

Uncertainty-aware Three-phase Optimal Power Flow based on Data-driven Convexification

Qifeng Li, *Senior Member, IEEE*

Abstract—This paper presents a novel optimization framework of formulating the three-phase optimal power flow that involves uncertainty. The proposed uncertainty-aware optimization (UaO) framework is: 1) a deterministic framework that is less complex than the existing optimization frameworks involving uncertainty, and 2) convex such that it admits polynomial-time algorithms and mature distributed optimization methods. To construct this UaO framework, a methodology of learning-aided uncertainty-aware modeling, with prediction errors of stochastic variables as the measurement of uncertainty, and a theory of data-driven convexification are proposed. Theoretically, the UaO framework is applicable for modeling general optimization problems under uncertainty.

Index Terms—Convex relaxation, data-driven, optimization under uncertainty, three-phase power flow.

I. INTRODUCTION

OPTIMIZATION technologies have been widely used in many decision-making processes in the operation, control, and planning of power systems, such as optimal power flow (OPF). However, the increasing uncertainty introduced by distributed energy resources (DER) makes it extremely hard for operators to make accurate optimal decisions ahead of real time. There mainly exist three types of frameworks for modeling power system optimization problems that involve uncertainty: 1) stochastic framework, 2) robust framework, and 3) chance-constrained framework [1]. Unfortunately, these frameworks are rather computationally expensive for large-scale, highly-nonconvex problems. As a result, a large portion of existing works investigate proper assumptions to simplify these frameworks for power system applications [2]. In contrast, based on regression analysis [3], this paper develops a novel uncertainty-aware optimization (UaO) framework using a new measurement of uncertainty that considers the prediction errors of stochastic variables (see Section III for more details).

Convex optimization [4] has applications in a broad range of disciplines including power system engineering, mainly because: 1) many classes of convex optimization problems are computationally tractable as they admit polynomial-time algorithms; and, 2) it plays a fundamental role in the theories of both distributed optimization and bi-level optimization. The general idea is to relax the nonconvex objective functions and constraints into convex ones. However, a solution of the resulting convex problem may be infeasible to the original

nonconvex problem due to the nature of relaxations, which now becomes one of the bottlenecks of this technology. In order to break these bottlenecks, this paper proposes a data-driven approach to construct convex relaxations with stronger tightness and lower complexity (see Subsection II-C for more details). The resulting convex relaxation is applied to convexify the developed UaO framework. The paper demonstrates the UaO framework on a three-phase optimal power flow (3 ϕ OPF) problem with uncertainty introduced by DERs and uncontrollable loads. The 3 ϕ OPF is balanced for transmission networks while unbalanced for distribution networks. It is worth noting that, theoretically, the proposed methods can be applied to general optimization problems under uncertainty.

II. THEORY OF DATA-DRIVEN CONVEXIFICATION

A. Three-Phase Power Flow Equations

In an OPF problem, the objective function is generally convex or linear. Thus, we focus on the main nonconvex constraints, i.e. the alternating current power flow (ACPF) equations, which are also considered as the mathematical model of power networks. Let \mathcal{N} , \mathcal{E} and Φ denote the sets of buses, lines, and phases respectively. For each $i, j \in \mathcal{N}$, $ij \in \mathcal{E}$ and $\phi, \phi' \in \Phi$, the formulation of three-phase ACPF (3 ϕ ACPF) equations [5] is given as

$$e_i^\phi \sum_{j \in \mathcal{N}} \sum_{\phi' \in \Phi} (G_{ij}^{\phi\phi'} e_j^{\phi'} - B_{ij}^{\phi\phi'} f_j^{\phi'}) + f_i^\phi \sum_{j \in \mathcal{N}} \sum_{\phi' \in \Phi} (B_{ij}^{\phi\phi'} e_j^{\phi'} + G_{ij}^{\phi\phi'} f_j^{\phi'}) = p_{D,i}^\phi + p_{S,i}^\phi \quad (1a)$$

$$f_i^\phi \sum_{j \in \mathcal{N}} \sum_{\phi' \in \Phi} (G_{ij}^{\phi\phi'} e_j^{\phi'} - B_{ij}^{\phi\phi'} f_j^{\phi'}) - e_i^\phi \sum_{j \in \mathcal{N}} \sum_{\phi' \in \Phi} (B_{ij}^{\phi\phi'} e_j^{\phi'} + G_{ij}^{\phi\phi'} f_j^{\phi'}) = q_{D,i}^\phi + q_{S,i}^\phi \quad (1b)$$

$$e_i^\phi \sum_{\phi' \in \Phi} [G_{ij}^{\phi\phi'} (e_j^{\phi'} - e_i^{\phi'}) - B_{ij}^{\phi\phi'} (f_j^{\phi'} - f_i^{\phi'})] + f_i^\phi \sum_{\phi' \in \Phi} [B_{ij}^{\phi\phi'} (e_j^{\phi'} - e_i^{\phi'}) + G_{ij}^{\phi\phi'} (f_j^{\phi'} - f_i^{\phi'})] = p_{ij}^\phi \quad (1c)$$

$$f_i^\phi \sum_{\phi' \in \Phi} [G_{ij}^{\phi\phi'} (e_j^{\phi'} - e_i^{\phi'}) - B_{ij}^{\phi\phi'} (f_j^{\phi'} - f_i^{\phi'})] - e_i^\phi \sum_{\phi' \in \Phi} [B_{ij}^{\phi\phi'} (e_j^{\phi'} - e_i^{\phi'}) + G_{ij}^{\phi\phi'} (f_j^{\phi'} - f_i^{\phi'})] = q_{ij}^\phi \quad (1d)$$

$$(e_i^\phi)^2 + (f_i^\phi)^2 = v_i, \quad (1e)$$

where e and f are the real and imaginary parts of voltage, while $p_{S,i}^\phi$ ($q_{S,i}^\phi$) and $p_{D,i}^\phi$ ($q_{D,i}^\phi$) denote the stochastic and

This work is supported by U.S. National Science Foundation under grant #1808988.

The author is with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, 32816 USA e-mail: qifeng.li@ucf.edu.

deterministic components of active (reactive) power injections, respectively, at each bus. Let $|\mathcal{M}| = (3 \times |\mathcal{N}| + 2 \times |\mathcal{E}|) \times |\Phi|$, equations (1) can be compactly formulated as

$$g_i(x) = x^T A_i x = y_i = z_i + u_i, \quad (i \in \mathcal{M}) \quad (2)$$

where the right-hand-side (RHS) vector $z = [\dots, p_{D,i}^\phi, q_{D,i}^\phi, \dots, p_{ij}^\phi, q_{ij}^\phi, \dots, v_i, \dots]^T$ denotes the RHS deterministic quantities while vector $u = [\dots, p_{S,i}^\phi, q_{S,i}^\phi, \dots, 0, \dots]^T$ denotes the RHS stochastic quantities. The coefficient matrices A_i in (1a)-(1d) are indefinite such that the quadratic functions in (2) are nonconvex (except the ones corresponding to (1e)). Further define a set $\Omega = \{(x, y) \mid p_i^\phi \leq p_{D,i}^\phi + p_{S,i}^\phi \leq \bar{p}_i^\phi, q_i^\phi \leq q_{D,i}^\phi + q_{S,i}^\phi \leq \bar{q}_i^\phi, (p_{ij}^\phi)^2 + (q_{ij}^\phi)^2 \leq \bar{S}_{ij}, \text{ and } \underline{v}_i \leq v_i \leq \bar{v}_i, \forall i \in \mathcal{N} \text{ and } \phi \in \Phi\}$, then the feasible set of 3ϕ ACPF (1) is $\Psi = \{(x, y) \in \Omega \mid (2), \forall i \in \mathcal{M}\}$. Note that Ω is convex while Ψ is not. Moreover, the three-phase *DistFlow* model [6] of radial networks is also a nonconvex quadratic system that can be represented in the form of (2). That means the proposed methods can be directly applied to *DistFlow*-based 3ϕ OPF.

B. State-of-the-Art Convex Relaxations for AC Power Flow

The ACPF can be formulated as quadratic models, like model (1), or trigonometric models. Most of the existing convex relaxations, e.g. the semidefinite (SD) [7] and conic relaxations [8], are based on the quadratic ACPF models. Mature convex relaxations, such as the conic [8] and convex hull [9] relaxations, have been developed for the simplest scenario, i.e. the balanced ACPF of radial networks which can be formulated as the *DistFlow* model [6]. However, despite the satisfactory performance of these convex relaxations, radial (distribution) networks are barely balanced. For the rest scenarios, such as balanced ACPF in meshed networks and unbalanced ACPF (i.e. 3ϕ ACPF) in radial networks, the most suitable one among the existing convex relaxations is the SD relaxation. However, the limitations and inexactness of the SD relaxation for ACPF have been widely observed [10]. Moreover, the SD relaxation of large-scale OPF problems is computationally expensive. To mitigate the issues associated with the existing convex relaxations for ACPF, this paper explores to utilize the emerging machine learning technology in the process of constructing convex relaxations.

C. Data-Driven Convex Relaxation

In this subsection, a methodology of data-driven convex relaxation (DDCR) is established and applied to construct a tight convex quadratic relaxation of 3ϕ ACPF model (2) (i.e. (1)). For simplicity, we start from a deterministic case, namely $u = 0$. Let \mathcal{D} denotes a training data set where the k th data point $D^{(k)} = (x^{(k)}, y^{(k)})$ denotes the real-time measurement of a historical operating point, we have $\mathcal{D} \subset \Psi$. The following regression algorithm is proposed to train \mathcal{D} to obtain a positive semi-definite (PSD) matrix P_i and a complementary vector B_i and scalar c_i for each quadratic equations in (1a)-(1d):

$$\min_{P_i, B_i, c_i} \frac{1}{|\mathcal{D}|} \sum_k r^{(k)} \quad (3a)$$

$$\text{s.t. } (x^{(k)})^T P_i x^{(k)} + B_i^T x^{(k)} + c_i - y_i^{(k)} = m^{(k)} \leq 0 \quad (3b)$$

$$P_i \succeq 0 \quad (3c)$$

$$\begin{bmatrix} 1 & m^{(k)} \\ m^{(k)} & r^{(k)} \end{bmatrix} \succeq 0, \quad (3d)$$

where $k = 1, \dots, |\mathcal{D}|$. Auxiliary variables $r^{(k)}$ are introduced for the purpose of formulating the regression model (3) as a standard semidefinite programming (SDP) problem. The dimensions of P_i , B_i , and c_i are consistent with the dimensions of the corresponding quadratic equations in (1a)-(1d). Note that the A matrix in (1e) is already PSD. Therefore, we don't need to train a P for (1e). The optimization model (3) is a standard SDP problem which can be effectively and globally solved by mature solvers like MOSEK, GUROBI, and CPLEX.

Define a quadratic convex set:

$$\Theta = \{(x, y) \in \Omega \mid x^T P_i x + B_i^T x + c_i \leq y_i, \forall i \in \mathcal{M}\},$$

we have the following theorem.

Theorem of Data-driven Convex Relaxation. *The set Θ is a convex relaxation of the feasible set Ψ of the original three-phase AC power flow (2) if:*

- the PSD matrices P_i , vectors B_i , and scalars c_i ($i \in \mathcal{M}$) are obtained by training \mathcal{D} using the regression algorithm (3),*
- \mathcal{D} contains all extreme points¹ of Ψ .*

Proof: Constraint (3b) guarantees that each $D^{(k)} \in \mathcal{D}$ satisfies

$$(x^{(k)})^T P_i x^{(k)} + B_i^T x^{(k)} + c_i \leq y_i^{(k)}, \quad i \in \mathcal{M}$$

which implies $\mathcal{D} \subset \Theta$. Therefore, Θ is a convex quadratic relaxation of \mathcal{D} since P_i ($i \in \mathcal{M}$) are PSD.

All extreme points of a feasible set are linearly independent according to the definition [11]. Suppose ψ is an arbitrary point in Ψ , there must exist a vector of extreme points $X = [\theta_1, \theta_2, \dots, \theta_l]^T$ of Ψ and a vector of multipliers $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$ that satisfy

$$\psi = \alpha^T X,$$

where $0 \leq \alpha_i \leq 1$ ($i = 1, \dots, l$), $\sum_i \alpha_i = 1$, and l equals to the dimension of the (x, y) -space. Since all $\theta_i \in \mathcal{D} \subset \Theta$ according to condition b), then $\psi \in \Theta$ due to the convexity of Θ . Therefore, $\Psi \subset \Theta$ as ψ is an arbitrary point in Ψ , which means Θ is convex relaxation of Ψ . \square

Remark 1. Condition b) in the theorem of data-driven convex relaxation is not easy to strictly satisfy. However, under the concept of *Big Data*, it is reasonable to assume that the data set \mathcal{D} is big enough to represent the original feasible set Ψ , which implies Θ is highly close to a strictly convex relaxation of Ψ . Moreover, regression (3) is a convex optimization problem that can be globally solved, which implies that Θ is the *tightest* quadratic convex relaxation of Ψ .

III. LEARNING-AIDED UNCERTAINTY-AWARE MODELING

This section introduces the idea and methodology of the proposed concept of learning-aided uncertainty-aware modeling (L-UaM) starting from the original 3ϕ ACPF model (2). For simplicity of the explanation, we ignore the controllable power injection z which is deterministic and does not impact the analysis of uncertainty, namely $y = u$.

¹An extreme point of a convex set is a point in this set that does not lie in any open line segment joining two points of this set [11]. We use this definition to define an extreme point of nonconvex sets.

A. Idea of L-UaM

Suppose \hat{u} is the real-time measurement of u at time t , its system response is \hat{x} as in equation (4a). Generally, the OPF is solved ahead of time t (e.g. five minutes \sim a day ahead) based on an ahead-of-real-time forecast \tilde{u} as in (4b). One can consider that the uncertainty originates from the prediction error $\|\tilde{u} - \hat{u}\|$, since the resulting model output error $\|\tilde{x} - \hat{x}\|$ may lead to failures in OPF. The objective of this research is to construct an uncertainty-aware model h that can approximate the actual state \hat{x} at time t , which is a system response to real-time parameter \hat{u} , with only an ahead-of-real-time forecast \tilde{u} as in (4c). For this purpose, a concept of L-UaM is proposed, of which the methodology is detailed in next subsection.

$$g_i(\hat{x}) = \hat{u}_i \quad (4a)$$

$$g_i(\tilde{x}) = \tilde{u}_i, \quad (i \in \mathcal{M}) \quad (4b)$$

$$h_i(\hat{x}) \approx \tilde{u}_i. \quad (4c)$$

In the stochastic and chance-constrained frameworks, uncertainty is measured by probability distributions, while it is captured by a deterministic sets under the robust framework [1], [2]. According to the above analysis, a ‘‘good prediction’’ and a ‘‘bad prediction’’ may have totally different impacts on the operation, control, and planning of power systems. However, the information of forecast is not considered in the measurement of uncertainty in these optimization frameworks, which increases the conservativeness. In this research, we use the prediction error $\|\tilde{u} - \hat{u}\|$ as the new measurement of uncertainty which will be incorporated into the following machine learning process. Researchers from the field of renewable generation and load forecast attempts to reduce the prediction errors $\|\tilde{u} - \hat{u}\|$, while this research focuses on reducing the model output errors $\|\tilde{x} - \hat{x}\|$ in the process of system modeling given the information of \tilde{u} .

B. Methodology of L-UaM

In order to learn an *uncertainty-aware model* (UaM) of 3ϕ ACPF, we first re-designed the historical training data set \mathcal{D} to include the information of uncertainty. To avoid confusion, we denote the new data set as $\tilde{\mathcal{D}}$ of which the k th data point $\tilde{D}^{(k)} = (\hat{u}^{(k)}, \hat{x}^{(k)}, \tilde{u}^{(k)})$. Then, the following regression algorithm is designed to learn a hypothesis function h :

$$\min_{h_i} \left\{ \frac{1}{|\tilde{\mathcal{D}}|} \sum_{\tilde{D}} \left[h_i(\hat{x}^{(k)}) - \tilde{u}_i^{(k)} \right]^2 \right\}, \quad (i \in \mathcal{M}) \quad (5)$$

where the form of $h(\cdot)$ is pre-determined, and the cost function represents the mean of squared Euclidean distance between data points and the hypothesis function. The deterministic model $h(x) = u$ learned by regression (5) is defined as a UaM of 3ϕ ACPF (2) (Note that the deterministic power component z is ignored for simplicity in explanation. In the UaM training process of a actual power system, the term z should be included). The UaM h can capture the uncertainty of u since it represents a mapping between \tilde{u} and \hat{x} . Namely, it can produce a close approximation to the actual system state \hat{x} at a future time t , with only an ahead-of- t forecast values \tilde{u} of the stochastic parameters u , as illustrated in equation (4c).

Remark 2. Model h resulting from (5) is deterministic with ability of capturing the uncertainty of u since it is obtained by training data set $\tilde{\mathcal{D}}$ that contains the historical forecast error $\|\tilde{u} - \hat{u}\|$ as the uncertainty measurement of u . Note that the forms of both cost function and the hypothesis function h in regression (5) are not unique, which can be customized as needed. For instance, to combine the L-UaM and DDCR processes, we will choose a convex quadratic hypothesis function as given in next section.

IV. UNCERTAINTY-AWARE CONVEX OPTIMIZATION

A. Uncertainty-aware Convex Modeling

By incorporating the DDCR feature into the L-UaM process, this subsection aims at training a data-driven model of 3ϕ ACPF that has two key features: convex and aware of uncertainty. For this purpose, like what we did in Section II, a hypothesis function h of the following convex quadratic form is chosen:

$$h_i(x) = x^T P_i x + B_i^T x + c_i. \quad (i \in \mathcal{M}) \quad (6)$$

Then, regression algorithms (3) and (5) are combined to produce the following regression algorithm

$$\min_{P, B, c} (3a) \quad \text{s.t.} \quad (3c), (3d) \quad \text{and} \quad (7a)$$

$$h_i(\hat{x}^{(k)}) - \hat{u}_i^{(k)} \leq 0 \quad (7b)$$

$$h_i(\hat{x}^{(k)}) - \tilde{u}_i^{(k)} = m^{(k)}, \quad (7c)$$

to train the new data set $\tilde{\mathcal{D}}$. Constraint (7c) together with the objective function (3a) implies that h is an optimal mapping between real-time system response \hat{x} and ahead-of-real-time forecast of stochastic parameter \tilde{u} . With the P_i , B_i and c_i inferred by regression (7), the epigraph (8) of h :

$$\mathbf{Epi}(h) = \{(x, y) \in \Omega \mid h_i(x) \leq u_i, \forall i \in \mathcal{M}\} \quad (8)$$

is a convex UaM of 3ϕ ACPF (2). Constraint (7b) indicates that convex set $\mathbf{Epi}(h)$ encloses all data points in $\tilde{\mathcal{D}}$ which represent real-time operating points. It is worth pointing out that UaM h uses the ahead-of-real-time forecast \tilde{u} as input, which matches the fact that OPF is generally solved ahead-of-real-time.

B. Uncertainty-aware 3ϕ OPF

With the typical objective that minimizes generation costs, the uncertainty-aware convex 3ϕ OPF can be compactly formulated as

$$\min_z \left\{ \sum_{i \in \mathcal{G}} \sum_{\phi \in \Phi} C(p_{D,i}^\phi) : (x, z + \tilde{u}) \in \mathbf{Epi}(h) \right\}, \quad (9)$$

where \mathcal{G} denotes the generator set, and P_i , B_i , c_i in $\mathbf{Epi}(h)$ are inferred by training $\tilde{\mathcal{D}}$ using regression (7). A typical objective of OPF in distribution systems is to minimize the active power from transmission grids. It is worth noting that (9) is a deterministic optimization problem which is much less-complex than the existing robust, stochastic, and chance-constrained optimization frameworks.

V. NUMERICAL EXPERIMENT

1) *Experiment design*: This section presents a numerical experiment on three scenarios of two balanced networks, i.e. a 5-bus and a 57-bus systems, and one unbalanced network, i.e. IEEE 34-bus distribution feeder. The three scenarios compared in this experiment are: 1) the original ACOPF with perfect predictions of stochastic power injections u ; 2) the original ACOPF with inaccurate predictions \tilde{u} ; and 3) the proposed convex UaO framework with the same inaccurate predictions \tilde{u} . Scenario 1 represents an ideal case which is considered as the reference. Scenario 2 simulates the current practice of solving OPF while scenario 3 is the proposed optimization framework for solving uncertainty-involved OPF. It is worth noting that this is a preliminary study based on simulated data since real-world data is currently not available. We will perform a more comprehensive study using real-world data in our future research.

2) *Training of convex uncertainty-aware 3ϕ ACPF models for the third scenario*: In the first step of generating the training data sets \tilde{D} for each test system, a set of power injection profiles $\hat{y}^{(k)} = \hat{z}^{(k)} + \hat{u}^{(k)}$ ($k = 1, \dots, 50000$) are randomly produced. For each $\hat{u}^{(k)}$, there is a one-hour ahead forecast $\tilde{u}^{(k)}$ which is also randomly generated assuming a maximum forecast error of $\pm 30\%$, namely $|\tilde{u}^{(k)} - \hat{u}^{(k)}| \leq 0.3|\hat{u}^{(k)}|$. Then, voltage profile $\hat{x}^{(k)}$ for each load profile of each system is obtained by solving PF (1). A convex uncertainty-aware 3ϕ ACPF model is obtained for each of the three systems by training the corresponding data set \tilde{D} using regression (7).

3) *Experiment results*: For each test system, the accuracy of scenario 3 is compared with scenario 2 considering scenario 1 as the reference on 50 load cases. Let $C_k^{(i)}$ denotes the optimal cost of the k th scenario in the i th load case, where $k = \{1, 2, 3\}$ and $i = 1, \dots, 50$. The following average errors of objective values are calculated for scenarios 2 and 3 respectively

$$E_k = \frac{1}{I} \sum_{i=1}^I \frac{|C_k^{(i)} - C_1^{(i)}|}{C_1^{(i)}},$$

where $k = \{2, 3\}$ and $I = 50$. The simulation results are tabulated in Table I.

TABLE I
AVERAGE ERRORS OF OPTIMAL COSTS

	5-bus	34-bus	57-bus
E_2	20.21%	25.54%	21.43%
E_3	3.72%	6.71%	1.66%

4) *Analysis*: It can be observed from the above table that the proposed convex uncertainty-aware optimization framework (i.e. scenario 3) outperforms the original ACOPF (i.e. scenario 2) in terms of approximating the ideal case (i.e. scenario 1). The large errors of scenario 2 comes from the inaccurate forecast \tilde{u} of the stochastic parameter u . Although the convex UaO framework also uses the inaccurate forecast \tilde{u} as input, it can provide better solutions since the forecast errors are taken into account in the training process of the 3ϕ ACPF constraints. The convex UaO framework is still not able to produce a strictly accurate solution. Nevertheless, it can be

improved by training a larger, better data set according to its learning-based nature. Another important observation is that the proposed approaches are applicable to both balanced and unbalanced networks, and both radial and meshed networks.

Approximating a nonconvex optimization problem with its convex relaxations can significantly improve the computational efficiency while tightness is another important performance of a convex relaxation which is used to measure the accuracy of this approximation. The concept of convex UaO models is first proposed in this paper such that there not exist counterparts for comparison. Therefore, an independent numerical experiment in the deterministic environment is needed to evaluate the tightness of the proposed data-driven convex relaxation being compared with exiting deterministic convex relaxations. Due to page limit, we will seek to present such an evaluation in future publications based on real-world data.

VI. CONCLUSION AND FUTURE WORK

This paper develops a uncertainty-aware optimization (UaO) framework for modeling the three-phase optimal power flow (3ϕ OPF) problem under uncertainty. The UaO is a convex, deterministic optimization framework that can capture uncertainty. A preliminary numerical experiment based on simulated data shows that the proposed framework has strong capability of mitigate the impacts of uncertainty on 3ϕ OPF. In our future research, we will explore advanced machine learning technologies, such as ensemble learning [3], to improve the efficiency of UaO framework. Moreover, we will apply the UaO framework to model other power system optimization problems other than OPF problems.

REFERENCES

- [1] Alireza Soroudi and Turaj Amraee. Decision making under uncertainty in energy systems: State of the art. *Renewable and Sustainable Energy Reviews*, 28:376–384, 2013.
- [2] Morteza Aien, Ali Hajebrahimi, and Mahmud Fotuhi-Firuzabad. A comprehensive review on uncertainty modeling techniques in power system studies. *Renewable and Sustainable Energy Reviews*, 57:1077–1089, 2016.
- [3] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [4] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [5] Ren Hu, Qifeng Li, and Feng Qiu. Ensemble learning based convex approximation of three-phase power flow. *IEEE Transactions on Power Systems*, 1st revision under review(A pre-print available at arXiv:1909.05748), 2019.
- [6] Mesut E Baran and Felix F Wu. Optimal capacitor placement on radial distribution systems. *IEEE Transactions on power Delivery*, 4(1):725–734, 1989.
- [7] Xiaoqing Bai, Hua Wei, Katsuki Fujisawa, and Yong Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6-7):383–392, 2008.
- [8] Rabih A Jabr. Radial distribution load flow using conic programming. *IEEE transactions on power systems*, 21(3):1458–1459, 2006.
- [9] Qifeng Li and Vijay Vittal. Convex hull of the quadratic branch ac power flow equations and its application in radial distribution networks. *IEEE Transactions on Power Systems*, 33(1):839–850, 2017.
- [10] Bernard C Lesieutre, Daniel K Molzahn, Alex R Borden, and Christopher L DeMarco. Examining the limits of the application of semidefinite programming to power flow problems. In *2011 49th annual Allerton conference on communication, control, and computing (Allerton)*, pages 1492–1499. IEEE, 2011.
- [11] Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan M Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013.