

# The Optimal Power Flow Operator: Theory and Computation

Fengyu Zhou, *Student Member, IEEE*, James Anderson, *Senior Member, IEEE*, and Steven H. Low, *Fellow, IEEE*

**Abstract**—Optimal power flow (OPF) problems are mathematical programs to determine how to distribute power over networks subject to power flow and operational constraints. In this paper, we treat an OPF problem as an operator that maps user demand to generated power, and allow the problem parameters to take values in some admissible set. We formalize this operator theoretic approach, define and characterize restricted parameter sets under which the mapping has a singleton output, independent binding constraints, and is differentiable. We show that for any power network, these analytical properties hold under almost all operating conditions and can thus be relied upon in applications. We further provide a closed-form expression for the Jacobian matrix of the OPF operator and describe how various derivatives can be computed using a recently proposed scheme based on homogenous self-dual embedding. In contrast to related work in the optimization literature, our results have a clear physical interpretation.

**Index Terms**—Analysis, optimal power flow, linear programming.

## I. INTRODUCTION

**G**IVEN a power network, the optimal power flow (OPF) problem seeks to find an operating point that minimizes an appropriate cost function subject to power flow and operational constraints such as Kirchhoff's laws, voltage and capacity limits, e.g. [1], [2], [3], [4], [5]. The decision variables in an OPF problem are typically voltages and generation power. Cost function choices include minimizing power loss, generation cost, and user disutility.

### A. Summary

We consider the direct current (DC) model [6], [7], [8] of the power flow equations. The DC-OPF problem is widely used in industry and takes the form of a linear program (LP). Consider a power network with  $N_G$  generators and  $N_L$  loads, the problem is formulated as

$$\underset{\mathbf{s}^g}{\text{minimize}} \quad \mathbf{f}^T \mathbf{s}^g \quad (1a)$$

$$\text{subject to} \quad \mathbf{A}_{\text{eq}} \mathbf{s}^g = \mathbf{b}_{\text{eq}}(\mathbf{s}^l, \mathbf{b}'), \quad (1b)$$

$$\mathbf{A}_{\text{in}} \mathbf{s}^g \leq \mathbf{b}_{\text{in}} \quad (1c)$$

This work was funded by NSF grants CCF 1637598, CPS 1739355, and ECCS 1619352, PNNL grant 424858, and through the ARPA-E GRID DATA program.

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).

James Anderson is now with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: james.anderson@columbia.edu).

where  $\mathbf{s}^g \in \mathbb{R}^{N_G}$  is the decision vector of power generations at each generator node in the network. The equality constraint function  $\mathbf{b}_{\text{eq}}$  is linear in  $\mathbf{s}^l$  (the vector of power demands at each node) and  $\mathbf{b}'$  (some constants). We view  $\mathbf{f}$ ,  $\mathbf{b}'$  and  $\mathbf{b}_{\text{in}}$  as problem parameters, whose values are allowed to change in a set  $\Omega$ . Matrices  $\mathbf{A}_{\text{eq}}$  and  $\mathbf{A}_{\text{in}}$  are determined by network topology and susceptances. We treat the power demand  $\mathbf{s}^l$ , which is allowed to take values in the set  $\Omega_{\text{sl}}$ , as an “input” and the optimal generations  $(\mathbf{s}^g)^*$  as an “output”. One of the main contributions of this work is to study the DC-OPF problem (1) as an operator, defined as:

$$\mathcal{OPF} : \Omega_{\text{sl}} \rightarrow 2^{\mathcal{S}}$$

where  $\mathcal{S} \subset \mathbb{R}^{N_G}$  and  $2^{\mathcal{S}}$  denotes the power set of  $\mathcal{S}$ .

We begin by deriving conditions such that:

- 1)  $\mathcal{OPF}$  maps to a singleton.
- 2)  $\mathcal{OPF}$  is continuous everywhere and differentiable almost everywhere.
- 3) All points in the image space of  $\mathcal{OPF}$  are optimal solutions to a DC-OPF problem with a fixed number of binding constraints (the number will be later derived as  $N_G - 1$ ).

Results such as those above have been shown to hold (with high probability) when the problem is simply viewed as a mathematical program (see for example; [9], [10], [11], [12]), of which the DC-OPF problem is a special case. However these results, while insightful in optimization theory convey little information about the process the optimization problem models. We note that Property 3 is equivalent to the concept of nondegeneracy [13]. In contrast to standard results, we narrate from the perspective of binding constraints. Such a perspective is beneficial because it provides insight into the physical meaning of OPF problem solutions. In particular, each binding constraint implies that either a generator or power line is idle or saturated. Furthermore, we use a sequence of restrictions to obtain subsets of “good” parameters (i.e., those which preserve the properties listed), thus our method is both constructive and interpretable.

With properties 1) – 3) established, we provide a closed-form expression for the Jacobian matrix of  $\mathcal{OPF}$ , that is, an approximate linear mapping from input (loads) to output (an optimal DC-OPF solution). We also show how one could equivalently view the Jacobian as a function of the set of binding constraints, i.e., a mapping which is independent of problem parameters. This new perspective presents advantages especially if we are more interested in the global behavior of the Jacobian map (such as the worst-case amplification of a

generator gain given a change in load), rather than just the generations given a specific load profile. Finally, we conclude by describing how various derivatives can be computed using recently developed ideas [14], [15].

From a power networks perspective there has been a lot of interest on the topic of characterizing the relationship between power demand and generation over the last few years, see for example [16], [17], [18], [19], [20], and the references therein. The purpose of this paper is to make these concepts rigorous and bridge optimization theory with practical power system operation and planning.

This work is a first step towards characterizing this complicated relationship. Specifically, establishing uniqueness of solution is a fundamental property of an operator as it provides the foundation for defining a derivative. Moreover, many numerical techniques require unique solutions to ensure convergence. Characterizing the set of independent binding constraints paves the road for many further topics of interest. For instance, in [21], it is shown that even under significant load variations, the number of binding line constraints in the DC-OPF problem is frequently a small proportion of the total number of constraints – an observation which has significant implications when it comes to long-term planning and assessing network vulnerability [22]. In [17], the set of binding constraints determines an area of load profiles (termed System Pattern Regions by the authors) within which the vector of locational marginal prices remains constant.

Sensitivity, more broadly is a useful term to quantify. In recent works we showed that “worst-case” sensitivity bounds provide privacy guarantees when releasing power flow data [16] and for data disaggregation [23]. In the context of real-time optimization where sensitivity is often assumed to be known and bounded [24], this work can be used to provide exactly these bounds (or rule them out).

## B. Related work

Our approach to sensitivity analysis differs from the standard perturbation approach which assumes the constraints are shifted from their nominal right-hand sides, and then looks at the Lagrange multipliers; see for example [25, Ch. 5.6]. In such a setting one considers the optimization problem

$$\{\text{minimize}_{\mathbf{x}} f_0(\mathbf{x}) : f_i(\mathbf{x}) \leq \mathbf{u}_i, h_j(\mathbf{x}) = \mathbf{v}_j, \forall i, j\}. \quad (2)$$

The *nominal* form of problem (2) has  $\mathbf{u}_i = 0$  and  $\mathbf{v}_j = 0$  for all  $i$  and  $j$ . Denote the nominal optimal value by  $\mathbf{x}^*$ .

Then we have the following well known lower-bound for the perturbed problem  $p^*(\mathbf{u}, \mathbf{v}) \geq p^*(0, 0) - \boldsymbol{\lambda}^{*\top} \mathbf{u} - \boldsymbol{\nu}^{*\top} \mathbf{v}$ , and the sensitivities are given by the relations  $\boldsymbol{\lambda}_i^* = -\partial_{\mathbf{u}_i} p^*(0, 0)$  and  $\boldsymbol{\nu}_j^* = -\partial_{\mathbf{v}_j} p^*(0, 0)$ . This standard perturbation approach is also applied to AC power systems to study the sensitivity and subdifferential of the optimal cost value [26]. This approach differs from our problem in the following ways. First, we focus on the perturbation of the optimal solution, rather than the optimal value. Second, we focus more on the set of binding constraints and whether they are independent, as opposed to appealing to duality.

A related body of work to compare our results to is that of robust optimization [27], [28], [29] and stochastic

optimization [30], [31], [32]. In both cases, the goal is to mitigate the effects of uncertainty. In contrast, our work seeks to determine how the optimal decision changes with respect to data perturbations. Our results can thus be considered complementary to the robust and stochastic optimization frameworks.

Within the optimization community there has been a lot of interest in sensitivity problems. Although not stated formally, some work has studied LPs as operators. We can trace these sensitivity-type results back to Hoffman [33] who showed that for vector  $\mathbf{x}$  which satisfies  $\mathbf{Ax} \leq \mathbf{b}$  and  $\mathbf{z}$  such that  $\mathbf{Az} \not\leq \mathbf{b}$ , then  $\text{dist}(\mathbf{x}, \mathbf{z})$  is bounded from above by  $\alpha \|(\mathbf{Az} - \mathbf{b})_+\|$ , where the constant  $\alpha$  depends on  $\mathbf{A}$ . The value of  $\alpha$  depends on the choices of norms and is typically difficult to compute. This line of work was generalized in [34] and extended in [35], where it was shown that perturbations to the cost vector can destroy Lipschitz continuity. A contribution of our work is to provide dense sets for which perturbations of the cost vector (and right-hand-side vectors) maintain differentiability.

With respect to uniqueness of the solution, it has been proved in [9], [10] that among all the linear program problem instances, almost all of those instances have unique (thus basic) and non-degenerate primal and dual solutions, and strict complementarity holds almost everywhere. In the context of OPF problems, our results are more specific. First, the traditional result shows nondegenerate instances are almost everywhere among all the problem instances, while our result says the good instances within a subset are also almost everywhere. Taking (1) as an example, traditional results show that for almost all the  $(\mathbf{f}, \mathbf{A}_{\text{eq}}, \mathbf{A}_{\text{in}}, \mathbf{b}_{\text{eq}}, \mathbf{b}_{\text{in}})$ , (1) maintains the three desirable properties above. In our work, we show that given any fixed  $(\mathbf{A}_{\text{eq}}, \mathbf{A}_{\text{in}})$  that describes a power network structure, for almost all  $(\mathbf{f}, \mathbf{b}', \mathbf{b}_{\text{in}})$ , all the good properties listed above hold for almost all instances of  $\mathbf{s}^l$ . In DC-OPF problems, many entries of  $\mathbf{b}_{\text{eq}}, \mathbf{b}_{\text{in}}$  are fixed or restricted due to their physical meanings. As a result, set of all the physical problem instance is of measure zero compared to the entire instance space of (1). Good instances being dense in the entire space in general does not imply they are still dense within a measure zero subset. Therefore, our results are not direct corollaries of the properties for general LP. It means that for any network, these analytical properties hold under almost all operating conditions and can thus be relied upon in applications. Second, traditional results are usually formulated in the canonical form in order to capture the general features of LPs. Our result, on the other hand, does not rely on introducing auxiliary slack variables and thus reveals how those properties link to the physical behavior of power systems.

There has been some work which specifically defines and studies the DC-OPF sensitivity. In [36], [37], the OPF problem is formulated as a parameterized optimization. Under heavy assumptions, sensitivity can be computed. In contrast, our methods are more general and offer structural insights. Furthermore, we provide numerical methods to compute the derivatives.

## C. Paper outline

In Section II, we formalize the DC-OPF problem, and characterize the parameter set of interest. In Section III, we

restrict the set of interest to those parameters that endow the OPF operator with desirable properties. We show that the restricted set is dense within the set of interest, so the restriction does not lose generality up to perturbation. In Section IV, we prove the operator is differentiable and derive the closed form expression of the Jacobian matrix in terms of the independent binding constraints. Section V shows how an algorithm introduced in recent work [14], [15] can help numerically evaluate the operator derivative. Finally, Section VI provides an illustrative example.

## II. BACKGROUND

Vectors and matrices are typically written in bold while scalars are not. For a set  $\mathcal{S}$ , we use  $\text{clos}(\mathcal{S})$  and  $\text{int}(\mathcal{S})$  to refer to its closure and interior, respectively. The power set of  $\mathcal{S}$  is the set of all subsets of  $\mathcal{S}$ . Given two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ ,  $\mathbf{a} \geq \mathbf{b}$  denotes the element-wise partial order  $a_i \geq b_i$  for  $i = 1, \dots, n$ . For a scalar  $k$ , we define the projection operator  $[k]^- := \min\{0, k\}$ . We define  $\|\mathbf{x}\|_0$  as the number of non-zero elements of the vector  $\mathbf{x}$ . Identity and zero matrices are denoted by  $\mathbf{I}^n$  and  $\mathbf{0}^{n \times m}$  while vectors of all ones are denoted by  $\mathbf{1}_n$  where superscripts and subscripts indicate their dimensions. To streamline notation, we omit the dimensions when the context makes it clear. The notation  $\mathbb{R}_+$  denotes the non-negative real set  $[0, +\infty)$ . For  $\mathbf{X} \in \mathbb{R}^{n \times m}$ , the restriction  $\mathbf{X}_{\{1,3,5\}}$  denotes the  $3 \times m$  matrix composed of stacking rows 1, 3, and 5 on top of each other. We will frequently use a set to describe the rows we wish to form the restriction from, in this case we assume the elements of the set are arranged in increasing order. We will use  $\mathbf{e}_m$  to denote the standard base for the  $m^{\text{th}}$  coordinate, and its dimension will be clear from the context. Let  $(\cdot)^\dagger$  be the Moore-Penrose inverse. Denote  $[m] := \{1, 2, \dots, m\}$  and  $[n, m] := \{n, n+1, \dots, m\}$ . Finally, for a convex set  $\mathcal{X} \subseteq \mathbb{R}^n$  and vector  $\mathbf{x} \in \mathbb{R}^n$ , we let  $\mathcal{P}_{\mathcal{X}}\mathbf{x}$  be the projection of  $\mathbf{x}$  onto the set  $\mathcal{X}$ . By isometry, the domain of the projection operator is extended to matrices when needed.

### A. System model

Consider a power network modeled by an undirected connected graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V} := \mathcal{V}_G \cup \mathcal{V}_L$  denotes the set of buses which can be further classified into subsets of generators  $\mathcal{V}_G$  and loads  $\mathcal{V}_L$ , and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  is the set of all branches linking those buses. Suppose  $\mathcal{V}_G \cap \mathcal{V}_L = \emptyset$  and there are  $|\mathcal{V}_G| =: N_G$  generator and  $|\mathcal{V}_L| =: N_L$  loads, respectively. For simplicity, let  $\mathcal{V}_G = [N_G]$ ,  $\mathcal{V}_L = [N_G + 1, N_G + N_L]$ . Let  $N = N_G + N_L$ . Without loss of generality,  $\mathcal{G}$  is a connected graph with  $|\mathcal{E}| =: E$  edges labelled as  $1, 2, \dots, E$ . Let  $\mathbf{C} \in \mathbb{R}^{N \times E}$  be the incidence matrix. We will use  $e$ ,  $(u, v)$  or  $(v, u)$  interchangeably to denote the same edge. Let  $\mathbf{B} = \text{diag}(b_1, b_2, \dots, b_E)$ , where  $b_e > 0$  is the susceptance of branch  $e$ . As we adopt a DC power flow model, all branches are assumed lossless. Further, we denote the generation and load as  $\mathbf{s}^g \in \mathbb{R}^{N_G}$ ,  $\mathbf{s}^l \in \mathbb{R}^{N_L}$ , respectively. Thus  $\mathbf{s}_i^g$  refers to the generation on bus  $i$  while  $\mathbf{s}_i^l$  refers to the load on bus  $N_G + i$ . We will refer to bus  $N_G + i$  simply as load  $i$  for simplicity. The power flow on branch  $e \in \mathcal{E}$  is denoted as  $\mathbf{p}_e$ , and  $\mathbf{p} := [\mathbf{p}_1, \dots, \mathbf{p}_E]^\top \in \mathbb{R}^E$  is the vector of all branch

power flows. To simplify analysis, we assume that there are no buses in the network that are both loads and generators.

### B. DC Optimal Power Flow

We focus on the DC-OPF problem with a linear cost function [6], [7], [8]. That is to say, the voltage magnitudes are assumed to be fixed and known and the lines are considered to be lossless. Without loss of generality, we assume all the voltage magnitudes to be 1. The decision variables are the voltage angles denoted by vector  $\boldsymbol{\theta} \in \mathbb{R}^N$  and power generations  $\mathbf{s}^g$ , given loads  $\mathbf{s}^l$ . The DC-OPF problem takes the form:

$$\underset{\mathbf{s}^g, \boldsymbol{\theta}}{\text{minimize}} \quad \mathbf{f}^\top \mathbf{s}^g \quad (3a)$$

$$\text{subject to} \quad \boldsymbol{\theta}_1 = 0 \quad (3b)$$

$$\mathbf{CBC}^\top \boldsymbol{\theta} = \begin{bmatrix} \mathbf{s}^g \\ -\mathbf{s}^l \end{bmatrix} \quad (3c)$$

$$\underline{\mathbf{s}}^g \leq \mathbf{s}^g \leq \bar{\mathbf{s}}^g \quad (3d)$$

$$\underline{\mathbf{p}} \leq \mathbf{BC}^\top \boldsymbol{\theta} \leq \bar{\mathbf{p}}. \quad (3e)$$

Here,  $\mathbf{f} \in \mathbb{R}_+^{N_G}$  is the unit cost for each generator, and bus 1 is selected as the slack bus with fixed voltage angle 0. In (3c), we define the injections for generators to be positive, while injections for loads are defined as  $-\mathbf{s}^l$ . The upper and lower limits on the generations are set as  $\bar{\mathbf{s}}^g$  and  $\underline{\mathbf{s}}^g$ , respectively, and  $\bar{\mathbf{p}}$  and  $\underline{\mathbf{p}}$  are the limits on branch power flows. We assume that (3) is well posed, i.e.  $\bar{\mathbf{s}}^g > \underline{\mathbf{s}}^g \geq 0$ ,  $\bar{\mathbf{p}} > \underline{\mathbf{p}}$ . Note that the LP (3) is a particular realization of (1).<sup>1</sup>

Let  $\boldsymbol{\tau} \in \mathbb{R}^{N+1}$  be the vector of Lagrangian multipliers associated with equality constraints (3b), (3c), and  $(\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_-)$  and  $(\boldsymbol{\mu}_+, \boldsymbol{\mu}_-)$  be the Lagrangian multipliers associated with inequalities (3d) and (3e) respectively. As (3) is a linear program, the following KKT condition holds at an optimal point when (3) is feasible:

$$(3b) - (3e) \quad (4a)$$

$$\mathbf{0} = \mathbf{M}^\top \boldsymbol{\tau} + \mathbf{CB}(\boldsymbol{\mu}_+ - \boldsymbol{\mu}_-) \quad (4b)$$

$$-\mathbf{f} = -[\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_{N_G}]^\top + \boldsymbol{\lambda}_+ - \boldsymbol{\lambda}_- \quad (4c)$$

$$\boldsymbol{\mu}_+, \boldsymbol{\mu}_-, \boldsymbol{\lambda}_+, \boldsymbol{\lambda}_- \geq 0 \quad (4d)$$

$$\boldsymbol{\mu}_+^\top (\mathbf{BC}^\top \boldsymbol{\theta} - \bar{\mathbf{p}}) = \boldsymbol{\mu}_-^\top (\underline{\mathbf{p}} - \mathbf{BC}^\top \boldsymbol{\theta}) = 0 \quad (4e)$$

$$\boldsymbol{\lambda}_+^\top (\mathbf{s}^g - \bar{\mathbf{s}}^g) = \boldsymbol{\lambda}_-^\top (\underline{\mathbf{s}}^g - \mathbf{s}^g) = 0, \quad (4f)$$

where

$$\mathbf{M} := \begin{bmatrix} \mathbf{CBC}^\top \\ \mathbf{e}_1^\top \end{bmatrix}$$

is an  $(N+1)$ -by- $N$  matrix with rank  $N$ , and  $\mathbf{e}_1$  denotes the standard first basis vector.

We adopt the DC model for many reasons. First, it is a good approximation of the AC model while being simple and more tractable. It also serves as a first step towards the study of an AC-OPF operator. Second, the DC model is widely used by system operators for transmission networks [6]. Therefore,

<sup>1</sup>Though two problems have different decision variables, one can always replace  $\boldsymbol{\theta}$  in (3) by  $(\mathbf{CBC}^\top)^\dagger[(\mathbf{s}^g)^\top, (-\mathbf{s}^l)^\top]^\top$  to absorb the additional decision variable  $\boldsymbol{\theta}$ .

studying the sensitivity of the DC model, even although it is an approximation, can have an impact on how the real system is operated.

The formulation in (3) is equipped with a linear cost function. In fact, our results can be applied to more general situations; in particular we can handle cost functions of the form  $\sum_{i \in \mathcal{V}_G} f_i(\mathbf{s}_i^g)$  where each  $f_i$  is strictly increasing, convex and piecewise linear. In this case, one can always convert each generator into multiple sub-generators, and each of which corresponds to a specific linear segment of the cost. The actual marginal cost for all those sub-generators will be determined by the *merit order*, which would take the same value as the original piecewise linear cost function [8]. Since any strictly increasing and convex function can be approximated by a piecewise linear function, our results can be applied (with arbitrarily small approximation error) to more general cost functions including quadratic cost.

### C. OPF as an operator: $\mathcal{OPF}$

We will now describe how to formulate the DC-OPF (3) as a mapping from load to (optimal) generation space. We assume throughout the paper that the topology of the network remains constant, as do the line susceptances. These assumptions imply that the graph Laplacian given by  $\mathbf{CBC}^\top$  does not change. Let  $\boldsymbol{\xi} := [(\bar{\mathbf{s}}^g)^\top, (\mathbf{s}^g)^\top, \bar{\mathbf{p}}^\top, \mathbf{p}^\top]^\top \in \mathbb{R}^{2N_G+2E}$  be the vector of system limits. Define

$$\Omega_{\boldsymbol{\xi}} := \{\boldsymbol{\xi} \mid \bar{\mathbf{s}}^g \geq 0, (3b) - (3e) \text{ are feasible for some } \mathbf{s}^l > 0\}.$$

The set  $\Omega_{\boldsymbol{\xi}}$  defines the set of power flow and generation limits such that the DC-OPF is primal-dual feasible and makes physical sense i.e. upper-limits are greater than lower-limits.

For each  $\boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}$ , define<sup>2</sup>

$$\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}) := \{\mathbf{s}^l \mid \mathbf{s}^l > 0, (3b) - (3e) \text{ are feasible}\}.$$

Then  $\Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$  is convex and nonempty. When we fix  $\boldsymbol{\xi}$  and there is no confusion, we simply write  $\Omega_{\mathbf{s}^l}$ .

**Definition 1:** Define  $\Omega := \{(\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l) \mid \mathbf{f} \in \mathbb{R}_+^{N_G}, \boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}, \mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})\}$ .

When  $\boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}$  and  $\mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$  the DC-OPF problem (3) is feasible. As (3b) fixes the angle  $\theta_1$  at the slack bus, and (3b) restricts the angle difference between any two adjacent buses we conclude that the feasible set of (3) is compact, and thus, by Weirstrass' Theorem, the optimal solutions to (3) always exist. We now define the operator  $\mathcal{OPF}$ , which is the central object of study in this paper.

**Definition 2:** For given  $\mathbf{f} \in \mathbb{R}_+^{N_G}$  and  $\boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}$ , the set valued operator  $\mathcal{OPF} : \Omega_{\mathbf{s}^l} \rightarrow 2^{\mathbb{R}^{N_G}}$  is the mapping such that  $\mathcal{OPF}(\mathbf{x})$  is the set of optimal solutions to (3) with parameter  $\mathbf{s}^l = \mathbf{x}$ .

In the following section we will establish various properties of the  $\mathcal{OPF}$  operator and show that it is a valuable tool for gaining insight into the sensitivity, robustness, and structure of the DC-OPF problem (3).

<sup>2</sup>In practice, if a load has 0 value, one could replace it by an arbitrarily small positive value so that the load profile is always strictly positive.

## III. OPERATOR PROPERTIES

The operator  $\mathcal{OPF}$  is parameterized by  $\mathbf{f}$ ,  $\boldsymbol{\xi}$  and acts on  $\mathbf{s}^l$ . The set  $\Omega$  defined in Definition 1 prescribes all the parameters under which (3) is feasible.

### A. Uniqueness

We are specifically interested in the case when the OPF operator defined above maps to a singleton. Adapting from the results in [9], one can easily show that OPF has a unique solution for almost all choices of  $\mathbf{f}$ . However, to pave the way for further properties in the following subsections, we in fact consider the vector  $\mathbf{f}$  under heavier constraints. Let  $\Omega_{\mathbf{f}}$  be the set of vectors  $\mathbf{f} \geq 0$  such that  $\forall \boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}, \forall \mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ :

- DC-OPF problem (3) has a unique solution.
- The KKT-system (4) satisfies

$$\|\boldsymbol{\mu}_+\|_0 + \|\boldsymbol{\mu}_-\|_0 + \|\boldsymbol{\lambda}_+\|_0 + \|\boldsymbol{\lambda}_-\|_0 \geq N_G - 1. \quad (5)$$

**Proposition 1:**  $\Omega_{\mathbf{f}}$  is dense in  $\mathbb{R}_+^{N_G}$ .

*Proof:* See Appendix A. ■

Proposition 1 shows that for a fixed network, it is easy to find an objective vector  $\mathbf{f}$  such that (3) not only has a unique solution for feasible  $\mathbf{s}^l$ , but also gives sufficiently many non-zero dual variables. For the remainder of the paper, the following assumption is in play:

**Assumption 1:** The objective vector  $\mathbf{f}$  is in  $\Omega_{\mathbf{f}}$ .

This assumption ensures that (3) has to have a unique solution. When Assumption 1 does not hold, Proposition 1 implies that we can always perturb  $\mathbf{f}$  such that the assumption is valid.

**Remark 1:** Under Assumption 1, the value of  $\mathcal{OPF}$  is always a singleton, so we can overload  $\mathcal{OPF}(\mathbf{x})$  as the function mapping from  $\mathbf{x}$  to the unique optimal solution of (3) with parameter  $\mathbf{s}^l = \mathbf{x}$ . Since the solution set to the parametric linear program is both upper and lower hemi-continuous [38],  $\mathcal{OPF}$  is continuous as well.

### B. Independent binding constraints

The analysis on the OPF operator can usually be simplified if the set of binding (active) constraints at the optimal point is independent. Here, binding constraints refer to the set of equality constraints (3b), (3c), and those inequality constraints (3d), (3e) for which either the upper or lower-bounds are active. Grouping the coefficients of these constraints into a single matrix  $\mathbf{Z}$  we refer to them as being independent if  $\mathbf{Z}$  is full-rank. Finally, define

$$\tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi}, \mathbf{f}) := \{\mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi}) \mid (3) \text{ has exactly } N_G - 1 \text{ binding inequalities at the optimal point, given } \mathbf{s}^l\}.$$

When  $\mathbf{f}$  is fixed, we shorten  $\tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi}, \mathbf{f})$  as  $\tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi})$ . Further, if  $\boldsymbol{\xi}$  is also fixed, then we will simply use  $\tilde{\Omega}_{\mathbf{s}^l}$ .

**Theorem 1:** For a fixed  $\mathbf{f} \in \Omega_{\mathbf{f}}$ , there exists a dense set  $\tilde{\Omega}_{\boldsymbol{\xi}}(\mathbf{f}) \subseteq \Omega_{\boldsymbol{\xi}}$  such that  $\forall \boldsymbol{\xi} \in \tilde{\Omega}_{\boldsymbol{\xi}}(\mathbf{f})$ , the following statements are true:

- $\text{clos}(\text{int}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))) = \text{clos}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))$ .
- $\tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi}, \mathbf{f})$  is dense in  $\Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ .

*Proof:* See Appendix B. ■

*Assumption 2:* The parameter  $\xi$  for the limits of generations and branch power flows is assumed to be in  $\tilde{\Omega}_\xi(\mathbf{f})$ , as proposed in Theorem 1.

Assumption 2 allows one to work with sets  $\tilde{\Omega}_\xi(\mathbf{f})$  that are well behaved (where “well behaved” is interpreted as  $\Omega_{s^l}$  and  $\text{int}(\Omega_{s^l})$  having the same closure and there being exactly  $N_G - 1$  binding constraints at the optimal point in the associated DC-OPF problem for almost every  $s^l$ ). This assumption is important as in Section IV it will be needed to show that the derivative of  $\mathcal{OPF}$  exists and can be efficiently represented almost everywhere. If Assumption 2 does not hold, Theorem 1 implies that we can always perturb  $\xi$  such that the assumption holds. In the context of DC-OPF problem, it also means for almost all the problem instances, there are exactly  $N_G - 1$  binding generators or power lines at the optimal point. The proof of Theorem 1 can directly extend to the following two corollaries:

*Corollary 1:* The set  $\Omega_{s^l} \setminus \tilde{\Omega}_{s^l}$  can be covered by the union of finitely many affine hyperplanes.

*Corollary 2:* For any  $s^l \in \Omega_{s^l}$ , the  $N_G - 1$  tight inequalities in (3), along with  $N + 1$  equality constraints, are independent.

*Definition 3:* Define  $\tilde{\Omega} := \{(\mathbf{f}, \xi, s^l) \mid \mathbf{f} \in \Omega_f, \xi \in \tilde{\Omega}_\xi(\mathbf{f}), s^l \in \tilde{\Omega}_{s^l}(\xi, \mathbf{f})\}$ .

In summary, the two sets  $\Omega$  and  $\tilde{\Omega}$  characterize sets of objective functions, problem parameters, and “inputs” that endow  $\mathcal{OPF}$  with desirable properties. In particular  $\Omega$  guarantees (3) is feasible and  $\mathcal{OPF}$  is thereby well-defined. The parameters in  $\tilde{\Omega}$  additionally guarantee that (3) has independent binding constraints and  $\mathcal{OPF}$  is singleton-valued, and as will be shown in the next section,  $\mathcal{OPF}$  is differentiable when  $(\mathbf{f}, \xi, s^l) \in \tilde{\Omega}$ . The relationship among the sets  $\Omega_f$ ,  $\Omega_{s^l}$ ,  $\tilde{\Omega}_{s^l}$ ,  $\Omega_\xi$ ,  $\tilde{\Omega}_\xi$  defined above is illustrated in Fig. 1. Recall that informally, the set  $\Omega_\xi$  contains all the  $\xi$  that make the OPF problem feasible, and  $\Omega_f$  contains  $\mathbf{f}$  that guarantee the unique optimal solution for feasible OPF problems and sufficiently many non-zero Lagrange multipliers. Proposition 1 shows  $\Omega_f$  is dense in  $\mathbb{R}_+^{N_G}$ . Each  $\xi \in \Omega_\xi$  maps to a set  $\Omega_{s^l}(\xi)$ , while each  $(\xi, \mathbf{f})$  maps to set  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$ , which is a subset of  $\Omega_{s^l}(\xi)$ . For fixed  $\mathbf{f}$ , by collecting all the  $\xi$  such that  $\Omega_{s^l}(\xi)$  has “good” topological property and  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$  is dense in  $\Omega_{s^l}(\xi)$ , we obtain a set  $\tilde{\Omega}_\xi(\mathbf{f})$  depending on  $\mathbf{f}$ , and Proposition 1 implies  $\tilde{\Omega}_\xi(\mathbf{f})$  is always dense in  $\Omega_\xi$ .

Since the sets that imply “good” properties ( $\Omega_f$ ,  $\tilde{\Omega}_{s^l}$ ,  $\tilde{\Omega}_\xi$ ) are all dense with respect to the corresponding whole sets of interest ( $\mathbb{R}_+^{N_G}$ ,  $\Omega_{s^l}$ ,  $\Omega_\xi$ ), one can always perturb the parameters to endow  $\mathcal{OPF}$  with these desirable properties.

Some practical implications of those results can be found in [39]. For example, one direct consequence is that for almost all problem instances, there must be at least one non-binding generator (i.e. a generator which is not inactive or saturated) in each subnetwork partitioned by binding branches. Therefore for radial power networks, if a single load is changed by a tiny amount, then only one generator will respond to this change in its optimal generation, which will change by the same amount. As a result, the locational marginal prices (LMP) for radial networks are always strictly positive. However, such result does not hold in general for meshed networks. In fact, for meshed networks the increase in a single loads can lead

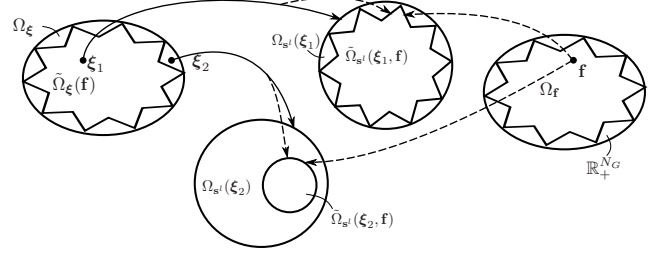


Fig. 1: Relationship among definitions in Section II-C. Solid arrows show the mapping from  $\xi$  to  $\Omega_{s^l}(\xi)$ , and dashed arrows show the mapping from  $(\xi, \mathbf{f})$  to  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$ . A star inscribed within an oval indicates the former set is dense within the latter.

to both increase and decrease in the optimal generation at multiple different locations, and this effect is well known as the cause of negative LMPs. For almost all the problem instances with “good” properties, i.e. those we restrict the parameter sets to  $(\Omega_f, \Omega_{s^l}, \Omega_\xi)$ , our work [39] further suggests that the computation of the sensitivity of the optimal generation with respect to loads can be decomposed and is thus tractable even for huge networks.

#### IV. OPF DERIVATIVE

In this section we show that  $\mathcal{OPF}$  is differentiable almost everywhere. We also provide an equivalent perspective from which to view the derivative (Jacobian matrix) of  $\mathcal{OPF}$  in terms of binding constraints, and derive its closed-form expression.<sup>3</sup>

##### A. Existence

Before deriving the expressions for the  $\mathcal{OPF}$  derivative, it is necessary to guarantee that the operator is in fact differentiable. The following lemma of [40] and [12] gives a sufficient condition of differentiability. We rephrase the lemma as follows.

*Lemma 1 ([40], [12]):* Consider a generic optimization problem parametrized by  $\Theta$ :

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} \quad f(\mathbf{x}; \Theta) \quad (6a)$$

$$\text{subject to} \quad g_i(\mathbf{x}; \Theta) \leq 0, i = 1, 2, \dots, m \quad (6b)$$

$$h_j(\mathbf{x}; \Theta) = 0, j = 1, 2, \dots, l. \quad (6c)$$

If  $(\mathbf{x}^*, \boldsymbol{\eta}^*, \boldsymbol{\nu}^*)$  is a primal-dual optimal solution for some  $\Theta_0$  and satisfies:

- 1)  $\mathbf{x}^*$  is a locally unique primal solution.
- 2)  $f, g_i, h_j$  are twice continuously differentiable in  $\mathbf{x}$  and differentiable in  $\Theta$ .
- 3) The gradients  $\nabla g_i(\mathbf{x}^*)$  for binding inequality constraints and  $\nabla h_j(\mathbf{x}^*)$  for equality constraints are independent.

<sup>3</sup>A word on notation is in order here. We denote the derivative of  $f(x)$  with respect to  $x$  by  $\partial_x f$ , however in some cases when there are complex dependencies on  $x$  we will use  $\frac{\partial f}{\partial x}$ . In Section V when we deal with derivatives of conic programs we use the notationally lighter differential operator  $D$ .

4) Strict complementary slackness holds.

Then the local derivative  $\partial_{\Theta} \mathbf{x}^*$  exists at  $\Theta_0$ , and the set of binding constraints is unchanged in some neighborhood of  $\Theta_0$ .

Using the set definitions from the previous section and the above lemma, we obtain the following result:

*Corollary 3:* Under Assumptions 1 and 2, for  $\mathbf{s}^l \in \tilde{\Omega}_{\mathbf{s}^l}$ , the derivative  $\partial_{\mathbf{s}^l} \mathcal{OPF}(\mathbf{s}^l)$  always exists, and the set of binding constraints stay unchanged in some neighborhood of  $\mathbf{s}^l$ .

*Proof:* By checking the conditions 1-4 in Lemma 1, the proof is established. Alternatively, the theorem can be proved by extending Proposition 3.2 in [11]. ■

Having established the existence of the derivative of  $\mathcal{OPF}$  we are now ready to study the associated Jacobian matrix.

### B. Jacobian matrix

The Jacobian is an important tool in sensitivity analysis as it provides the best linear approximation of an operator from input to output space. The results of the previous section ensure that the partial derivatives exist almost everywhere. Let

$$\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi}) := \partial_{\mathbf{s}^l} \mathcal{OPF}(\mathbf{s}^l) \quad (7)$$

for  $(\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l) \in \tilde{\Omega}$  denote the Jacobian of  $\mathcal{OPF}$  at  $\mathbf{s}^l$ . To reduce the notational burden, we will simply use  $\mathbf{J}$  or  $\mathbf{J}(\mathbf{s}^l)$  for short when the value of  $(\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l)$  or  $(\mathbf{f}, \boldsymbol{\xi})$  is clear from context. Suppose at point  $\mathbf{s}^l$ , the set of generators corresponding to binding inequalities is  $\mathcal{S}_G \subseteq \mathcal{V}_G$ , while the set of branches corresponding to binding inequalities is  $\mathcal{S}_B \subseteq \mathcal{E}$ . From Theorem 1 and Assumption 2 the following corollary is immediate:

*Corollary 4:*  $|\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1$ .

As Lemma 1 implies that generators  $\mathcal{S}_G$  and branches  $\mathcal{S}_B$  still correspond to binding constraints near  $\mathbf{s}^l$ , there is a local relationship between  $(\mathbf{s}^g)^* = \mathcal{OPF}(\mathbf{s}^l)$  and  $\mathbf{s}^l$ :

$$\mathbf{H} \begin{bmatrix} \mathbf{s}^g \\ \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{0}^{N_G \times 1} \\ -\mathbf{s}^l \\ \gamma^T \boldsymbol{\xi} \\ 0 \end{bmatrix}, \mathbf{H} := \left[ \begin{array}{c|c} -\mathbf{I}^{N_G} & \mathbf{I}_{\mathcal{V}_G}^N \mathbf{CBC}^T \\ \hline \mathbf{0}^{N_L \times N_G} & \mathbf{I}_{\mathcal{V}_L}^N \mathbf{CBC}^T \\ \mathbf{I}_{\mathcal{S}_G}^{N_G} & \mathbf{0}^{|\mathcal{S}_G| \times N} \\ \mathbf{0}^{|\mathcal{S}_B| \times N_G} & \mathbf{I}_{\mathcal{S}_B}^E \mathbf{BC}^T \\ \hline \mathbf{0}^{1 \times N_G} & \mathbf{e}_1^T \end{array} \right] \quad (8)$$

When there is no danger of confusion, we use  $\mathbf{s}^g$  to denote  $(\mathbf{s}^g)^*$ .

On the right hand side,  $\gamma \in \mathbb{R}^{(2N_G + 2E) \times (N_G - 1)}$  where each column of  $\gamma$  is a basis vector such that  $\gamma^T \boldsymbol{\xi}$  gives a vector of capacity limits that binding generations and branch power flows hit. By Corollary 2, the first  $N + N_G - 1$  rows of  $\mathbf{H}$  are independent, and clearly the last row  $[0, \mathbf{e}_1^T]$  does not depend on the first  $N + N_G - 1$  rows. Hence  $\mathbf{H}$  is invertible, and using the block matrix inversion formula, we have

$$\begin{bmatrix} \mathbf{s}^g \\ \boldsymbol{\theta} \end{bmatrix} = \mathbf{H}^{-1} \begin{bmatrix} \mathbf{0}^{N_G \times 1} \\ -\mathbf{s}^l \\ \gamma^T \boldsymbol{\xi} \\ 0 \end{bmatrix} = \begin{bmatrix} * & \mathbf{H}_1 \\ * & * \end{bmatrix} \begin{bmatrix} \mathbf{0}^{N_G \times 1} \\ -\mathbf{s}^l \\ \gamma^T \boldsymbol{\xi} \\ 0 \end{bmatrix} \quad (9)$$

with  $\mathbf{H}_1 = \mathbf{I}_{\mathcal{V}_G}^N \mathbf{CBC}^T (\mathbf{R}(\mathcal{S}_G, \mathcal{S}_B)^T)^{-1}$  and

$$\mathbf{R}(\mathcal{S}_G, \mathcal{S}_B)^T := \begin{bmatrix} \mathbf{I}_{\mathcal{V}_L}^N \mathbf{CBC}^T \\ \mathbf{I}_{\mathcal{S}_G}^N \mathbf{CBC}^T \\ \mathbf{I}_{\mathcal{S}_B}^E \mathbf{BC}^T \\ \mathbf{e}_1^T \end{bmatrix}. \quad (10)$$

Recall (7) that  $\mathbf{s}^g = \mathcal{OPF}(\mathbf{s}^l)$  in (9), so the Jacobian matrix  $\mathbf{J}$  is

$$\mathbf{J} = -\mathbf{H}_1 (\mathbf{I}_{[N_L]}^N)^T. \quad (11)$$

It is worth noting that the value of  $\mathbf{J}$  computed via (8)-(11) depends on knowing the binding constraints  $\mathcal{S}_G$  and  $\mathcal{S}_B$  for given  $(\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l)$ . We abuse notation slightly and let  $\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi})$  be the Jacobian matrix when  $(\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l) \in \tilde{\Omega}$  is known and let  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$  be the Jacobian when  $(\mathcal{S}_G, \mathcal{S}_B)$  is known. When it is clear from context or not relevant we simply use  $\mathbf{J}$ .

### C. Range of OPF derivative

The previous subsection has shown that the value of  $\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi})$  is equivalent to  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$  for certain choices of  $\mathcal{S}_G$  and  $\mathcal{S}_B$ . The following theorem also implies the equivalence between the range of  $\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi})$  and  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$ .<sup>4</sup>

*Theorem 2:*

$$\begin{aligned} & \{\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi}) \mid (\mathbf{f}, \boldsymbol{\xi}, \mathbf{s}^l) \in \tilde{\Omega}\} \\ &= \{\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B) \mid \mathcal{S}_G \in \mathcal{V}_G, \mathcal{S}_B \in \mathcal{E}, \mathcal{S}_G \perp \mathcal{S}_B, \\ & \quad |\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1\}. \end{aligned} \quad (12)$$

Here, we use  $\mathcal{S}_G \perp \mathcal{S}_B$  to denote that in (3), all the inequality constraints corresponding to  $\mathcal{S}_G$  and  $\mathcal{S}_B$ , as well as equality constraints, are independent of each other. Notice that the left hand side of (12) is induced by the DC-OPF problem and hence involves physical parameters such as the cost function, generation and load. The right hand side, however, purely depends on the graph topology. Theorem 2 shows the equivalence between the value ranges of  $\mathbf{J}(\mathbf{s}^l)$  and  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$ . While one direction,  $\text{Range}(\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)) \supseteq \text{Range}(\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi}))$ , trivially hold for general linear programs with a graph structure, the other direction,  $\text{Range}(\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)) \subseteq \text{Range}(\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi}))$ , is in fact non-trivial and does not hold for general LPs. This is because it requires any valid combination of generators and branches to be achievable by a problem with physically meaningful parameters (e.g., non-negative  $\mathbf{f}$  and  $\mathbf{s}^g$ ). However, Theorem 2 shows for DC-OPF, any valid combination is indeed achievable.

We first provide the following lemmas in order to build up to the final proof for Theorem 2.

*Lemma 2:* For any  $\mathcal{S}_G \in \mathcal{V}_G, \mathcal{S}_B \in \mathcal{E}$  such that  $|\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1$  and  $\mathcal{S}_G \perp \mathcal{S}_B$ , there exist  $(\mathbf{f}_*, \boldsymbol{\xi}_*, \mathbf{s}_*^l) \in \Omega$  such that (3) has unique solution and all the binding constraints at the solution point exactly correspond to  $\mathcal{S}_G$  and  $\mathcal{S}_B$ .

*Proof:* See [41]. ■

*Lemma 3:* For any  $\mathcal{S}_G \in \mathcal{V}_G, \mathcal{S}_B \in \mathcal{E}$  such that  $|\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1$  and  $\mathcal{S}_G \perp \mathcal{S}_B$ , there exist  $\mathbf{f}_{**} \in \Omega_{\mathbf{f}}, \boldsymbol{\xi}_{**} \in \Omega_{\boldsymbol{\xi}}(\mathbf{f}_{**})$  and an open ball  $W \subseteq \tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi}_{**}, \mathbf{f}_{**})$  such that all the

<sup>4</sup>Here, the range refers to the set of values that  $\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi})$  or  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$  could take, rather than the column space of  $\mathbf{J}(\mathbf{s}^l; \mathbf{f}, \boldsymbol{\xi})$  or  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$ .

binding constraints exactly correspond to  $\mathcal{S}_G$  and  $\mathcal{S}_B$  whenever  $s^l \in W$ .

*Proof:* See [41]. ■

*Proof:* (Theorem 2) For any  $(\mathbf{f}, \boldsymbol{\xi}, s^l) \in \tilde{\Omega}$ , by definition the binding constraints  $\mathcal{S}_G$  and  $\mathcal{S}_B$  must satisfy  $|\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1$  and  $\mathcal{S}_G \perp \mathcal{S}_B$ . Thus the left hand side of (12) is a subset of the right hand side of (12). As for the opposite direction, Lemma 3 implies for any  $(\mathcal{S}_G, \mathcal{S}_B)$  such that  $|\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1$  and  $\mathcal{S}_G \perp \mathcal{S}_B$  we can always find  $(\mathbf{f}, \boldsymbol{\xi}, s^l) \in \tilde{\Omega}$  whose associated binding constraints exactly correspond to  $(\mathcal{S}_G, \mathcal{S}_B)$ . Hence (12) holds. ■

The result of Theorem 2 indicates there exists a surjection from  $\tilde{\Omega}$  to the set  $\{(\mathcal{S}_G, \mathcal{S}_B) \mid |\mathcal{S}_G| + |\mathcal{S}_B| = N_G - 1, \mathcal{S}_G \perp \mathcal{S}_B\}$  and the derivative of the operator (depending on the parameters) and the Jacobian matrix (depending on the binding constraints) take the same value under such surjection. If one is only interested in the range of  $\mathbf{J}(s^l; \mathbf{f}, \boldsymbol{\xi})$  such as the worst-case analysis instead of the value at a specific point, then it is tight to study  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$  instead. One benefit of studying  $\mathbf{J}(\mathcal{S}_G, \mathcal{S}_B)$  is it has a closed-form expression and only depends on the graph topology of the system. For instance, it helps us study the OPF sensitivity bounds in the “worst case” [39], which provides privacy guarantees when releasing power flow data [16].

## V. COMPUTATION

In this section, we will show how recent results on conic problem differentiation can be applied to the OPF operator. To compare this section and the result in Subsection IV-B, this method provides an efficient way to compute the derivative of the optimal solution with respect to more general parameters, and could also be extended to other power flow models, while Subsection IV-B is specifically for  $\mathcal{OPF}$  operator under DC model and focuses more on deriving the closed-form expression for analytical rather than computational studies.

### A. Differentiating a General Conic Program

The method of computation we pursue largely follows that presented in [14] which considers general convex conic optimization problems that are solved using the homogenous self-dual embedding framework [42], [43]. Consider a standard primal-dual pair written in conic form:

$$\begin{aligned} \text{(P)} \quad & \underset{(\mathbf{x}, \mathbf{s}) \in \mathbb{R}^n \times \mathcal{K}}{\text{minimize}} && \mathbf{c}^\top \mathbf{x} \\ & \text{subject to} && \mathbf{A}\mathbf{x} + \mathbf{s} = \mathbf{b}, \\ \text{(D)} \quad & \underset{(\mathbf{r}, \mathbf{y}) \in \{0\}^n \times \mathcal{K}^*}{\text{minimize}} && \mathbf{b}^\top \mathbf{y} \\ & \text{subject to} && \mathbf{A}^\top \mathbf{y} + \mathbf{c} = \mathbf{r}. \end{aligned}$$

In this setting the problem data consists of the triple  $(\mathbf{A}, \mathbf{b}, \mathbf{c}) \in \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n$ . The primal variable is  $\mathbf{x} \in \mathbb{R}^n$ , the primal slack variable is  $\mathbf{s} \in \mathbb{R}^m$ , and the dual variable is  $\mathbf{y} \in \mathbb{R}^m$ , with  $\mathbf{r} \in \mathbb{R}^n$  the dual slack variable. The set  $\mathcal{K}$  in a non-empty, closed, convex cone with  $\mathcal{K}^*$  its dual. Linear programming falls into this class of conic problems by setting  $\mathcal{K}$  to be the positive orthant. We will only provide the main technical results from [14] in this subsection, the detailed derivation can be found in the original paper.

The homogenous self-dual embedding formulation is expressed as

$$\begin{aligned} & \text{find} && (\mathbf{u}, \mathbf{v}) \\ & \text{subject to} && \mathbf{v} = \mathbf{Q}\mathbf{u} \\ & && (\mathbf{u}, \mathbf{v}) \in \mathcal{C} \times \mathcal{C}^* \end{aligned} \quad (13)$$

with cones  $\mathcal{C} = \mathbb{R}^n \times \mathcal{K}^* \times \mathbb{R}_+$  and its dual  $\mathcal{C}^* = \{0\}^n \times \mathcal{K} \times \mathbb{R}_+$ . The variables  $\mathbf{u}$  and  $\mathbf{v}$  correspond to variables in (P) and (D) and two augmented variables  $\kappa$  and  $\tau$ , and satisfy:

$$\underbrace{\begin{bmatrix} \mathbf{r} \\ \mathbf{s} \\ \kappa \end{bmatrix}}_{\mathbf{v}} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{A}^\top & \mathbf{c} \\ -\mathbf{A} & \mathbf{0} & \mathbf{b} \\ -\mathbf{c}^\top & -\mathbf{b}^\top & \mathbf{0} \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \tau \end{bmatrix}}_{\mathbf{u}}, \quad (\tau, \kappa) \in \mathbb{R}_+ \times \mathbb{R}_+$$

Using Minty’s parametrization [44], one can reformulate (13) as (where  $\mathbf{z}$  denotes  $\mathbf{u} - \mathbf{v}$ )

$$\begin{aligned} & \text{find} && \mathbf{z} = (\mathbf{z}_1 \in \mathbb{R}^n, \mathbf{z}_2 \in \mathbb{R}^m, \mathbf{z}_3 \in \mathbb{R}) \in \mathbb{R}^{n+m+1} \\ & \text{subject to} && -\mathcal{P}_{\mathcal{C}^*}\mathbf{z} = \mathbf{Q}\mathcal{P}_{\mathcal{C}}\mathbf{z}, \quad \mathbf{z}_3 > 0. \end{aligned} \quad (14)$$

The *solution map* is defined as  $\mathcal{S} : \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^{2m+n}$  which “pushes” the problem data  $(\mathbf{A}, \mathbf{b}, \mathbf{c})$  through optimization problem (13) to return  $(\mathbf{x}, \mathbf{y}, \mathbf{s})$  – the primal-dual solutions. The task here is to find  $(d\mathbf{x}, d\mathbf{y}, d\mathbf{s})$  from  $(d\mathbf{A}, d\mathbf{b}, d\mathbf{c})$ . The following result is taken from [14]. One can first derive the expression for  $d\mathbf{z}$  and then recover  $(d\mathbf{x}, d\mathbf{y}, d\mathbf{s})$  from  $d\mathbf{z}$ . Numerically  $d\mathbf{z} = -\mathbf{M}^{-1}\mathbf{g}$ , where

$$\begin{aligned} \mathbf{M} &= ((\mathbf{Q} - \mathbf{I})D\mathcal{P}_{\mathcal{C}}\mathbf{z} + \mathbf{I})/\mathbf{z}_3 \\ \mathbf{g} &= d\mathbf{Q}\mathcal{P}_{\mathcal{C}}(\mathbf{z}/\mathbf{z}_3) \\ d\mathbf{Q} &= \begin{bmatrix} \mathbf{0} & d\mathbf{A}^\top & d\mathbf{c} \\ -d\mathbf{A} & \mathbf{0} & d\mathbf{b} \\ -d\mathbf{c}^\top & -d\mathbf{b}^\top & \mathbf{0} \end{bmatrix}. \end{aligned} \quad (15)$$

Here for any set  $\mathcal{X}$ ,  $D\mathcal{P}_{\mathcal{X}}(\cdot)$  is the derivative of the projection operator  $\mathcal{P}_{\mathcal{X}}(\cdot)$ . Note that for large systems it may be preferable to not invert  $\mathbf{M}$  and instead solve a least squares problem. Finally, partition  $d\mathbf{z}$  conformally as  $(d\mathbf{z}_1, d\mathbf{z}_2, d\mathbf{z}_3)$  and compute

$$\begin{bmatrix} d\mathbf{x} \\ d\mathbf{y} \\ d\mathbf{s} \end{bmatrix} = \begin{bmatrix} d\mathbf{z}_1 - (d\mathbf{z}_3)\mathbf{x} \\ (D\mathcal{P}_{\mathcal{K}^*}(\mathbf{z}_2))d\mathbf{z}_2 - (d\mathbf{z}_3)\mathbf{y} \\ (D\mathcal{P}_{\mathcal{K}^*}(\mathbf{z}_2))d\mathbf{z}_2 - d\mathbf{z}_2 - (d\mathbf{z}_3)\mathbf{s} \end{bmatrix}. \quad (16)$$

The method outlined above provides us with more information than we have considered to this point. Specifically, it leverages information about the primal and dual conic forms and provides derivative information with respect to all problem data rather than just load changes.

### B. DC Optimal Power Flow

The DC OPF problem (3) can easily be written in the form (P) by introducing appropriate slack variables:

$$\begin{aligned} & \underset{\mathbf{x} := [(s^g)^\top, \boldsymbol{\theta}^\top]^\top, \mathbf{s}}{\text{minimize}} && [\mathbf{f}^\top, \mathbf{0}^\top]\mathbf{x} \\ & \text{subject to} && \begin{bmatrix} \mathbf{A}_{\text{eq}} \\ \mathbf{A}_{\text{in}} \end{bmatrix} \mathbf{x} + \mathbf{s} = \begin{bmatrix} \mathbf{b}_{\text{eq}} \\ \mathbf{b}_{\text{in}} \end{bmatrix} \\ & && (\mathbf{x}, \mathbf{s}) \in \mathbb{R}^{N_G+N} \times (\{0\}^{N+1} \times \mathbb{R}_+^{2N_G+2E}) \end{aligned}$$



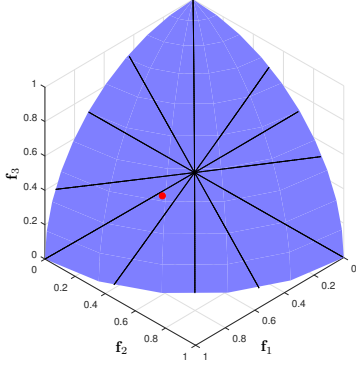


Fig. 2: Region for the cost vector  $\mathbf{f}$ . Black lines denote the values of  $\mathbf{f}$  in  $\mathbb{R}_+^{N_G} \setminus \Omega_{\mathbf{f}}$ . The red dot denotes the cost vector which will be used throughout this example.

where  $(\mathbf{A}_{\text{eq}}, \mathbf{A}_{\text{in}}, \mathbf{b}_{\text{eq}}, \mathbf{b}_{\text{in}})$  are as defined in (20). Noting that  $\mathcal{K}^* = \mathbb{R}^{N+1} \times \mathbb{R}_+^{2N_G+2E}$ .

Here we note that the derivative of the projection operator  $\mathcal{DP}_{\mathcal{C}}$  appearing in (15) is decomposed as

$$\mathcal{DP}_{\mathbb{R}^{N_G+N}} \times \mathcal{DP}_{\mathbb{R}^{N+1}} \times \mathcal{DP}_{\mathbb{R}_+^{2N_G+2E}} \times \mathcal{DP}_{\mathbb{R}_+}$$

and  $\mathcal{DP}_{\mathcal{K}^*}$  appearing in (16) is decomposed as

$$\mathcal{DP}_{\mathbb{R}^{N+1}} \times \mathcal{DP}_{\mathbb{R}_+^{2N_G+2E}}.$$

Specifically,  $\mathcal{DP}_{\mathbb{R}_+}$  is differentiable everywhere but at  $\{0\}$ , elsewhere  $\mathcal{DP}_{\mathbb{R}_+} \mathbf{x} = \frac{1}{2}(\text{sign}(\mathbf{x}) + 1)$ .

## VI. ILLUSTRATIVE EXAMPLES

In this section, we use the IEEE 9-bus test network as an example to illustrate what the sets  $(\Omega_{\mathbf{f}}, \Omega_{\xi}, \tilde{\Omega}_{\xi}, \Omega_{s^l}, \tilde{\Omega}_{s^l})$  in Fig. 1 look like. It has three generators (white circles) and 6 loads (black circles). The susceptances (edge weights) of power lines are taken from the MATPOWER toolbox [45]. The system parameters not provided in the text can be found in Table 1 of [41]. The data for the capacity limits and the loads are either directly taken from MATPOWER or perturbed to satisfy our assumptions.

First, we visualize and illustrate the sets  $\mathbb{R}_+^{N_G}$  and  $\Omega_{\mathbf{f}}$  where the cost vector  $\mathbf{f}$  resides. As we ignore the trivial case when  $\mathbf{f} = \mathbf{0}$ , we restrict  $\mathbf{f}$  to the unit sphere for visual clarity. As a result,  $\mathbb{R}_+^{N_G}$  is visualized by the blue region including the boundary and black curve segments shown in Fig. 2. The black curve segments represent the set of  $\mathbf{f}$  which may potentially make the OPF problem have multiple solutions or violate (5). Thereby the blue region excluding the black curve segments is the restriction of a subset of  $\Omega_{\mathbf{f}}$  onto the unit sphere. Figure 2 provides a visualization that  $\Omega_{\mathbf{f}}$  is dense in  $\mathbb{R}_+^{N_G}$  ( $N_G = 3$  in this example). If the cost vector  $\mathbf{f}$  is randomly chosen in  $\mathbb{R}_+^{N_G}$ , then we will almost surely obtain a well-behaved  $\mathbf{f}$  not aligned with the black curves. In the rest of this example, we randomly chose  $\mathbf{f} = [0.7191, 0.5066, 0.4758]^T$  and is shown as the red point in Fig. 2.

We will now visualize the sets  $\Omega_{\xi}$  and  $\tilde{\Omega}_{\xi}(\mathbf{f})$  for our choice of  $\mathbf{f}$ , and illustrate how different points in those two sets endow the OPF problem with different properties. Consider that there are 3 generators and 9 branches in the network, and each

generator and branch has both the upper and lower bounds for its generation and branch power flow, the vector  $\xi$  thus has 24 dimensions. In order to make visualization possible, we fix all the capacity limits except for the power flow limits at branch (2, 8). A positive power flow at branch  $(u, v)$  means that power is transmitted from  $u$  to  $v$ . Conversely, a negative value implies power is transmitted in the opposite direction. Figure 3 shows when  $\mathbf{f}$  and other capacity limits are fixed, how the upper and lower bounds for branch (2, 8) affect the OPF operator. In other words, Fig. 3 visualizes a slice of sets  $\Omega_{\xi}$  and  $\tilde{\Omega}_{\xi}(\mathbf{f})$ . The purple region, including the boundaries and black lines, is the slice of  $\Omega_{\xi}$ . Picking any point in the purple region as the capacity limits for branch (2, 8), there exist some  $s^l > 0$  such that the constraints (3b)-(3e) are feasible. However, for some points on the black lines or boundaries, the associated set  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$  might be not dense in  $\Omega_{s^l}(\xi)$ . We collect all the points in the purple region excluding the black lines and boundaries to form a slice of  $\tilde{\Omega}_{\xi}(\mathbf{f})$ , which is dense in  $\Omega_{\xi}$ . We now pick the red point in  $\tilde{\Omega}_{\xi}(\mathbf{f})$  (not on the black lines) and the black point in  $\Omega_{\xi} \setminus \tilde{\Omega}_{\xi}(\mathbf{f})$  (on the black line) as shown in Fig. 3, and will show their difference. Recall that in Fig. 1, we plot two points  $\xi_1 \in \tilde{\Omega}_{\xi}(\mathbf{f})$  and  $\xi_2 \in \Omega_{\xi} \setminus \tilde{\Omega}_{\xi}(\mathbf{f})$ , so the red point visualizes  $\xi_1$  while the black point visualizes  $\xi_2$ .

First, we pick the red point in Fig. 3, i.e., set the lower and upper bounds for branch (2, 8) at  $(-2.5490, 2.5695)$ , respectively. Since it is difficult to visualize all 6 loads, we fix buses 5, 6, 8 and 9, and visualize the region for buses 4 and 7 in Fig. 4a. The whole hexagon excluding the axes represents the slice of  $\Omega_{s^l}(\xi)$ , within which any point corresponds to a load profile which makes the OPF problem feasible. The whole region is further divided into seven colored subregions, and each of them refers to the set of load profiles under which the binding constraints of (3) do not change. In the interior of those subregions, there will be exactly  $N_G - 1 = 2$  independent binding inequality constraints. Depending on the physical meaning of binding inequalities, we use three colors to distinguish different subregions. Red indicates two binding constraints refer to two binding generators, green indicates one generator and one branch are binding, and purple indicates two binding branches. Only the interior of those colored subregions contribute to the set  $\Omega_{s^l}(\xi, \mathbf{f})$ , which guarantees the number of and independence among all the binding constraints. The operator  $\mathcal{OPF}$  is also guaranteed to be differentiable when the loads are picked in  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$ , and here the Jacobian matrix is given in Section IV-B in a closed form. From Fig. 4a we can see that when the red point is picked, the interior of all the subregions (i.e.,  $\tilde{\Omega}_{s^l}(\xi, \mathbf{f})$ ) is dense in the whole hexagon (i.e.,  $\Omega_{s^l}(\xi)$ ).

Next, we pick the black point in Fig. 3, i.e., set the lower and upper bounds for branch (2, 8) at  $(-3.0758, 3.0758)$ . In this case, the whole hexagon contains a large chunk of shaded area. For the load profile in the shaded area, there might be more than  $N_G - 1 = 2$  binding inequality constraints, and all the binding constraints are not independent any more. The Jacobian matrix we derived in Section IV-B is no longer valid. As the shaded area is non-negligible, the interior of all the subregions is not dense in the whole hexagon any more.



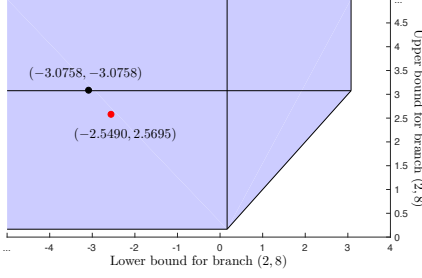
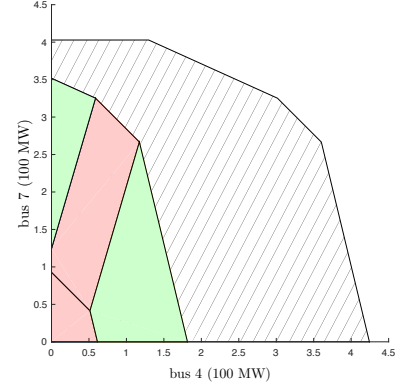
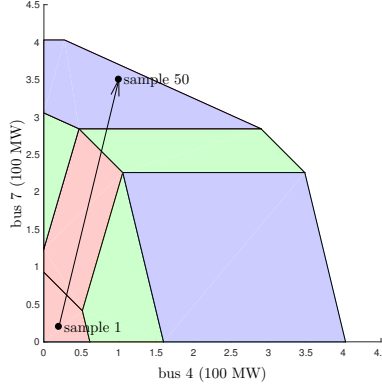


Fig. 3: Feasibility region for the power flow limits at branch (2, 8). The polytope (including its boundaries and the black lines) is (a slice of)  $\Omega_{\xi}$ . The set  $\tilde{\Omega}_{\xi}(\mathbf{f})$  is given by the purple region excluding the black lines and boundaries.



(a) When  $\xi$  is selected at the red point. (b) When  $\xi$  is selected at the black point.

Fig. 4: Region for the loads (bus 4 and bus 7).

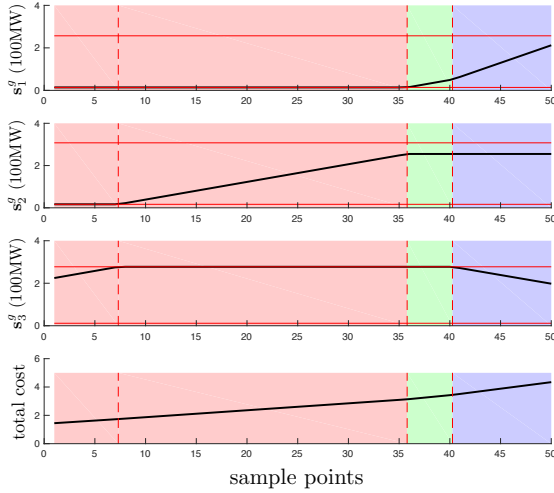


Fig. 5: Optimal generations and costs for different load profiles along the path plotted in Fig. 4a. The solid red horizontal lines indicate the upper and lower bounds for the generation.

Fortunately, both our proof and Fig. 3 show that for almost all the capacity limits, they will behave like the red point in the above example and guarantee the independence among binding constraints for almost all the feasible load profiles.

Finally, we consider a path in Fig. 4a which goes through four different subregions, and pick 50 sample points along the path. Each sample point corresponds to a specific load profile for the power system. In Fig. 5, we show how the optimal generations and costs change for those 50 sample load profiles. In each subregion, the gradient of the optimal solution stays unchanged until the load profile enters a new subregion.

## VII. CONCLUSION

We presented an approach for analyzing a linear program that solves the DC optimal power flow problem based on operator theoretic view of a linear program. Sets were defined upon which the OPF operator has a unique solution, is continuous, induces independent binding constraints, and the derivative exists (almost everywhere). Two equivalent perspectives on Jacobian matrix were given. The first was from the problem

data and the second from knowledge of the binding constraints. A closed-form expression of the Jacobian matrix is derived in terms of the binding constraints sets. Finally a numerical method based upon differentiating the solution map of a homogeneous self-dual conic program was described.

It is hoped that this formulation will provide practitioners with new tools for analyzing the robustness of their networks. Simultaneously, it opens up many interesting theoretical questions, in particular, the study of AC optimal power flow sensitivity. The main challenge to study AC-OPF operator is the non-convexity of AC power flow equations. As a first step, it is promising to extend our results to the SDP relaxation by leveraging the techniques on differentiating through a cone program. However, to recover a solution from the relaxation, a rank-1 solution is required. Therefore, it will be an interesting but also challenging task to study how the perturbation in problem parameters may affect the rank of relaxation solutions.

We are currently investigating how to compute the worst-case sensitivity of the DC-OPF problem as this appears in a diverse range of applications including differential privacy, real-time optimization, and locational marginal pricing.

## APPENDIX A PROOF OF PROPOSITION 1

We first define

$$\Omega_{\mathbf{f}}^{(1)} = \{\mathbf{f} \in \mathbb{R}_+^{N_G} \mid \forall \xi \in \Omega_{\xi}, \forall \mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\xi), (3) \text{ has unique optimal solution}\} \quad (17a)$$

$$\Omega_{\mathbf{f}}^{(2)} = \{\mathbf{f} \in \mathbb{R}_+^{N_G} \mid \forall \xi \in \Omega_{\xi}, \forall \mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\xi), \text{all solutions of (4) satisfy (5)}\}, \quad (17b)$$

then  $\Omega_{\mathbf{f}} = \Omega_{\mathbf{f}}^{(1)} \cap \Omega_{\mathbf{f}}^{(2)}$ . For  $\mathcal{S} \subseteq \mathcal{E}$ ,  $\mathcal{T} \subseteq \mathcal{V}_G$  such that  $|\mathcal{S}| + |\mathcal{T}| \leq N_G - 2$ , we construct  $\mathcal{Q}(\mathcal{S}, \mathcal{T})$  to be the set of  $\mathbf{f}$  such that  $\exists \boldsymbol{\tau} \in \mathbb{R}^{N+1}, \boldsymbol{\mu} \in \mathbb{R}^E, \boldsymbol{\lambda} \in \mathbb{R}^{N_G}$  satisfying:

$$\mathbf{0} = \mathbf{M}^T \boldsymbol{\tau} + \mathbf{C} \mathbf{B} \boldsymbol{\mu} \quad (18a)$$

$$-\mathbf{f} = -[\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_{N_G}]^T + \boldsymbol{\lambda} \quad (18b)$$

$$\boldsymbol{\mu}_i \neq 0 \Rightarrow i \in \mathcal{S} \quad (18c)$$

$$\boldsymbol{\lambda}_i \neq 0 \Rightarrow i \in \mathcal{T}. \quad (18d)$$

When  $\mathcal{S}$  and  $\mathcal{T}$  are fixed, the vector  $\mathbf{CB}\boldsymbol{\mu}$  takes value in an  $|\mathcal{S}|$  dimensional subspace. Since  $\text{rank}(\mathbf{M}) = N$ , the possible values of  $\boldsymbol{\tau}$  must fall within an  $|\mathcal{S}| + 1$  dimensional subspace. Therefore, (18b) implies that  $\mathbf{f}$  must be in an  $|\mathcal{S}| + 1 + |\mathcal{T}| \leq N_G - 1$  dimensional subspace, and hence  $\text{int}(\text{clos}(\mathcal{Q}(\mathcal{S}, \mathcal{T}))) = \emptyset$ . Denote

$$\mathcal{Q}_U := \left( \bigcup_{\substack{\mathcal{S} \subseteq \mathcal{E}, \mathcal{T} \subseteq \mathcal{V}_G \\ |\mathcal{S}| + |\mathcal{T}| \leq N_G - 2}} \mathcal{Q}(\mathcal{S}, \mathcal{T}) \right),$$

then,  $\mathcal{Q}_U \cap \mathbb{R}_+^{N_G}$  is nowhere dense in  $\mathbb{R}_+^{N_G}$ .

On one hand, (17b) and (18) imply that

$$\mathbb{R}_+^{N_G} \setminus \Omega_{\mathbf{f}}^{(2)} = \{\mathbf{f} \in \mathbb{R}_+^{N_G} \mid \exists \boldsymbol{\xi} \in \Omega_{\boldsymbol{\xi}}, \mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi}), \text{one solution of (4) violates (5)}\}$$

is a subset of  $\mathcal{Q}_U$ . Thereby,  $\mathbb{R}_+^{N_G} \setminus \mathcal{Q}_U \subseteq \Omega_{\mathbf{f}}^{(2)}$ .

On the other hand, we reformulate (3) as

$$\begin{aligned} & \underset{\mathbf{x} := [(\mathbf{s}^g)^T, \boldsymbol{\theta}^T]^T}{\text{minimize}} && [\mathbf{f}^T, \mathbf{0}^T] \mathbf{x} \end{aligned} \quad (19a)$$

$$\text{subject to} \quad \mathbf{A}_{\text{eq}} \mathbf{x} = \mathbf{b}_{\text{eq}} \quad (19b)$$

$$\mathbf{A}_{\text{in}} \mathbf{x} \leq \mathbf{b}_{\text{in}} \quad (19c)$$

where

$$\mathbf{A}_{\text{eq}} := \begin{bmatrix} \mathbf{0}^{1 \times N_G} & \mathbf{e}_1 \\ -\mathbf{I}^{N_G} & \mathbf{CBC}^T \\ \mathbf{0}^{N_L \times N_G} & \end{bmatrix}, \quad \mathbf{b}_{\text{eq}} := \begin{bmatrix} \mathbf{0}^{(1+N_G) \times 1} \\ -\mathbf{s}^l \end{bmatrix}, \quad (20a)$$

$$\mathbf{A}_{\text{in}} := \begin{bmatrix} \mathbf{0}^{E \times N_G} & \mathbf{BC}^T \\ \mathbf{0}^{E \times N_G} & -\mathbf{BC}^T \\ \mathbf{I}^{N_G} & \mathbf{0}^{N_G \times N} \\ -\mathbf{I}^{N_G} & \mathbf{0}^{N_G \times N} \end{bmatrix}, \quad \mathbf{b}_{\text{in}} := \begin{bmatrix} \bar{\mathbf{p}} \\ -\mathbf{p} \\ \bar{\mathbf{s}}^g \\ -\mathbf{s}^g \end{bmatrix}. \quad (20b)$$

Geometrically, an LP has multiple optimal solutions if and only if the objective vector is normal to the hyperplane defined by equality constraints and the set of inequality constraints which are binding for all the optimal solutions (i.e., corresponding rows in  $\mathbf{A}_{\text{eq}}$  and  $\mathbf{A}_{\text{in}}$ ). We collect the rows in  $\mathbf{A}_{\text{in}}$  which correspond to binding inequality constraints (for all the optimal solutions) and form a new matrix  $\tilde{\mathbf{A}}_{\text{in}}$ . Formally, let  $\mathcal{X}$  be the set of indices  $i$  such that the  $i^{\text{th}}$  row of  $\mathbf{A}_{\text{in}}$  corresponds to a binding constraint for all the optimal solutions, then  $\tilde{\mathbf{A}}_{\text{in}} = \mathbf{I}_{\mathcal{X}} \mathbf{A}_{\text{in}}$ . In our case, the objective vector  $[\mathbf{f}^T, \mathbf{0}^T]^T$  is an  $N_G + N$  dimensional vector, thus the row space of  $[\mathbf{A}_{\text{eq}}^T, \tilde{\mathbf{A}}_{\text{in}}^T]^T$  must have dimension  $\leq N_G + N - 1$  and  $[\mathbf{f}^T, \mathbf{0}^T]^T$  must be within this row space. As  $\mathbf{A}_{\text{eq}}$  has  $N + 1$  linearly independent rows, we can always find  $\leq N_G - 2$  independent rows of  $\tilde{\mathbf{A}}_{\text{in}}$  to form a new matrix  $\hat{\mathbf{A}}_{\text{in}}$  such that  $[\mathbf{A}_{\text{eq}}^T, \hat{\mathbf{A}}_{\text{in}}^T]^T$  and  $[\mathbf{A}_{\text{eq}}^T, \tilde{\mathbf{A}}_{\text{in}}^T]^T$  share the same row space. As a result,  $[\mathbf{f}^T, \mathbf{0}^T]^T$  can be represented as the linear combination of rows in  $[\mathbf{A}_{\text{eq}}^T, \hat{\mathbf{A}}_{\text{in}}^T]^T$ , and one can always find  $(\mathcal{S}, \mathcal{T}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\lambda})$  satisfying (18) and also  $|\mathcal{S}| + |\mathcal{T}| \leq N_G - 2$ . Hence  $\mathbb{R}_+^{N_G} \setminus \Omega_{\mathbf{f}}^{(1)}$  is also a subset of  $\mathcal{Q}_U$  and thus  $\mathbb{R}_+^{N_G} \setminus \mathcal{Q}_U \subseteq \Omega_{\mathbf{f}}^{(1)}$ .

Above all,  $\mathbb{R}_+^{N_G} \setminus \mathcal{Q}_U \subseteq \Omega_{\mathbf{f}}^{(1)} \cap \Omega_{\mathbf{f}}^{(2)} = \Omega_{\mathbf{f}}$ . Since  $\mathcal{Q}_U \cap \mathbb{R}_+^{N_G}$  is nowhere dense in  $\mathbb{R}_+^{N_G}$ ,  $\Omega_{\mathbf{f}}$  is dense in  $\mathbb{R}_+^{N_G}$ .

## APPENDIX B PROOF OF THEOREM 1

The following results are used in the proof of Theorem 1. Their proofs can be found in [41].

*Lemma 4:* The set  $\Omega_{\boldsymbol{\xi}}$  satisfies  $\text{clos}(\text{int}(\Omega_{\boldsymbol{\xi}})) = \text{clos}(\Omega_{\boldsymbol{\xi}})$ .

*Lemma 5:* Suppose the set  $\mathcal{S} \subseteq \mathbb{R}^n$  satisfies the condition that  $\text{clos}(\text{int}(\mathcal{S})) = \text{clos}(\mathcal{S})$ , and  $\mathcal{T}$  is an affine hyperplane with dimension strictly less than  $n$ . Then  $\mathcal{T}$  is nowhere dense in  $\mathcal{S}$ .

We now prove Theorem 1. Our strategy is to construct the set  $\tilde{\Omega}_{\boldsymbol{\xi}}(\mathbf{f})$  first, then prove  $\text{clos}(\text{int}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))) = \text{clos}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))$ , and finally show that  $\tilde{\Omega}_{\mathbf{s}^l}(\boldsymbol{\xi}, \mathbf{f})$  is dense in  $\Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ .

Consider the power flow equations below:

$$\mathbf{T}\boldsymbol{\theta} := \begin{bmatrix} \mathbf{CBC}^T \\ \mathbf{BC}^T \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} \mathbf{s}^g \\ -\mathbf{s}^l \\ \mathbf{p} \end{bmatrix}. \quad (21)$$

Proposition 1 and Assumption 1 show that there will always be at least  $N_G - 1$  binding inequality constraints. We use the following procedure to construct the set  $\tilde{\Omega}_{\boldsymbol{\xi}}$ .

- I. Let  $\tilde{\Omega}_{\boldsymbol{\xi}}$  be the subset of  $\Omega_{\boldsymbol{\xi}}$  where upper limits for generation/line flow are strictly larger than lower limits.
- II. For each  $\mathcal{S} \subseteq \mathcal{V}_G \cup [N + 1, N + E]$ , construct  $\mathbf{T}_{\mathcal{S}}$ .
  - a) If  $\text{rank}(\mathbf{T}_{\mathcal{S}}) = |\mathcal{S}|$ , then continue to another  $\mathcal{S}$ .
  - b) If  $\text{rank}(\mathbf{T}_{\mathcal{S}}) < |\mathcal{S}|$ , then consider

$$\Gamma := \prod_{i \in \mathcal{S} \cap \mathcal{V}_G} \{\mathbf{e}_i, \mathbf{e}_{N_G+i}\} \times \prod_{\substack{j \in \mathcal{E} \\ j+N \in \mathcal{S}}} \{\mathbf{e}_{2N_G+j}, \mathbf{e}_{2N_G+E+j}\}. \quad (22)$$

Now update  $\tilde{\Omega}_{\boldsymbol{\xi}}$  as

$$\tilde{\Omega}_{\boldsymbol{\xi}} \leftarrow \tilde{\Omega}_{\boldsymbol{\xi}} \setminus \bigcup_{\gamma \in \Gamma} \{\boldsymbol{\xi} \mid \exists \boldsymbol{\theta}, \text{ such that } \gamma^T \boldsymbol{\xi} = \mathbf{T}_{\mathcal{S}} \boldsymbol{\theta}\}. \quad (23)$$

III. Return  $\tilde{\Omega}_{\boldsymbol{\xi}}$ .

In the above procedure, an  $n$ -tuple of vectors is also regarded as a matrix of  $n$  columns and the product in (22) is Cartesian product.<sup>5</sup> Since  $\gamma \in \Gamma$  is of rank  $|\mathcal{S}|$  and  $\mathbf{T}_{\mathcal{S}} \boldsymbol{\theta}$  with  $\boldsymbol{\theta} \in \mathbb{R}^N$  defines a subspace of  $\leq |\mathcal{S}| - 1$  dimensions, each set of  $\{\boldsymbol{\xi} \mid \exists \boldsymbol{\theta}, \text{ such that } \gamma^T \boldsymbol{\xi} = \mathbf{T}_{\mathcal{S}} \boldsymbol{\theta}\}$  in (23) is a subspace with dimension strictly lower than  $2N_G + 2E$ , and is thereby nowhere dense in  $\Omega_{\boldsymbol{\xi}}$  by Lemma 5. As a result, we have that  $\tilde{\Omega}_{\boldsymbol{\xi}}$  is dense in  $\Omega_{\boldsymbol{\xi}}$ . It is sufficient to show that two conditions in Proposition 1 are satisfied.

To show  $\text{clos}(\text{int}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))) = \text{clos}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))$ , it is sufficient to prove that fix  $\boldsymbol{\xi} \in \tilde{\Omega}_{\boldsymbol{\xi}}$ ,  $\forall \hat{\mathbf{s}}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ , there exists a sequence  $(\mathbf{s}_{(n)}^l)_{n=1}^{\infty}$  such that  $\lim_{n \rightarrow \infty} \mathbf{s}_{(n)}^l = \hat{\mathbf{s}}^l$  and each  $\mathbf{s}_{(n)}^l$  has an open neighborhood  $U(\mathbf{s}_{(n)}^l)$  such that  $U(\mathbf{s}_{(n)}^l) \subseteq \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ . By definition, there exists  $\hat{\mathbf{s}}^g$  and  $\hat{\boldsymbol{\theta}}$  such that (3b)-(3e) are satisfied for  $\hat{\mathbf{s}}^l$ . We also use  $\hat{\mathbf{p}}$  to denote the branch power flow associated with  $(\hat{\mathbf{s}}^g, \hat{\mathbf{s}}^l)$ . Here we overload  $\mathcal{S} \subseteq \mathcal{V}_G \cup [N + 1, N + E]$  to denote the indices of all the binding inequality constraints for  $(\hat{\mathbf{s}}^g, \hat{\mathbf{s}}^l)$  and let  $\mathcal{S}^c$  denote the indices of all the non-binding inequalities.<sup>6</sup> By construction, we have  $\text{rank}(\mathbf{T}_{\mathcal{S}}) = |\mathcal{S}| \leq$

<sup>5</sup>Hence, each  $\gamma \in \Gamma$  can also be regarded as a  $(2N_G + 2E)$ -by- $|\mathcal{S}|$  matrix. An explicit example of  $\gamma$  could be found in [41].

<sup>6</sup>In this section, the index of a constraint associated with generator  $i$  is  $i$  and the index of a constraint associated with branch  $i$  is  $i + N$ .

$\text{rank}(\mathbf{T}) = N - 1$ . There are two situations to discuss:  $|\mathcal{S}| = 0$  and  $1 \leq |\mathcal{S}| \leq N - 1$ .

In the first case, if  $|\mathcal{S}| = 0$ , let  $\rho_1$  be the matrix norm of

$$\mathbf{T}_{\mathcal{V}_G \cup [N+1, N+E]} \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{T}_{\mathcal{V}_L} \end{bmatrix}^\dagger$$

induced by the  $\ell_1$  vector norm. Let

$$\epsilon_1 = \min \left\{ \min_{\substack{i \in \mathcal{V}_G \\ \hat{\mathbf{s}}_i^g > \underline{\mathbf{s}}_i^g}} \hat{\mathbf{s}}_i^g - \underline{\mathbf{s}}_i^g, \min_{\substack{i \in \mathcal{V}_G \\ \bar{\mathbf{s}}_i^g > \hat{\mathbf{s}}_i^g}} \bar{\mathbf{s}}_i^g - \hat{\mathbf{s}}_i^g, \right. \\ \left. \min_{\substack{i \in \mathcal{E} \\ \hat{\mathbf{p}}_i > \underline{\mathbf{p}}_i}} \hat{\mathbf{p}}_i - \underline{\mathbf{p}}_i, \min_{\substack{i \in \mathcal{E} \\ \bar{\mathbf{p}}_i > \hat{\mathbf{p}}_i}} \bar{\mathbf{p}}_i - \hat{\mathbf{p}}_i \right\} \quad (24)$$

$$\epsilon_2 = \min_i \hat{\mathbf{s}}_i^l. \quad (25)$$

Here, we have used  $\min.$  as short hand for minimize. Now we can construct  $\mathbf{s}_{(n)}^l \equiv \hat{\mathbf{s}}^l$ , and

$$U(\mathbf{s}_{(n)}^l) = \left\{ \mathbf{s}^l \mid |\mathbf{s}^l - \mathbf{s}_{(n)}^l| < \frac{1}{2} \min \left\{ \frac{\epsilon_1}{N_L \rho_1}, \epsilon_2 \right\} \mathbf{1}_{N_L} \right\}.$$

It is trivial that  $\lim_{n \rightarrow \infty} \mathbf{s}_{(n)}^l = \hat{\mathbf{s}}^l$ . For any  $\mathbf{s}^l \in U(\mathbf{s}_{(n)}^l)$ , we have

$$\mathbf{s}^l > \mathbf{s}_{(n)}^l - \frac{1}{2} \epsilon_2 \mathbf{1}_{N_L} = \hat{\mathbf{s}}^l - \left( \frac{1}{2} \min_i \hat{\mathbf{s}}_i^l \right) \mathbf{1}_{N_L} > 0.$$

Further, it is easy to check that for

$$\boldsymbol{\theta} = \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{T}_{\mathcal{V}_L} \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ -\mathbf{s}^l \end{bmatrix}, \mathbf{s}^g = \mathbf{T}_{\mathcal{V}_G} \boldsymbol{\theta}, \quad (26)$$

(3b)-(3e) are satisfied. As a result,  $\mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$  and thus  $U(\hat{\mathbf{s}}^l) \subseteq \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ .

In the second case, we have  $1 \leq |\mathcal{S}| \leq N - 1$ , then define

$$\mathbf{T}'(\mathcal{R}) := \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{T}_{\mathcal{S}} \\ \mathbf{T}_{\mathcal{R}} \end{bmatrix}, \text{ for } \mathcal{R} \subseteq \mathcal{V}_L.$$

Let  $\mathcal{R}^*$  be a smallest  $\mathcal{R}$  (in cardinality) such that  $\text{rank}(\mathbf{T}'(\mathcal{R})) = \text{rank}(\mathbf{T}'(\mathcal{V}_L))$ . We further define

$$\mathbf{T}''(\mathcal{T}) := \begin{bmatrix} \mathbf{T}'(\mathcal{R}^*) \\ \mathbf{T}_{\mathcal{T}} \end{bmatrix}, \text{ for } \mathcal{T} \subseteq \mathcal{S}^c.$$

Let  $\mathcal{T}^*$  be a smallest  $\mathcal{T}$  such that  $\text{rank}(\mathbf{T}''(\mathcal{T})) = \text{rank}(\mathbf{T}''(\mathcal{S}^c))$ . There are two simple observations:

- All rows of the matrix  $\mathbf{T}''(\mathcal{T}^*)$  are independent.
- All rows of  $\mathbf{T}$  are in the row space of  $\mathbf{T}''(\mathcal{T}^*)$ .

Let  $\rho_2$  be the matrix norm of  $\mathbf{T}(\mathbf{T}''(\mathcal{T}^*))^\dagger$  induced by the  $\ell_1$  vector norm. Let  $\epsilon_1$  and  $\epsilon_2$  be the same as in (24) and (25), and we define the direction vector  $\mathbf{v} \in \mathbb{R}^{|\mathcal{S}|}$  as

$$\mathbf{v} := \text{sgn} \left( \mathbf{T}_{\mathcal{S}} \hat{\boldsymbol{\theta}} - \begin{bmatrix} \bar{\mathbf{s}}_i^g \\ \mathbf{o}_{\mathcal{V}_L \times 1}^g \\ \bar{\mathbf{p}} \end{bmatrix}_{\mathcal{S}} \right) + \text{sgn} \left( \mathbf{T}_{\mathcal{S}} \hat{\boldsymbol{\theta}} - \begin{bmatrix} \bar{\mathbf{s}}_i^g \\ \mathbf{o}_{\mathcal{V}_L \times 1}^g \\ \underline{\mathbf{p}} \end{bmatrix}_{\mathcal{S}} \right)$$

where  $\text{sgn}$  applies the sign function to each coordinate of the vector. We then construct

$$\mathbf{s}_{(n)}^l := \mathbf{T}_{\mathcal{V}_L} (\mathbf{T}''(\mathcal{T}^*))^\dagger \begin{bmatrix} 0 \\ \mathbf{T}_{\mathcal{S}} \hat{\boldsymbol{\theta}} - \frac{\min\{\epsilon_1, \epsilon_2\}}{2nN\rho_2} \mathbf{v} \\ \mathbf{T}_{\mathcal{R}^*} \hat{\boldsymbol{\theta}} \\ \mathbf{T}_{\mathcal{T}^*} \hat{\boldsymbol{\theta}} \end{bmatrix}.$$

It is easy to check  $\lim_{n \rightarrow \infty} \mathbf{s}_{(n)}^l \rightarrow \hat{\mathbf{s}}^l$  and  $\mathbf{s}_{(n)}^l > 0$ . We then construct the associated  $\boldsymbol{\theta}_{(n)}$ ,  $\mathbf{s}_{(n)}^g$  and  $\mathbf{p}_{(n)}$  as

$$\boldsymbol{\theta}_{(n)} := (\mathbf{T}''(\mathcal{T}^*))^\dagger \begin{bmatrix} 0 \\ \mathbf{T}_{\mathcal{S}} \hat{\boldsymbol{\theta}} - \frac{\min\{\epsilon_1, \epsilon_2\}}{2nN\rho_2} \mathbf{v} \\ \mathbf{T}_{\mathcal{R}^*} \hat{\boldsymbol{\theta}} \\ \mathbf{T}_{\mathcal{T}^*} \hat{\boldsymbol{\theta}} \end{bmatrix},$$

$$\mathbf{s}_{(n)}^g := \mathbf{T}_{\mathcal{V}_G} \boldsymbol{\theta}_{(n)}, \mathbf{p}_{(n)} := \mathbf{T}_{[N+1, N+E]} \boldsymbol{\theta}_{(n)}.$$

This construction guarantees that  $\mathbf{s}_{(n)}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$  and  $(\mathbf{s}_{(n)}^g, \mathbf{s}_{(n)}^l)$  does not introduce new binding constraints. Similar to the first case, there is always a neighborhood  $U(\mathbf{s}_{(n)}^l) \subseteq \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ . This completes the proof for  $\text{clos}(\text{int}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))) = \text{clos}(\Omega_{\mathbf{s}^l}(\boldsymbol{\xi}))$ .

Next, we will show  $\tilde{\Omega}_{\boldsymbol{\xi}}(\boldsymbol{\xi})$  is dense in  $\Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ . In fact,  $\forall \boldsymbol{\xi} \in \tilde{\Omega}_{\boldsymbol{\xi}}$ , if for some  $\mathbf{s}^l \in \Omega_{\mathbf{s}^l}(\boldsymbol{\xi})$ , the optimal solution to (3) has  $\geq N_G$  tight inequality constraints, then we use  $\mathcal{S} \subseteq [N + E] \setminus \mathcal{V}_L$ ,  $|\mathcal{S}| = N_G$  again to denote the indices of any  $N_G$  tight inequality constraints. As those  $N_G$  inequality constraints are tight, there must exist  $\gamma \in \Gamma$ , as defined in (22), such that  $\gamma^\top \boldsymbol{\xi} = \mathbf{T}_{\mathcal{S}} \boldsymbol{\theta}^*$  for the optimal  $\boldsymbol{\theta}^* \in \mathbb{R}^N$ . According to (23),  $\text{rank}(\mathbf{T}_{\mathcal{S}})$  must be exactly  $N_G$ . We now have

$$\gamma^\top \boldsymbol{\xi} = \mathbf{T}_{\mathcal{S}} \boldsymbol{\theta}^*, \quad -\mathbf{s}^l = \mathbf{T}_{\mathcal{V}_L} \boldsymbol{\theta}^*. \quad (27)$$

For each  $\gamma \in \Gamma$ , as  $\text{rank}(\mathbf{T}_{\mathcal{S}}) = N_G$  but  $\text{rank}(\mathbf{T}) = N - 1$ , the set  $\{\mathbf{s}^l \mid \exists \boldsymbol{\theta}^*, (27) \text{ holds}\}$  is a strict subspace in  $\mathbb{R}^{N_L}$  and thereby nowhere dense in  $\Omega_{\mathbf{s}^l}$  according to Proposition 4 and Lemma 5. As the result, we have

$$\tilde{\Omega}_{\mathbf{s}^l} \supseteq \Omega_{\mathbf{s}^l} \setminus \bigcup_{\substack{|\mathcal{S}|=N_G \\ \mathcal{S} \subseteq [N+E] \setminus \mathcal{V}_L}} \bigcup_{\gamma \in \Gamma} \{\mathbf{s}^l \mid \exists \boldsymbol{\theta}^*, (27) \text{ holds for } \gamma\}$$

must be dense in  $\Omega_{\mathbf{s}^l}$ .

## 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] 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.
- [3] W. H. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on power apparatus and systems*, no. 10, pp. 1866–1876, 1968.
- [4] S. Frank, I. Steponavice, and S. Rebennack, "Optimal power flow: a bibliographic survey i," *Energy Systems*, vol. 3, no. 3, pp. 221–258, 2012.
- [5] S. Frank and S. Rebennack, "An introduction to optimal power flow: Theory, formulation, and examples," *IIE Transactions*, vol. 48, no. 12, pp. 1172–1197, 2016.
- [6] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.
- [7] J. Sun and L. Tesfatsion, "DC optimal power flow formulation and solution using QuadProgJ," *Iowa State University Digital Repository*, 2010.
- [8] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. John Wiley & Sons, 2012.
- [9] G. Pataki and L. Tunçel, "On the generic properties of convex optimization problems in conic form," *Mathematical Programming*, vol. 89, no. 3, pp. 449–457, 2001.
- [10] M. Dür, B. Jargalsaikhan, and G. Still, "Genericity results in linear conic programming—a tour d’horizon," *Mathematics of operations research*, vol. 42, no. 1, pp. 77–94, 2016.
- [11] D. T. Luc and P. H. Dien, "Differentiable selection of optimal solutions in parametric linear programming," *Proceedings of the American Mathematical Society*, vol. 125, no. 3, pp. 883–892, 1997.

- [12] A. V. Fiacco, "Introduction to sensitivity and stability analysis in nonlinear programming," 1983.
- [13] Y. Wang and R. D. Monteiro, "Nondegeneracy of polyhedra and linear programs," *Computational Optimization and Applications*, vol. 7, no. 2, pp. 221–237, 1997.
- [14] A. Agrawal, S. Barratt, S. Boyd, E. Busseti, and W. M. Moursi, "Differentiating through a conic program," *arXiv preprint arXiv:1904.09043*, 2019.
- [15] B. Amos and J. Z. Kolter, "Optnet: Differentiable optimization as a layer in neural networks," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, 2017, pp. 136–145.
- [16] F. Zhou, J. Anderson, and S. H. Low, "Differential privacy of aggregated DC optimal power flow data," *Accepted to the 2019 American Control Conference*, *arXiv preprint arXiv:1903.11237*, 2019.
- [17] X. Geng and L. Xie, "Learning the LMP-load coupling from data: A support vector machine based approach," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1127–1138, 2016.
- [18] Y. Ji, R. J. Thomas, and L. Tong, "Probabilistic forecasting of real-time LMP and network congestion," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 831–841, 2017.
- [19] S. Misra, L. Roald, and Y. Ng, "Learning for constrained optimization: Identifying optimal active constraint sets," *arXiv preprint arXiv:1802.09639*, 2018.
- [20] Y. Ng, S. Misra, L. A. Roald, and S. Backhaus, "Statistical learning for dc optimal power flow," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–7.
- [21] L. Roald and D. K. Molzahn, "Implied constraint satisfaction in power system optimization: The impacts of load variations," *arXiv preprint arXiv:1904.01757*, 2019.
- [22] T. Kim, S. J. Wright, D. Bienstock, and S. Harnett, "Analyzing vulnerability of power systems with continuous optimization formulations," *IEEE Transactions on Network Science and Engineering*, vol. 3, no. 3, pp. 132–146, 2016.
- [23] J. Anderson, F. Zhou, and S. H. Low, "Disaggregation for networked power systems," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–7.
- [24] Y. Tang, E. Dall'Anese, A. Bernstein, and S. Low, "Running primal-dual gradient method for time-varying nonconvex problems," *arXiv preprint arXiv:1812.00613*, 2018.
- [25] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [26] 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, 2014.
- [27] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009, vol. 28.
- [28] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.
- [29] J. M. Mulvey, R. J. Vanderbei, and A. S. Zenios, "Robust optimization of large-scale systems," *Operations research*, vol. 43, no. 2, pp. 264–281, 1995.
- [30] D. P. Heyman and M. J. Sobel, *Stochastic models in operations research: stochastic optimization*. Courier Corporation, 2004, vol. 2.
- [31] P. Kall and S. W. Wallace, *Stochastic programming*. Springer, 1994.
- [32] B. L. Miller and H. M. Wagner, "Chance constrained programming with joint constraints," *Operations Research*, vol. 13, no. 6, pp. 930–945, 1965.
- [33] A. J. Hoffman, "On approximate solutions of systems of linear inequalities," in *Selected Papers Of Alan J Hoffman: With Commentary*. World Scientific, 2003, pp. 174–176.
- [34] S. M. Robinson, "Stability theory for systems of inequalities. Part I: Linear systems," *SIAM Journal on Numerical Analysis*, vol. 12, no. 5, pp. 754–769, 1975.
- [35] O. L. Mangasarian and T.-H. Shiau, "Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems," *SIAM Journal on Control and Optimization*, vol. 25, no. 3, pp. 583–595, 1987.
- [36] P. R. Gribik, D. Shirmohammadi, S. Hao, and C. L. Thomas, "Optimal power flow sensitivity analysis," *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 969–976, 1990.
- [37] C. Yu, "Sensitivity analysis of multi-area optimum power flow solutions," *Electric Power Systems Research*, vol. 58, no. 3, pp. 149–155, 2001.
- [38] X.-S. Zhang and D.-G. Liu, "A note on the continuity of solutions of parametric linear programs," *Mathematical Programming*, vol. 47, no. 1-3, pp. 143–153, 1990.
- [39] J. Anderson, F. Zhou, and S. H. Low, "Worst-case sensitivity of dc optimal power flow problems," in *To appear in Proc. of the 2020 American Control Conference*. IEEE, 2020.
- [40] A. V. Fiacco, "Sensitivity analysis for nonlinear programming using penalty methods," *Mathematical programming*, vol. 10, no. 1, pp. 287–311, 1976.
- [41] F. Zhou, J. Anderson, and S. H. Low, "The optimal power flow operator: Theory and computation," *arXiv preprint arXiv:1907.02219*, 2019.
- [42] Y. Ye, M. J. Todd, and S. Mizuno, "An  $O(\sqrt{n})$ -iteration homogeneous and self-dual linear programming algorithm," *Mathematics of Operations Research*, vol. 19, no. 1, pp. 53–67, 1994.
- [43] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *Journal of Optimization Theory and Applications*, vol. 169, no. 3, pp. 1042–1068, 2016.
- [44] R. T. Rockafellar, *Convex analysis*. Princeton university press, 2015.
- [45] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2011.

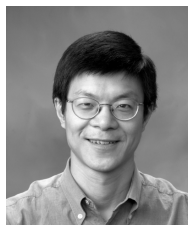


**Fengyu Zhou** received the B.E. degree in electronics engineering from Tsinghua University, Beijing, China, in 2016 and the M.S. degree in electrical engineering from Caltech, Pasadena, CA, USA, in 2018. He is currently pursuing the Ph.D. degree in the Department of Electrical Engineering at Caltech.

His research interests are on the modeling, optimization and control of cyber-physical systems.



**James Anderson** is an assistant professor in the Department of Electrical Engineering and member of the Data Science Institute at Columbia University in the City of New York. He was previously a senior postdoctoral scholar in the Department of Computing & Mathematical Sciences at Caltech and a Junior Research Fellow in Engineering Science at St John's College in the University of Oxford. He received his DPhil (PhD) from the University of Oxford and the MSc and BSc degrees from the University of Reading.



**Steven H. Low** is the Gilloon Professor of Computing & Mathematical Sciences and Electrical Engineering at Caltech and honorary professor of Melbourne University, Australia. He is a co-recipient of IEEE best paper awards and an IEEE Fellow. His research on networks has been accelerating more than 1TB of Internet traffic every second since 2014. His research on smart grid is providing large scale electric vehicle charging to workplaces, from K-12 and universities to Fortune Global 500 companies. He received his B.S. from Cornell and PhD from

Berkeley.