenvalue of multiplicity 2 (due to Condition C2). Therefore, x_r is a differentiable point lying on the boundary of $\mathcal{D}_r$ and hence satisfies condition C1.

Since OPF is feasible and the total power loss is zero, the total generation must be equal to the total demand. Hence, it follows from the weak duality theorem that

$$h(x, r) \;\leq\; \sum_{l=1}^{m} P_l^g \;=\; \sum_{k=1}^{n} P_k^d$$

for every feasible point (x, r) of Optimization 2. Then $h(x, r) = \sum_{k=1}^{n} P_k^d$ means that $(x, r) = (x^{\text{opt}}, r^{\text{opt}})$ is indeed a maximizer of Optimization 2. ∎

Theorem 5 says that condition C1 holds when the objective is to minimize total generated power and active power loss is zero at optimality. In that case, the theorem exhibits explicitly an optimal point that lies on a smooth face of the feasible set. The assumption on the objective can be relaxed to allow more general cost functions; see the IEEE 118-bus system discussed in Section VI-A for an example. Active power loss is, however, nonzero in practice. In that case, if the Lagrange multipliers $\boldsymbol{\lambda}_k^{\text{opt}}$, $\bar{\boldsymbol{\lambda}}_k^{\text{opt}}$ and $\boldsymbol{\mu}_k^{\text{opt}}$ are treated as prices for active and reactive powers as well as voltage levels, then the optimal point is likely to be in the vicinity of $(1, 0, 0)$, and hence, by continuity, also a differentiable point on the boundary of $\mathcal{D}_r$. As we will see in Section VI, this is indeed the case for the IEEE benchmark systems.

C. Summary

Summarizing the ideas in this section, condition C1 is mainly the consequence of two properties of power systems: (i) the particular "sign" structure of the Y matrix as described in condition C0(ii) (due to the positivity of the physical quantities, namely resistance, capacitance and inductance), (ii) the positivity of the Lagrange multipliers $\boldsymbol{\lambda}^{\text{opt}}$ representing the cost of active power generation

We showed in a recent paper [20] that an antenna design problem can be cast as an optimization problem with the same structure as Optimization 1, which was proven to be NP-complete. Hence, the duality gap of Optimization 1 is in general nonzero. However, due to the special structure of the admittance matrix Y of a power system, a subset of Optimization 1 has zero duality gap and is hence efficiently solvable.

V. Discussion: Generalizations

We have shown that the solution of the highly nonconvex OPF problem can be found by solving its convex dual problem when condition C1 holds. Most of the existing algorithms for solving the OPF problem can also be adapted to solving Dual OPF. For instance, powerful primal-dual algorithms can be deployed to solve the "dual" and the "dual of the dual" of the OPF problem, i.e. Optimizations 2 and 3, iteratively. This contrasts with the common technique of using a primal-dual algorithm to solve OPF and its dual concurrently. Note that although the number of variables in Optimization 2 is linear in n, Optimization 3 has a matrix variable that makes the number of its scalar variables quadratic in n. In other words, Optimization 3 might be costly to solve directly for very large values of n, in which case it is recommended to use some sub-gradient techniques [27].

Optimization 2 has the interesting property that the given loads and limits on power/voltage variables ($P_k^d, P_k^d, P_l^{\min}, P_l^{\max}, Q_l^{\min}, Q_l^{\max}, V_k^{\min}, V_k^{\max}$ for $k = 1, 2, ..., n$ and $l = 1, 2, ..., m$) only appear in the objective function, whereas the network topology (the matrix Y) appears only in its linear matrix constraint. Therefore, there is a natural decomposition between the load profile and the network topology in Optimization 2. This useful property, besides the linearity of Optimization 2, makes it possible to solve many more sophisticated problems efficiently, such as solving the OPF problem in the case when the load is stochastic and time-varying, designing the best network topology (using certain switches) to minimize power loss, etc.

The classical OPF problem formulated in (1) and (2) encompasses the basic and most important constraints needed to find an operating point of a power network. Nonetheless, more constraints are imposed in practice, e.g. stability limits, thermal limits and line flow constraints. The techniques developed in this paper can be used to prove that the incorporation of such constraints into Optimization 2 does not create a nonzero duality gap.

As the simplest example, assume that there are extra conditions $|V_i - V_j| < V_{ij}^{\max}$ in the OPF problem, for some given limits $V_{ij}^{\max}$ and indices $i,j = 1,\dots,n$ to ensure that the voltages at buses i and j are sufficiently close to each other. Define the matrix M_{ij} as

$$\begin{bmatrix} e_ie_i^* - e_ie_j^* - e_je_i^* + e_je_j^* & 0 \\ 0 & e_ie_i^* - e_ie_j^* - e_je_i^* + e_je_j^* \end{bmatrix}$$

By considering a scalar nonnegative variable μ_{ij} (Lagrange multiplier) associated with the constraint $|V_i - V_j|^2 < (V_{ij}^{\max})^2$, the only modifications needed in Optimization 2 are (i) to add the linear term $-\mu_{ij}(V_{ij}^{\max})^2$ to the objective function, and (ii) to add the matrix term $\mu_{ij}M_{ij}$ to the left side of the matrix constraint (4a). The duality gap is still expected to be zero (under appropriate conditions) because the parameter μ_{ij}^{opt} is nonnegative and the off-diagonal entries of the matrix M_{ij} are non-positive. Similarly, one can incorporate any constraints on the magnitudes of line currents into Optimization 2.

VI. Power System Examples

This section illustrates our results through two examples. Example 1 uses the IEEE benchmark systems archived at [22] to show the practicality of our result. Since the systems analyzed in Example 1 are so large that the specific values of the optimal solution cannot be provided in the paper, some smaller examples are analyzed in Example 2 with more details.

There are two main findings from this exercise. First, the duality gap is zero for all the systems we have tried, even when the sufficient condition C1 is not satisfied. We verify this by following the Algorithm in Section II-B to solve Dual OPF and compute the voltages. In all cases, the voltages obtained are feasible for Optimization 1 and achieve a primal objective value that is equal to the optimal objective value of Optimization 2. By weak duality theorem, the duality gap is zero and the voltages are optimal for OPF. Second, condition C1 is essentially satisfied: when it is violated, the violation is due to the simplifying modeling assumption that transformers have zero resistance. If a small resistance (10^{-5} per unit) is added to each of these transformers, condition C1 is satisfied for all IEEE benchmark systems.

The results of this section are attained using the following software tools:

- The MATLAB-based toolbox "YALMIP" (together with the solver "SEDUMI") is used to solve the dual of the OPF problem (Optimization 2), which is in the form of a linear-matrix-inequality optimization problem [29].
- The software toolbox "MATPOWER" is used to solve the OPF problem in Example 1 for the sake of comparison. The data for the IEEE benchmark systems analyzed in this example is extracted from the library of this toolbox [30].
- The software toolbox "PSAT" is used to draw and analyze the power networks given in Example 2 [31].

A. Example 1: IEEE benchmark systems

We have solved all IEEE systems with 14, 30, 57, 118 and 300 buses using the method developed in this paper, where the goal is to minimize either the total generation cost or the power loss. However, due to space restrictions, the details will be provided here only for two cases: (i) the loss minimization for the IEEE 30-bus system, and (ii) the total generation cost minimization for the IEEE 118-bus system.

1) IEEE 30-bus system: First, consider the OPF problem for the IEEE 30-bus system, where the objective is to minimize the total power generated by the generators. When the original Optimization 2 is solved, the four smallest eigenvalues of the matrix

$$A^{\text{opt}} = \begin{bmatrix} H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) & H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) \\ -H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) & H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) \end{bmatrix}$$

would be obtained as $0, 0, 0, 0$. Since the number of zero eigenvalues is 4, condition C1 required in Theorem 1 is violated. To explore the underlying reason, consider the circuit of this power system that is depicted in Figure 2. The circuit is composed of three regions connected to each other via some transformers. This implies that if each line of the circuit is replaced by its resistive part, the resulting resistive graph will not be connected (since the lines with transformers are assumed to have no resistive parts). Thus, the graph induced by $\text{Re}\{Y\}$ is not strongly connected and the zero eigenvalue of $\text{Re}\{Y\}$ has multiplicity larger than 1, violating condition C2. This is an issue with all the IEEE benchmark systems. This can be easily fixed by adding a little resistance to each transformer, say on the order of 10^{-5} (per unit). After this modification to the real part of Y, the four smallest eigenvalues of the matrix A^{opt} turn out to be $0, 0, 0.0075, 0.0075$; i.e. the zero eigenvalues resulting from the non-connectivity of the resistive graph have disappeared. Condition C1 is satisfied and the corresponding vector of optimal voltages can be recovered using the algorithm described after Theorem 1.

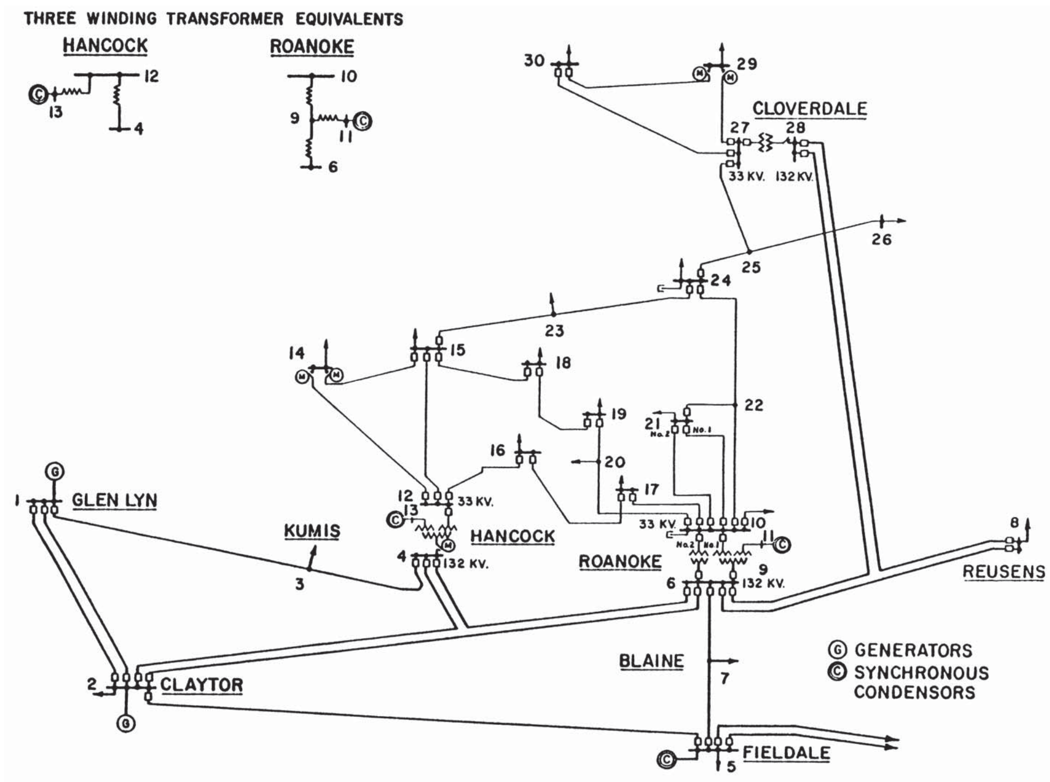
To illustrate the discussion in Section IV, we note that, for $k = 1, \dots, n$,

$$\boldsymbol{\lambda}_k \in [1, 1.0426], \quad \bar{\boldsymbol{\lambda}}_k \in [0, 0.0152], \quad \boldsymbol{\mu}_k \in [0, 0.0098],$$

Hence

- $\boldsymbol{\lambda}_k$'s are all positive and around 1.
- $\bar{\boldsymbol{\lambda}}_k$'s are all positive and around 0.
- $\boldsymbol{\mu}_k$'s are all very close to 0.

confirming the properties discussed in Section IV-B. Moreover, the maximum absolute values of the entries of $H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}})$ is 0.0867, whereas the average absolute values of the nonzero entries of $H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}})$ is 4.1201. This confirms the claim in Section IV-A that the matrix H_2 is expected to be negligible compared to H_1.



(a)

Fig. 2. The circuit of the IEEE 30-bus system taken from [22].

2) IEEE 118-bus system: Consider now the problem of minimizing the total generation cost for the IEEE 118-bus system. After adding some small resistance to certain entries of Re$\{Y\}$ to make the induced graph strongly connected, the four smallest eigenvalues of the matrix

$$\begin{bmatrix} H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) & H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) \\ -H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) & H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}}) \end{bmatrix}$$

are $0, 0, 1.3552, 1.3552$. Hence, condition C1 is satisfied and OPF can be solved by solving Dual OPF. Since the cost function f_l is not total generated power, the condition discussed in Section IV-B $(\boldsymbol{\lambda}_k, \bar{\boldsymbol{\lambda}}_k, \boldsymbol{\mu}_k) \sim (1, 0, 0)$ needs to be modified: the optimal variables normalized by $c_{l1} = 40$ satisfy, for $k = 1, \ldots, n$,

$$\frac{\boldsymbol{\lambda}_k}{c_{l1}} \in [0.8858, 1.0356], \quad \frac{\bar{\boldsymbol{\lambda}}_k}{c_{l1}} \in [-0.0063, 0.0118], \\ \frac{\boldsymbol{\mu}_k}{c_{l1}} \in [0, 0.1894]$$

As before, $(\frac{\boldsymbol{\lambda}_k}{c_{l1}}, \frac{\bar{\boldsymbol{\lambda}}_k}{c_{l1}}, \frac{\boldsymbol{\mu}_k}{c_{l1}})$ are around (1, 0, 0). In addition, $\boldsymbol{\lambda}_k$'s are all positive and most of $\bar{\boldsymbol{\lambda}}_k$ are positive (more than 100 of them). As the last property, the maximum of the absolute values of the entries of $H_2(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}})$ is 13.8613, whereas the average of the absolute values of the nonzero entries of $H_1(\mathbf{\Lambda}^{\text{opt}}, \bar{\mathbf{\Lambda}}^{\text{opt}}, \mathbf{\Gamma}^{\text{opt}})$ is 237.3938. Thus, H_2 is negligible compared to H_1 as before.

The computation on the IEEE benchmark examples were all finished in a few seconds and the number of iterations for each example was between 5 and 20. Note that although Optimization 2 is convex and there is no convergence problem regardless of what initial point is used, the number of iterations needed to converge mainly depends on the choice of starting point. It is worth mentioning that when different algorithms implemented in Matpower were applied to these systems, some of the constraints are violated at the optimal point probably due to the large-scale and non-convex nature of the OPF problem. However, no constraint violation have occurred by solving the dual of the OPF problem due to its convexity.

B. Example 2: small systems

The IEEE test systems in the previous example operate in a normal condition when the optimal bus voltages are close to each other both in magnitude and phase. This example illustrates that condition C1 is satisfied even in the absence of such a normal operation. Consider three distributed power systems, referred to as Systems 1, 2 and 3, depicted in Figure 3. Note that Systems 2 and 3 are radial, while System 1 has a loop. The detailed specifications of these systems are provided in Table I in per unit for the voltage rating 400kV and the power rating 100MVA, in which $\bar{z}_{ij}$ and $\bar{y}_{ij}$ denote the series impedance and the shunt capacitance of the Π model of the transmission line connecting buses $i, j \in \{1, 2, 3, 4\}$. The goal is to minimize the active power injected at slack bus 1 while satisfying the constraints given in Table II.

Optimization 2 is solved for each of these systems, and it is observed that condition C1 always holds. The optimal solution

TABLE I

PARAMETERS OF THE SYSTEMS GIVEN IN FIGURE 3.

Parameters	System 1	System 2	System 3
$\bar{z}_{12}$	$0.05+0.25\mathrm{i}$	$0.1+0.5\mathrm{i}$	$0.10+0.1\mathrm{i}$
$\bar{z}_{13}$	$0.04+0.40\mathrm{i}$	None	None
$\bar{z}_{23}$	$0.02+0.10\mathrm{i}$	$0.02+0.20\mathrm{i}$	$0.01+0.1\mathrm{i}$
$\bar{z}_{14}$	None	None	$0.01+0.2\mathrm{i}$
$\bar{y}_{12}$	0.06i	0.02i	0.06i
$\bar{y}_{13}$	0.05i	None	None
$\bar{y}_{23}$	0.02i	0.02i	0.02i
$\bar{y}_{14}$	None	None	0.02i

TABLE II

CONSTRAINTS TO BE SATISFIED FOR THE SYSTEMS GIVEN IN FIGURE 3.

Constraints	System 1	System 2	System 3
$P_2^d+Q_2^d\mathrm{i}$	$0.95+0.4\mathrm{i}$	$0.7+0.02\mathrm{i}$	$0.9+0.02\mathrm{i}$
$P_3^d+Q_3^d\mathrm{i}$	$0.9+0.6\mathrm{i}$	$0.65+0.02\mathrm{i}$	$0.6+0.02\mathrm{i}$
$P_4^d+Q_4^d\mathrm{i}$	None	None	$0.9+0.02\mathrm{i}$
$V_1^{\max}$	1.05	1.4	1

of OPF recovered from the solution of Optimization 2 are provided in Table III (P_{loss} and Q_{loss} in the table represent the total active and reactive power losses, respectively). It is interesting to note that although different buses have very disparate voltage magnitudes and phases, the duality gap is still zero. The optimal solution of Optimization 2 is summarized in Table IV to demonstrate that the Lagrange multipliers corresponding to active and reactive power constraints are positive.

As another scenario, let the desired voltage magnitude at the slack bus of System 1 be changed from 1.05 to 1. It can be verified that the optimal value of Optimization 2 becomes $+\infty$, which simply implies that the corresponding OPF problem is infeasible.

TABLE III

PARAMETERS OF THE OPF PROBLEM RECOVERED FROM THE SOLUTION OF OPTIMIZATION 2.

Recovered Parameters	System 1	System 2	System 3
V_1	$1.05\angle 0°$	$1.4\angle 0°$	$1\angle 0°$
V_2	$0.71\angle{-20.11}°$	$1.10\angle{-25.73}°$	$0.78\angle{-10.58}°$
V_3	$0.68\angle{-21.94}°$	$1.08\angle{-31.96}°$	$0.76\angle{-16.31}°$
V_4	None	None	$0.95\angle{-10.82}°$
P_{loss}	0.2193	0.1588	0.3877
Q_{loss}	1.2944	0.7744	0.5343

TABLE IV

LAGRANGE MULTIPLIERS OBTAINED BY SOLVING OPTIMIZATION 2 FOR THE SYSTEMS GIVEN IN FIGURE 3.

Lagrange Multipliers	System 1	System 2	System 3
$\boldsymbol{\lambda}_2$	1.3809	1.4028	1.7176
$\boldsymbol{\lambda}_3$	1.4155	1.4917	1.7900
$\boldsymbol{\lambda}_4$	None	None	1.0207
$\bar{\boldsymbol{\lambda}}_2$	0.4391	0.2508	0.1764
$\bar{\boldsymbol{\lambda}}_3$	0.4955	0.2633	0.1858
$\bar{\boldsymbol{\lambda}}_4$	None	None	0.0061
$\boldsymbol{\mu}_1$	0.0005	0.0001	0.0005

We repeated several hundred times this example by randomly choosing the parameters of the systems given in Figure 3 over a wide range of values. In all these trials, the Algorithm prescribed in Section II always found a globally optimal solution of the OPF problem or detected its infeasibility.

VII. CONCLUSIONS

We study the optimal power flow (OPF) problem that has been studied for about half a century and is notorious for its high nonconvexity. We have derived the dual of OPF as a convex linear matrix inequality optimization which can be efficiently solved. We have provided a sufficient condition under which the duality gap is zero and a globally optimal solution of the OPF problem can be recovered from a dual optimal solution. This condition is satisfied for the IEEE benchmark systems with 14, 30, 57, 118 and 300 buses, after a small resistance (10^{-5} per unit) is added to every transformer that originally assumes zero resistance. We have provided an informal justification from algebraic and geometric perspectives on why the condition might hold widely in practice. The main underlying reason for zero duality gap is that physical quantities, such as resistance, capacitance and inductance, are all positive.

ACKNOWLEDGMENT

The authors would like to gratefully acknowledge John C. Doyle for fruitful discussions on this topic.

APPENDIX

A. LMI Optimization Problems

The area of convex optimization has seen remarkable progress in the past two decades, particularly in linear matrix inequalities (LMIs) and semidefinite programming where the goal is to minimize a linear function subject to some linear matrix inequalities [21], [23]. The book [24] describes several difficult control problems that can be cast as LMI problems and then solved efficiently. The recent advances in this field have been successfully applied to different problems in other areas, e.g. circuit and communications [25], [26]. A powerful property in semidefinite programming is that the dual of an LMI optimization problem is again an LMI problem and, moreover, strong duality often holds [23].

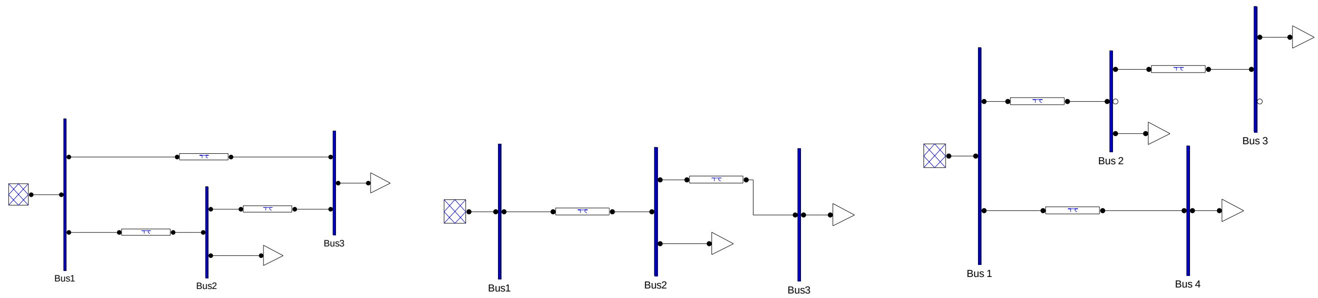
Given the scalar variables $x_1, ..., x_n$, consider the problem of minimizing

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \tag{19}$$

subject to the constraint

$$A_0 + A_1x_1 + \cdots + A_nx_n \preceq 0 \tag{20}$$

where $a_1, ..., a_n$ are given real numbers and $A_0, ..., A_n$ are given symmetric matrices in $\mathbf{R}^{n_0\times n_0}$, for some natural number n_0. Notice that the objective of the above optimization problem is a linear scalar function, and its constraint is a linear matrix inequality. Therefore, the above optimization problem is referred to as *an LMI problem*, which belongs to the category of convex optimization problems that can be solved



(a) (b) (c)

Fig. 3. Figures (a), (b) and (c) depict Systems 1, 2 and 3 studied in Example 2, respectively.

efficiently. To write the Lagrangian for the above optimization problem, a Lagrange multiplier should be introduced for the inequality (20). In light of the generalized Lagrangian theory, the multiplier associated with the inequality (20) is a symmetric matrix W in $\mathbf{R}^{n_0\times n_0}$ that must be positive semidefinite. The corresponding Lagrangian will be as follows:

$$\sum_{k=1}^{n} a_k x_k + \text{trace}\left\{ W\left(A_0 + \sum_{k=1}^{n} A_k x_k \right)\right\}$$

Note that the trace operator performs the multiplication between the expression in the constraint (20) and its associated Lagrange multiplier. Minimizing the above Lagrangian over $x_1, ..., x_n$ and then maximizing the resulting term over $W \succeq 0$ lead to the optimization problem of maximizing

$$\text{trace}\{WA_0\}$$

subject to the constraints

$$\text{trace}\{WA_k\} + a_k = 0, \quad k = 1, 2, ..., n$$

for a symmetric matrix variable $W \succeq 0$. This optimization problem is the dual of the initial optimization problem formulated in (19) and (20). If some mild conditions (such as Slater's conditions) hold, then the duality gap between the solutions of these two optimization problems becomes zero, meaning that the optimal objective values obtained by these problems will be identical. In this case, it is said that "strong duality" holds; otherwise, only "weak duality" holds in which case the optimal value of the dual problem is only a lower bound on the optimal value of the original problem. One can refer to [21] and [23] for detailed discussions on LMI problems.

B. Proofs

We prove in this subsection Theorems 2 and 3 that are summarized pictorially in Figure 1.

Proof of Theorem 2:

1) As discussed before the definition of Optimization 1, constraints (8a) and (8b) are derived by eliminating dependent variables P^g and Q^g using (2d), (2e) and the convention $P_k^{\min} = P_k^{\max} = Q_k^{\min} = Q_k^{\max} = 0$ for $k \in \{m+1, ..., n\}$. To show that the objective function (1) in OPF is tantamount to the objective function (7) with the extra constraint (8d) in Optimization 1, observe that minimizing $\sum_{l=1}^{m} f_l(P_l^g)$ is the same as minimizing $\sum_{l=1}^{m} \alpha_l$ subject to the constraint

$$c_{l2}(P_l^g)^2 + c_{l1}P_l^g + c_{l0} \leq \alpha_l, \quad l = 1, 2, ..., m$$

Using Schur complement, this constraint is equivalent to

$$\begin{bmatrix} c_{l0} + c_{l1}P_l^g - \alpha_l & \sqrt{c_{l2}}P_l^g \\ \sqrt{c_{l2}}P_l^g & -1 \end{bmatrix} \preceq 0$$

Substituting $P_l^g = \text{Re}\{Y_l I_l^*\} + P_l^d$ yields the condition (8d).

2) For $k = 1, \ldots, n$, let $\lambda_k^{\min}, \bar{\lambda}_k^{\min}, \mu_k^{\min}$ denote the respective Lagrange multipliers associated with the lower inequalities in (8a), (8b), (8c); similarly, let $\lambda_k^{\max}, \bar{\lambda}_k^{\max}, \mu_k^{\max}$ denote the Lagrange multipliers for the upper inequalities in (8a), (8b), (8c). These Lagrange multipliers must all be nonnegative. Introduce a (symmetric) positive semidefinite matrix

$$\begin{bmatrix} r_{l0} & r_{l1} \\ r_{l1} & r_{l2} \end{bmatrix}$$

as the Lagrange multiplier for the inequality (8d). Then the Lagrangian corresponding to Optimization 1 is (after some simplifications)

$$\begin{aligned} &\sum_{k=1}^{m} \left(\boldsymbol{\lambda}_k \text{Re}\{V_k I_k^*\} + \bar{\boldsymbol{\lambda}}_k \text{Im}\{V_k I_k^*\} + \boldsymbol{\mu}_k |V_k|^2 \right) \\ &\qquad\qquad + h(x,r) + \sum_{l=1}^{m}(1 - r_{l0})\alpha_l \end{aligned} \tag{21}$$

Substituting (9), (11) and (12) into (21), the Lagrangian can be written as

$$\begin{aligned} &\text{trace}\left\{ \sum_{k=1}^{m} \left(\hat{\boldsymbol{\lambda}}_k \mathbf{Y}_k + \bar{\boldsymbol{\lambda}}_k \bar{\mathbf{Y}}_k + \boldsymbol{\mu}_k M_k \right) UU^T \right\} \\ &\qquad\qquad + h(x,r) + \sum_{l=1}^{m}(1 - r_{l0})\alpha_l \end{aligned}$$

To obtain the dual of Optimization 1, the Lagrangian should first be minimized over U, α and then maximized over the Lagrange multipliers. Observe that

- The minimum of $(1 - r_{l0})\,\alpha_l$ over the variable α_l is $-\infty$ unless $r_{l0} = 1$, in which case the minimum is zero.

- The minimum of the term
$$\text{trace}\left\{\sum_{k=1}^{m}\left(\hat{\boldsymbol{\lambda}}_k\text{Re}\mathbf{Y}_k+\bar{\boldsymbol{\lambda}}_k\text{Im}\bar{\mathbf{Y}}_k+\boldsymbol{\mu}_k M_k\right)UU^T\right\}$$
over U is $-\infty$ unless
$$\sum_{k=1}^{m}\left(\hat{\boldsymbol{\lambda}}_k\text{Re}\mathbf{Y}_k+\bar{\boldsymbol{\lambda}}_k\text{Im}\bar{\mathbf{Y}}_k+\boldsymbol{\mu}_k M_k\right)\succeq 0$$
in which case the minimum is zero.

Hence, we must have $r_{l0}=1$. The dual objective function is therefore as stated.

This completes the proof of Theorem 2. ■

Proof of Theorem 3:

1) Given a feasible vector V of Optimization 1, define W as UU^T, where U is expressed in terms of V via (10). The matrix W defined this way is positive semidefinite and has rank at most 1. On the other hand, one can use singular value decomposition to show that every positive semidefinite matrix W with rank at most 1 can be decomposed as UU^T for some vector U. Hence, this change of variable is a bijective map (up to the sign of U). To prove the equivalence between Optimizations 1 and 4, it suffices to show that the constraints of these optimization problems will be mapped to each other using this change of variable. But this follows directly from (9), (11) and (12).
2) One can derive the dual of Optimization 2 following the standard procedure to obtain Optimization 3, noting that W in Optimization 3 plays the role of Lagrange multiplier for the matrix constraint (4a) in Optimization 2; see Appendix A and, e.g., [21], [24]. The details are omitted for brevity. Since Optimizations 2 and 3 are both semidefinite programs and hence convex, strong duality holds if Optimization 2 has a finite optimal value and a strictly feasible point (Slater condition). Since OPF is feasible and equivalent to Optimization 1, Optimization 1 has a finite optimal value. Optimization 2 is its dual by Theorem 2, and is therefore upper bounded by the finite optimal value of Optimization 1 (weak duality). To show that Optimization 2 has a strictly feasible point, consider the point (x,r) specified by: for $k=1,\ldots,n$,
$$\begin{aligned}&\lambda_k^{\min}:=\begin{cases}c_{k1}+1 & \text{if } k=1,...,m\\ 1 & \text{otherwise}\end{cases},\\ &\lambda_k^{\max}=1,\quad \bar{\lambda}_k^{\min}=\bar{\lambda}_k^{\max}=1,\\ &\mu_k^{\min}=1,\quad \mu_k^{\max}=2,\quad r_{l1}=0,\quad r_{l2}=1\end{aligned}\tag{22}$$
Then $\boldsymbol{\lambda}_k=\bar{\boldsymbol{\lambda}}_k=0$ and $\boldsymbol{\mu}_k=1$. Now, observe that
   - The variable x specified in (22) is strictly positive componentwise.
   - The relations
$$\sum_{k=1}^{n}\left(\boldsymbol{\lambda}_k\mathbf{Y}_k+\bar{\boldsymbol{\lambda}}_k\bar{\mathbf{Y}}_k+\boldsymbol{\mu}_k M_k\right)=I\succ 0$$
and
$$\begin{bmatrix}1 & r_{l1}\\ r_{l11} & r_{l2}\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\succ 0$$
both hold.

   Hence (x,r) given in (22) is strictly feasible and strong duality holds.

This completes the proof of Theorem 3. ■

References

[1] J. A. Momoh, "Electric power system applications of optimization," *Markel Dekker*, New York, USA, 2001.
[2] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin Society Francaise Electriciens*, vol. 3, no. 8, pp. 431-447, 1962.
[3] M. Huneault and F. D. Galiana, "A survey of the optimal power flow literature," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 762-770, 1991.
[4] G. L. Torres and V. H. Quintana, "Optimal power flow by a nonlinear complementarity method," *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp. 1028-1033, 2000.
[5] H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185-1193, 2007.
[6] J. A. Momoh M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. Part I: Nonlinear and quadratic programming approaches," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 96-104, 1999.
[7] J. A. Momoh M. E. El-Hawary, and R. Adapa, "A review of selected optimal power flow literature to 1993. Part II: Newton, linear programming and interior point methods," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 105-111, 1999.
[8] K. S. Pandya and S. K. Joshi, "A survey of optimal power flow methods," *Journal of Theoretical and Applied Information Technology*, vol. 4, no. 5, pp. 450-458, 2008.
[9] H. Wei, H. Sasaki, J. Kubokawa, and R. Yokoyama, "An interior point nonlinear programming for optimal power flow problems with a novel data structure," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 870-877, 1998.
[10] W. M. Lin, C. H. Huang, and T. S. Zhan, "A hybrid current-power optimal power flow technique," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 177-185, 2008.
[11] Q. Y. Jiang, H. D. Chiang, C. X. Guo, and Y. J. Cao, "Power-current hybrid rectangular formulation for interior-point optimal power flow," *IET Generation, Transmission & Distribution*, vol. 3, no. 8, pp. 748-756, 2009.
[12] Q. Jiang, G. Geng, C. Guo, and Y. Cao, "An efficient implementation of automatic differentiation in interior point optimal power flow," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 147-155, 2010.
[13] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458-1459, 2006.
[14] R. A. Jabr, "Optimal power flow using an extended conic quadratic formulation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1000-1008, 2008.
[15] D. Gan, R. J. Thomas, and R. D. Zimmerman, "Stability-constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 535-540, 2000.
[16] H. R. Cai, C. Y. Chung, and K. P. Wong, "Application of differential evolution algorithm for transient stability constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 719-728, 2008.
[17] K. Xie and Y. H. Song, "Dynamic optimal power flow by interior point methods," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 148, no. 1, pp. 76-84, 2002.
[18] Y. Xia and K. W. Chan, "Dynamic constrained optimal power flow using semi-infinite programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1455-1457, 2006.
[19] P. M. Pardalos and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *Journal of Global Optimization*, vol. 1, no. 1, pp. 15-22, 1991.
[20] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle, "Solving Large-Scale Linear Circuit Problems via Convex Optimization," in *Proceedings of the 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009.

[21] S. Boyd and L. Vandenberghe, “Convex optimization,” *Cambridge University Press*, 2004.

[22] University of Washington, Power Systems Test Case Archive, http://www.ee.washington.edu/research/pstca.

[23] E. De Klerk, “Aspects of semidefinite programming: interior point algorithms and selected applications,” Applied Optimization Series, Vol. 65, *Kluwer Academic Publishers*, 2002.

[24] S. Boyd, LE Ghaoui, E. Feron and V. Balakrishnan, “Linear matrix inequalities in system and control theory,” *SIAM*, Philadelphia, PA, 1994.

[25] L. Vandenberghe, S. Boyd and A. E. Gamal, “Optimizing dominant time constant in RC circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 2, pp. 110-125, 1998.

[26] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle, “Programmable antenna design using convex optimization,” in *19th International Symposium on Mathematical Theory of Networks and Systems*, Budapest, Hungary, 2010.

[27] B. Recht, M. Fazel and P. A. Parrilo, “Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization,” to appear in *SIAM Review* (available online at http://pages.cs.wisc.edu/∼brecht/publications.html).

[28] H. Wolkowitz, R. Saigal, and L. Vandenberghe, editors, “Handbook of semidefinite programming: theory, algorithms and applications,” *Kluwer*, 2000.

[29] J. Löfberg, “A toolbox for modeling and optimization in MATLAB,” *in Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[30] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, “MATPOWER's extensible optimal power flow architecture,” *IEEE Power and Energy Society General Meeting*, 2009.

[31] F. Milano, “An open source power system analysis toolbox,” *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1199-1206, 2005.

[32] C. Godsil and G. Royle, “Algebraic graph theory,” *volume 207 of Graduate Texts in Mathematics. Springer*, 2001.