A Sample-Based Approach for Computing Conservative Linear Power Flow Approximations

Paprapee Buason, Daniel K. Molzahn
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA, USA
{pbuason6, molzahn}@gatech.edu

Sidhant Misra
Los Alamos National Laboratory
Los Alamos, NM, USA
sidhant@lanl.gov

***Abstract*—Non-convexities induced by the non-linear power flow equations challenge solution algorithms for many power system optimization and control problems. Linear approximations are often used to address these challenges by trading off modeling accuracy for tractability. The accuracy of a power flow linearization depends on the characteristics of the power system and the operational range where the linearization is applied. However, rather than exploiting knowledge of these characteristics for a particular system, many existing power flow linearizations are based on general assumptions for broad classes of systems, thus limiting their accuracy. Moreover, since existing linearizations do not consistently overestimate or underestimate quantities of interest such as voltage magnitudes and line flows, algorithms based on these linearizations may lead to constraint violations when applied to the system. In contrast, this paper computes *conservative linear approximations* of the power flow equations, i.e., linear approximations that intend to overestimate or underestimate a quantity of interest in order to enable tractable algorithms that avoid constraint violations. Using a sample-based approach, we compute these conservative linearizations by solving a constrained linear regression problem. We analyze and improve the conservative linear approximations via an iterative sampling approach, optimizing over functions of the quantities of interest, and a sample-complexity analysis. Considering the relationships between the voltage magnitudes and the active and reactive power injections, we characterize the performance of the conservative linear approximations for a range of test cases.**

***Index Terms*—Conservative linear approximation; sample selection; power flow approximation.**

I. Introduction

The AC power flow equations represent the physical laws governing the voltages and power flows in a power network. These equations are thus an essential part of the mathematical formulations for a variety of planning, optimization, and control problems. The implicit nonlinear nature of the power flow equations often results in computational and theoretical challenges when solving these problems. These challenges lead to the widespread use of simplified power flow representations to improve tractability and enable solution times that are within the problem's latency requirements.

P. Buason and D.K. Molzahn were supported by the National Science Foundation under grant number 2023140.

Linear approximations, due to their simplicity and efficiency, are among the most widely used class of approximations. Examples of linear approximations include the DC power flow for transmission systems [1], the LinDistFlow for distribution systems [2], and approximations based on the first-order Taylor expansion, such as the power transfer distribution factor (PTDF) representation, among others [3]. Employing these power flow linearizations significantly improves computational tractability, but comes at the price of accuracy which can result in unsafe or sub-optimal decision making. The magnitude of error introduced by a linearization is system and problem specific. However, most linearizations tend to be agnostic to these details and are derived based on general principles. For example, the DC approximation assumes that the voltage magnitudes throughout the system are close to their nominal values and that overall resistive losses are negligible, while the PTDF representation is only accurate within a neighborhood of a given operating point.

To address these drawbacks, [4] introduces the notion of an *optimal adaptive* linearization of the power flow equations. These linearizations are *adaptive* because they are tailored to a specific system and operating range, and *optimal* since they are constructed to minimize their deviation from the nonlinear AC equations with respect to a chosen error metric. A worse-case metric representing the maximum absolute error is considered in [4], whereas [5] minimizes the expected squared error with respect to an underlying probability distribution over the operating range. The mechanisms proposed to construct these linearizations are also able to *quantify* the error of the resulting approximations, thus allowing for more informed decision making when employing these linearizations.

Various recent studies have proposed methods to construct improved linearizations. In [6], a model that considers voltage magnitude and reactive power in addition to the voltage angle-active power relationship in the DC power flow (PF) is proposed. Although this relationship includes a squared term of voltage magnitude, it is seen as an independent variable, which keeps the relationship linear. In [7], an optimization strategy is proposed to choose an appropriate set of independent variables from a list of available independent variables (e.g., V, V^2, V^3, etc., where V is the voltage magnitude). Additionally, data-driven linear approximations have been studied in [7],

[8]. However, use of these linearizations in power flow-based optimization problems can still lead to constraint violations.

In this paper, we consider the notion of a *conservative linear approximation (CLA)*, which in addition to being optimal and adaptive, has the feature of being *conservative*. We say that an approximation is *conservative* if it preferentially under- or over- estimates the approximating quantity of interest. The motivation behind conservativeness is a natural one: almost all of our target applications seek to ensure that quantities such as voltages magnitudes, line flows, generator set-points, etc., remain within given upper and lower limits. For instance, bilevel optimization problems (see, e.g., [9]) and optimization problems with discrete variables (e.g., AC unit commitment [10]) include inequality constraints on voltage magnitudes, line flows, and generator outputs. However, directly enforcing these limits by modeling the nonlinear power flow equations (1) in these problems leads to significant computational challenges. CLAs address the need for constraint satisfaction during the construction of the linear approximations rather than deferring it to the problem-solving stage of the target application, thus achieving tractability while ensuring constraint satisfaction. CLAs, therefore, not only simplify the enforcement of bounds in the target problem but also enable the use of linearizations whose error properties are better suited for bound enforcement.

To construct these linear approximations, we propose a sample-based approach where the CLAs are obtained as the solution to a constrained regression problem. This differs from the iterative optimization-based approach of [4] and the polynomial chaos method used in [5], although it is closely related to the latter. The main reason for this choice is scalability, as sample-based regression problems are efficient to solve and can exploit parallelizability and high-performance distributed computing hardware. Further, we propose a basic sample-selection procedure that significantly boosts convergence speed and a Monte-Carlo validation in each selection step for obtaining theoretically guaranteed confidence bounds.

To further improve the quality of the approximations, we propose a method to optimally modify the quantity being approximated while retaining all of the desirable features of the CLA. For example, instead of approximating the voltage magnitude at a bus as a linear function of the active and reactive power injections, our method will automatically determine a different function (such as a monotonic quadratic function) of voltage magnitude that has a smaller linearization error. The procedure to determine the modified function is directly built into the constrained regression used to construct the CLA. While imposing a small computational cost, this can significantly reduce the approximation error. We also show that when the modified function is monotonic in the quantity of interest, this modification introduces no additional complexity when the constructed CLA is used in problems relevant to our target applications.

In summary, the main contributions of this paper are:

(*i*) A CLA formulation that is adaptive to the relevant system and operating range, optimal with respect to an error metric, and conservative in that it under- or over-estimates the quantity of interest.
(*ii*) A procedure to modify the quantity of interest being approximated that improves error without sacrificing beneficial properties of the CLA.
(*iii*) A data-driven approach to construct the CLAs using a constrained regression formulation and a sample selection method that improves the convergence rate.
(*iv*) Numerical analysis of the CLA for a variety of test cases and a demonstration of its application to the optimal power flow problem.

This paper is organized as follows. Section II provides background on the power flow equations. Section III formulates and solves an optimization problem for computing CLAs using a sample-based approach. Section IV presents our numerical tests. Section V concludes the paper and discusses future work.

II. The Power Flow Equations

Consider an n-bus power system with a fixed network topology. The sets of buses and lines are $\mathcal{N} = \{1, \ldots, n\}$ and $\mathcal{L}$, respectively. One bus in the system is specified as the slack bus (also known as the reference bus), where the voltage magnitude and angle is set to be $1\angle 0^\circ$ per unit. Each of the remaining buses is classified as either a PV bus, which specifies the active power and the voltage magnitude, or a PQ bus, which specifies the active and reactive power. The set of buses that are neighbors to bus i is denoted as $\mathcal{N}_i := \{k \,|\, (i,k) \in \mathcal{L}\}$. Subscript $(\cdot)_i$ denotes that quantity at bus i, and subscript $(\cdot)_{ik}$ denotes that quantity from or connecting bus i to k. The power flow equations at bus i are:

$$P_i = V_i^2 G_{ii} + \sum_{k \in \mathcal{N}_i} V_i V_k (G_{ik} \cos\theta_{ik} + B_{ik} \sin\theta_{ik}), \quad \text{(1a)}$$

$$Q_i = -V_i^2 B_{ii} + \sum_{k \in \mathcal{N}_i} V_i V_k (G_{ik} \sin\theta_{ik} - B_{ik} \cos\theta_{ik}), \quad \text{(1b)}$$

where V_i is the voltage magnitude at bus i, P_i (Q_i) is the active (reactive) power injection at bus i, $\theta_{ik} := \theta_i - \theta_k$ is the voltage angle difference between buses i and k, and G_{ik} (B_{ik}) is the real (imaginary) part of the network admittance matrix entry associated with buses i and k. Equations (1a)–(1b) are called the full AC-PF equations.

III. Problem Formulation and Solution Approach

This section introduces the conservative linear approximations (CLAs), formulates a constrained regression problem to compute the CLAs using a sample-based approach, and presents methods for improving the accuracy of the CLAs by modifying a function of the quantity being approximated and by using a sample selection procedure to speed up calculations.

Without loss of generality and for ease of exposition, we use the voltage magnitude at a bus as our quantity of interest and consider the CLAs that express the voltage magnitude as a function of active and reactive power injections.

A. Conservative linear approximations of the power flow equations

Linear approximations of the power flow equations are widely used to reduce the complexity of formulating many operation and planning problems. We propose to develop linear approximations that are *conservative*, meaning that the computed linear approximation is greater or equal to the actual quantity for an overestimating approximation or the other way around for an underestimating approximation.

Using voltage magnitude as our quantity to approximate and active and reactive power injections as input variables, our goal is to compute approximations that satisfy:

$$V_i - g_i(\mathbf{P}, \mathbf{Q}) \leq 0 \ (\geq 0); \text{ if overestimate (underestimate)} \tag{2}$$

where

$$g_i(\mathbf{P}, \mathbf{Q}) = a_{i,0} + \boldsymbol{a}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix} \tag{3}$$

is a linear function of active power injections at PQ buses ($\mathbf{P}$) and reactive power injections at PV and PQ buses ($\mathbf{Q}$).

B. Data-driven power flow approximations

Instead of an analytic or purely optimization-based approach, we propose a data-driven approach to compute the CLAs. The data-driven approach is highly scalable, can utilize advanced computing hardware, permits a parallel implementation, and can take advantage of improved sensing and communication technology for on-the-fly updates.

Using our example, we compute the data-driven power flow approximations by randomly sampling load demands (or generator outputs) within a given operating range and calculating the corresponding bus voltages by solving the power flow equations for each of these samples. The samples for load demands are obtained using a given probability distribution $\mathbb{P}_{\mathcal{S}}$ over a specified operating range $\mathcal{S}$. For example, one may consider $\mathcal{S} = \{P_{L_d}^{\min} \leq P_{L_d} \leq P_{L_d}^{\max},\ Q_{L_d}^{\min} \leq Q_{L_d} \leq Q_{L_d}^{\max}$ for all $d \in \mathcal{N}_{\mathcal{D}}\}$ where $(\cdot)_{L_d}$ denotes the load demand, $\mathcal{N}_{\mathcal{D}}$ is the set of load buses, $\mathbb{P}_{\mathcal{S}}$ is the uniform distribution, and superscript max (min) is an upper (lower) limit. The power injections at PQ buses change with varying load demands change, which also affects the reactive power injections at PV buses (see (3)).

In this paper, we approximate voltage magnitudes as linear functions of the power injections. Using uniformly generated samples, we propose to minimize errors (e.g., the 1-norm error) between the CLA and the samples. Using M power flow samples, we consider the following regression problem that computes a CLA while minimizing a chosen loss function $\mathcal{L}(\cdot)$ representing the approximation error.

$$\min_{g(\cdot)} \sum_{m=1}^{M} \mathcal{L}\left(V_{i,m} - g_i(\mathbf{P}_m, \mathbf{Q}_m)\right) \tag{4a}$$

$$\text{s.t.} \quad V_{i,m} - g_i(\mathbf{P}_m, \mathbf{Q}_m) \leq 0; \text{ if overestimate}, \tag{4b}$$

$$V_{i,m} - g_i(\mathbf{P}_m, \mathbf{Q}_m) \geq 0; \text{ if underestimate}, \tag{4c}$$

for all $m = 1, 2, \ldots, M$. Subscript m denotes the m-th sample. All bold quantities are vectors. The set of equations (4b) (or (4c) as chosen) enforces conservativeness and ensures that the CLA remains above (or below, as appropriate) the actual voltage magnitudes for all samples. Examples of $\mathcal{L}(\cdot)$ include the ℓ_1 and ℓ_2 loss functions.

Like any linear function, we note that the computed CLAs are defined for all inputs, i.e., every choice of power injections $\mathbf{P}_m, \mathbf{Q}_m$ is mapped to a voltage magnitude $g_i(\mathbf{P}_m, \mathbf{Q}_m)$. Conversely, the nonlinear power flow equations can be insolvable for some power injections, i.e., there may be power injections without corresponding voltages. When constructing the CLAs, we ignore any sampled power injections for which the power flow equations are insolvable. Thus, our goal is to compute the CLAs with respect to power injections for which the power flow equations are solvable.

C. Optimal output function design

For many applications, one can significantly reduce the linearization error for some quantity of interest by instead approximating an appropriately chosen function of that quantity. Using our voltage magnitude example, instead of approximating the voltage magnitude V_i itself, we propose to linearly approximate a function of the voltage magnitude, $f(V_i)$, in terms of $\mathbf{P}$ and $\mathbf{Q}$. We call the function $f(\cdot)$ an *output* function. To motivate the idea of an output function, we consider a simple two-bus system and examine the relationship between the voltage magnitudes and the power injections.

Multiplying (1a) by B_{ik} and (1b) by G_{ik} to eliminate $\cos\theta_{ik}$, the power flow equations for a two-bus system are:

$$P_1 B_{12} + Q_1 G_{12} = V_1^2(G_{11}B_{12} - G_{12}B_{11}) + V_1 V_2 (B_{12}^2 + G_{12}^2)\sin\theta_{12}, \tag{5a}$$

$$P_2 B_{21} + Q_2 G_{21} = V_2^2(G_{22}B_{21} - G_{21}B_{22}) + V_1 V_2 (B_{21}^2 + G_{21}^2)\sin\theta_{21}. \tag{5b}$$

Consider the case where the admittance matrix is symmetric, bus 1 is a slack bus, and bus 2 is a PQ bus. Adding (5a) to (5b) gives:

$$\begin{aligned} V_2^2(G_{22}B_{21} - G_{21}B_{22}) =& (P_1 + P_2)B_{12} + (Q_1 + Q_2)G_{12} \\ &+ \beta(G_{11}B_{12} - G_{12}B_{11}), \end{aligned} \tag{6}$$

where $\beta = -V_1^2$ is a constant, which is typically -1. Here, we model the power injections at both buses 1 and 2 (P_1, P_2, Q_1, and Q_2) as input variables for the linear approximation (3). Thus, (6) shows that the relationship between V_2 and power injections at buses 1 and 2 is quadratic. Hence, a linear output function $f(V_2) = V_2$ will result in a non-zero approximation error, whereas using a quadratic output function $f(V_2) = V_2^2$ could result in zero approximation error. The relationship in (6) is specific to the two-bus system, and more complicated systems will have non-zero approximation errors even with quadratic output functions. However, the two-bus example motivates the use of non-linear output functions. While there are many possible output functions, this paper considers two

natural choices: (i) polynomials and (ii) piecewise linear functions. Polynomial output functions are:

$$f(V_i) = b_n V_i^n + b_{n-1} V_i^{n-1} + \ldots + b_1 V_i, \tag{7}$$

where b_k is a polynomial coefficient for all $k = 1, 2, \ldots, n$. Given n segments, piecewise linear output functions are:

$$f(V_i) = c_{0j} + c_{1j} V_i, \quad \text{if } p_j \leq V_i \leq p_{j+1},\ 0 \leq j \leq n-1, \tag{8}$$

where p_j are the specified breakpoints, and c_{0j} and c_{1j} are coefficients corresponding to each piece.

To avoid calculating non-meaningful and trivial solutions, enable efficient computation of the CLA using the constrained regression in (4), and preserve the advantages of using the CLA in our target applications, we enforce the following characteristics on the output function $f(\cdot)$:

Characteristic 1. *The function $f(\cdot)$ is monotonically increasing within a specified range.*

Characteristic 2. *The range of $f(\cdot)$ is bounded below; i.e., $f(\overline{V_i}) - f(\underline{V_i}) \geq c$, for some $c > 0$, where $\overline{(\cdot)_i}$ and $\underline{(\cdot)_i}$ denote the maximum and minimum values of that quantity at bus i, respectively, among the sampled power flow solutions.*

Characteristic 1 means that $f(\cdot)$ is invertible. More importantly, it is required to ensure that when the CLA is used in an optimization problem, the resulting constraint *remains linear* even when the output function $f(\cdot)$ is nonlinear. This idea is demonstrated in detail using the optimal power flow (OPF) problem as an example application in Section III-F. Characteristic 2 is necessary to avoid trivial output functions since $f(\cdot) = 0$ will always result in zero approximation error.

For piecewise linear $f(\cdot)$, we impose additional characteristics to make the approximation well behaved.

Characteristic 3. *A piecewise $f(\cdot)$ is continuous within an operating range.*

Characteristic 4. *Let p be the set of all breakpoints. The slope of a piecewise function is bounded from above except at breakpoints; i.e., $f'(V) \leq d$, $\forall V \notin p$, for some specified $d \geq 0$, where $f'(V)$ denotes the first derivative of $f(V)$.*

Characteristic 3 ensures continuity of the piecewise function. Characteristic 4 prevents the piecewise function from having sharp changes within its range, especially towards the end of its range. This is also a desirable property for the polynomial output functions. The number and locations of the breakpoints in the piecewise linear function also impact the accuracy of the corresponding CLA. We chose the breakpoints such that each segment of the piecewise linear function contains the same number of sampled voltages.

It is also possible to include the following additional characteristic which may be beneficial for certain applications.

Characteristic 5. *(Optional) An output function is either convex or concave.*

We note that the output function of each voltage magnitude can be different. Let $f_i(V_i)$ denote the output function of V_i. We rewrite the regression in (4) incorporating the above characteristics (when applicable) to obtain the full version of the constrained regression to construct a CLA at bus i:

$$\min_{f_i(\cdot), g_i(\cdot)} \sum_{m=1}^{M} \mathcal{L}\left(f_i(V_{i,m}) - g_i(\mathbf{P}_m, \mathbf{Q}_m)\right) \tag{9a}$$

$$\text{s.t.} \quad f_i'(V_{i,m}) \geq 0 \tag{9b}$$

$$f_i(\overline{V_i}) - f_i(\underline{V_i}) \geq c \tag{9c}$$

$$f_i'(V_i) \leq d \text{ (if piecewise)} \tag{9d}$$

$$f_i(V_i) \text{ is continuous (if piecewise)} \tag{9e}$$

$$f_i(V_{i,m}) - g_i(\mathbf{P}_m, \mathbf{Q}_m) \leq 0 \text{ (if overestimate)} \tag{9f}$$

$$f_i(V_{i,m}) - g_i(\mathbf{P}_m, \mathbf{Q}_m) \geq 0; \text{ (if underestimate)}, \tag{9g}$$

for all $m = 1, \ldots, M$. Note that the decision variables in (9) are the coefficients of $f_i(\cdot)$ and $g_i(\cdot)$. We observe that when the output function $f_i(\cdot)$ can be linearly parameterized, which is the case for both polynomial and piecewise linear output functions, the constraints (9b)–(9g) are *linear* in these parameters (e.g., coefficients of the polynomial). When the loss function is chosen to be $\mathcal{L}(x) = |x|$ representing an ℓ_1-loss, the regression (9) can be cast as a *linear program* by using the standard slack formulation for minimizing the ℓ_1-norm, which enables an efficient solution. For the rest of this section and for our experiments, we will select $\mathcal{L}(x) = |x|$.

We define the minimum value of the objective function in (9a) as the **CLA error**. We also define the **actual error** as the error obtained by inverting the output function. Since characteristic 1 allows us to invert $f_i(\cdot)$, we can compute the voltage predicted by the CLA for each sample, $V_{i,m}^C$, as

$$V_{i,m}^C = f_i^{-1}\left(g_i(\mathbf{P}_m, \mathbf{Q}_m)\right). \tag{10}$$

The average actual error, $e_{a,i}$, is defined as

$$e_{a,i} = \frac{1}{M} \sum_{m=1}^{M} |V_{i,m}^C - V_{i,m}|. \tag{11}$$

As the name suggests, we would ideally want to directly minimize $e_{a,i}$ in (11) as opposed to the objective function (9a). However, this can make (9) a (possibly non-convex) nonlinear program that is much more challenging to solve than the linear program in (9). We therefore instead use the CLA error (9a) as a proxy for minimizing the actual error (11) that is significantly faster to compute. Our numerical experiments support the effectiveness of this approach.

D. Sample selection method and confidence bound

Since the calculated CLA is based on random samples, it is possible that newly drawn samples could be on the wrong side of CLA (e.g., a new sample may have a higher voltage than the overestimate provided by a CLA). Increasing the number of samples M used to compute the CLA in (9) improves confidence with respect to its conservativeness.

To reduce the number of samples needed to achieve a certain amount of confidence, we propose a sample selection method that iteratively updates a CLA using only violating samples (see Fig. 1). To perform the sample selection, we iterate between:

1) Drawing additional samples and adding points using the sample selection method, which only considers points that violate the CLA calculated in the previous iteration.
2) Updating the CLA by solving (9) and repeating these steps until termination conditions are satisfied.

All sampled points in step 1) can be further analyzed, so only meaningful points are added in each iteration. Note that the overestimating CLA and the underestimating CLA are updated separately based on violated points, i.e., a point that violates the overestimating CLA does not violate the underestimating CLA. Thus, that point is not included in the re-calculation of the underestimating CLA. At each step of the iteration, we use M^{out} samples to perform an out-of-sample validation which allows us to provide confidence bounds on the over- or under- estimation property of the CLA. Let e^{out} be the fraction of the M^{out} samples for which there is a violation of the over- or under- estimation property of the CLA in the out-of-sample testing. For a randomly drawn set of $(\mathbf{P}, \mathbf{Q})$, let V_i denote the actual voltage and let V_i^C denote the predicted voltage using the CLA. Then the probability $\mu^{err} = \mathbb{P}_{\mathcal{S}}(V_i^C < V_i)$ that a randomly drawn sample violates the over- or under- estimation property for the CLA can be bounded using the confidence bound

$$\mathbb{P}\left(\mu^{err} > e^{out} + \delta\right) \leq e^{-2\delta^2 M^{out}}, \tag{12}$$

which can be derived using the Hoeffding inequality [11].

E. Computation time

Using the CLA instead of the non-linear AC-PF equations can significantly improve the computational efficiency of problems for many applications. The regression computing the CLA (9) can be pre-computed offline given the system topology, line limits, range of load demands, and other pre-defined constraints. Moreover, this offline computation can be parallelized for each quantity of interest independently. Since the CLAs are constructed separately for each constrained quantity, CLA-based constraints can be imposed adaptively using a scenario and constraint generation approach that only enforces a constraint after determining that it may be binding. Conversely, when using the implicit AC-PF equations, if an operating scenario has even a single binding constraint, then all AC-PF equality constraints must be added for that scenario.

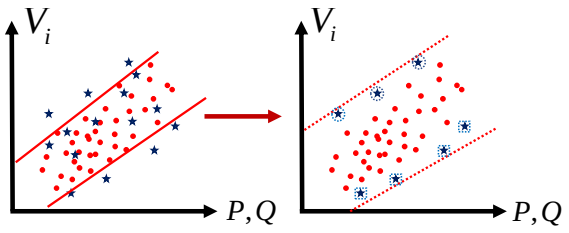


Fig. 1. An example of the sample selection method. The red circles and the blue stars are the previous and the new samples, respectively. The top (bottom) lines illustrate over- (under-) estimating CLAs. The right picture shows that only new samples which violate the previous CLA are included to compute the updated CLA. Specifically, the stars with the dashed circles (squares) are included for the updated over- (under-) estimating CLA.

F. Application: Optimal power flow

While the advantages of using the CLA are best seen in more complex optimization problems such as bilevel problems [9] and the AC unit commitment problem [10], a detailed analysis of such applications is out of scope of this paper. Instead, we use a simplified AC-OPF problem as a proof-of-concept demonstration of CLAs. The simplified OPF problem we consider here is identical to the classical OPF problem except that the line flow limits are neglected such that enforcing bus voltage limits is the primary goal in order to match the exhibition earlier in this paper. We emphasize that although we do not focus on line flow limits here, one may include them by constructing CLAs that approximate the line flows.

Let $\mathcal{N}_S$, $\mathcal{N}_{PV}$, and $\mathcal{N}_{PQ}$ denote the sets of slack, PV, and PQ buses, respectively. To compute the CLAs for this simplified OPF problem, we change the inputs of the linear function in (3) to be P_i and V_i at all $i \in \mathcal{N}_{PV}$. The approximated quantities are (i) Q_i, for all $i \in \mathcal{N}_{PV} \cup \mathcal{N}_S$, (ii) V_i for all $i \in \mathcal{N}_{PQ}$, and (iii) P_s at a slack bus $s \in \mathcal{N}_S$. Let $\widetilde{g}(\mathbf{P}, \mathbf{V})$ ($\underset{\sim}{g}(\mathbf{P}, \mathbf{V})$) be the function that overestimates (underestimates) quantities of interest, $\widetilde{f}(\,\cdot\,)$ ($\underset{\sim}{f}(\,\cdot\,)$) be the polynomial output function corresponding to $\widetilde{g}(\mathbf{P}, \mathbf{V})$ ($\underset{\sim}{g}(\mathbf{P}, \mathbf{V})$), and $\mathcal{N}_{\mathcal{G}}$ be the set of all generators. We denote the power generation and generation cost as P_G and c, respectively. By formulating voltage and power generation limits with CLAs, the simplified OPF problem becomes:

$$\min \sum_{k \in \mathcal{N}_{\mathcal{G}}} c_k P_{G_k} \tag{13a}$$

$$\text{s.t.} \quad \widetilde{f}_i(\overline{V}_i) \geq \widetilde{g}_i(\mathbf{P}, \mathbf{V}), \quad \underset{\sim}{f}_i(\underline{V}_i) \leq \underset{\sim}{g}_i(\mathbf{P}, \mathbf{V}), \tag{13b}$$

$$\widetilde{f}_j(\overline{Q}_j) \geq \widetilde{g}_j(\mathbf{P}, \mathbf{V}), \quad \underset{\sim}{f}_j(\underline{Q}_j) \leq \underset{\sim}{g}_j(\mathbf{P}, \mathbf{V}), \tag{13c}$$

$$\widetilde{f}_s(\overline{P}_s) \geq \widetilde{g}_s(\mathbf{P}, \mathbf{V}), \quad \underset{\sim}{f}_s(\underline{P}_s) \leq \underset{\sim}{g}_s(\mathbf{P}, \mathbf{V}), \tag{13d}$$

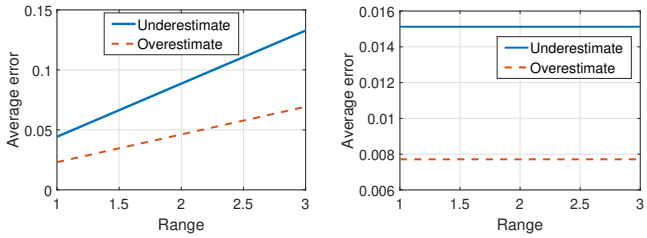
$$P_s = \widetilde{g}_{s1}(\mathbf{P}, \mathbf{V}), \tag{13e}$$

$$P_{G_k}^{\min} \leq P_{G_k} \leq P_{G_k}^{\max}, \; Q_{G_k}^{\min} \leq Q_{G_k} \leq Q_{G_k}^{\max}, \tag{13f}$$

for all $i \in \mathcal{N}_{PQ}$, $j \in \mathcal{N}_{PV} \cup \mathcal{N}_S$, $s \in \mathcal{N}_S$, and $k \in \mathcal{N}_{\mathcal{G}}$. We call (13) a **CLA-OPF** problem. Constraints (13b)–(13d) ensure that all approximated quantities are within their limits. The overestimated value of the power generation at the slack bus in (13e) is used in the objective function. The resulting CLA-OPF is a linear program that nevertheless ensures satisfaction of the inequality constraints in the non-linear AC-OPF problem so long as the computed conservative linear approximations are indeed conservative.

IV. Simulation and Results

In this section, we use numerical experiments on a number of test cases to analyze various properties of the CLA, demonstrate the advantages of output function design and sample



(a) CLA errors (b) Actual errors

Fig. 2. Plot of (a) CLA errors, (b) actual errors at bus 6 in the 6-bus system using the quadratic output functions. The solid and dashed lines represent the errors from underestimating and overestimating the voltage magnitude V_6, respectively

selection procedures, and demonstrate the effectiveness of the CLA in enforcing constraints using the simplified OPF.

The test cases we use in the simulations are a 6-bus system (adapted from *case6ww*), whose parameters can be found in [12], the 33-bus system *case33bw* from MATPOWER [13], and the IEEE 300-bus system. Note that the samples used to create each table and figure below are drawn independently, and 5000 samples are drawn in each test unless stated. All load demands vary within -50% to 200% of the values given in the MATPOWER test systems, except for the OPF simulations where the loads are fixed and the generator outputs vary. Only samples for which the power flow solver converges are included in CLA calculation, and the voltage magnitude at one representative bus per system is chosen for illustration. The loss function represents the ℓ_1 error.

A. Range constraint for polynomial output functions

As discussed in Section III-C, a lower bound c is enforced on the range of the output function (characteristic 2) to avoid trivial solutions. The value of c is an arbitrary parameter chosen when constructing a CLA, and it is therefore important to analyze the impact that this choice has on the CLA's linearization error.

For this purpose, Fig. 2(a) shows a representative example of the relationship between c and the CLA error for the 6-bus system. We observe that the CLA errors vary linearly with the value of c. This is reasonable since decreasing c relaxes constraint (9c), leading to reductions in the optimal objective value of (9). Conversely, Fig. 2(b) shows that the actual errors obtained using (10) do *not* depend on the value of c as the plots in this figure are constant. The results from cubic and quartic output functions, which are not shown in this paper, are similar. This analysis suggest that the value of c can be arbitrarily chosen and does not affect the optimal solution for over/underestimated voltage magnitude.

B. Error reduction with polynomial output functions

We next investigate the effect of using polynomials (7) with various degrees n as output functions as discussed in Section III-C. While the CLA error is guaranteed to be non-increasing with the degree of the polynomial, the same is not necessarily true for the actual error in (11).

Moreover, as discussed above, the choice of c does not affect the actual error of the CLAs. We arbitrarily set c to 1.5. The results from the first-order Taylor approximation around the operating point (given in the MATPOWER test cases) are also included for reference. The results are reported in Table I.

TABLE I
RESULTS COMPARING AVERAGE CLA AND ACTUAL ERRORS FROM CLAS AT BUS 6 IN THE 6-BUS SYSTEM AND AT BUS 33 IN THE 33-BUS SYSTEM

Case	Function	Average errors ($\times 10^{-2}$) in per unit			
		Underestimate		Overestimate	
		CLA	Actual (*)	CLA	Actual (*)
6-bus	*Quartic*	4.57	1.451 (54.70%)	2.29	0.655 (47.93%)
	Cubic	4.63	1.358 (57.60%)	2.29	0.657 (47.77%)
	Quadratic	4.64	1.327 (58.57%)	2.33	0.662 (47.38%)
	Linear	11.73	3.203 (0%)	4.61	1.258 (0%)
	Taylor	1.800§			
case33bw	*Quartic*	0.90	0.059 (58.74%)	0.32	0.021 (51.16%)
	Cubic	0.91	0.059 (58.74%)	0.32	0.021 (51.16%)
	Quadratic	0.91	0.059 (58.74%)	0.32	0.021 (51.16%)
	Linear	2.15	0.143 (0%)	0.64	0.043 (0%)
	Taylor	0.259§			

*The values in ( ) show the percentage reductions in the actual errors compared to the actual errors from linear function.
§ The errors from first-order Taylor approximation are the average absolute value of the errors caused by the approximation.

Table I presents the averages of the CLA errors and the actual errors from the quartic, cubic, quadratic, and linear functions for the CLA associated with the voltage magnitude at bus 6 in the 6-bus system and bus 33 in the 33-bus system. Note that using a linear output function is equivalent to not designing an output function. The average CLA errors are smaller when the degree of the polynomial is higher. This is due to the fact the optimal solution of a lower-degree polynomial is a feasible solution for the higher-degree polynomial. The percentage reductions in the average actual errors with respect to the linear function are also shown in Table I. The average actual errors from the quadratic, cubic, and quartic functions are significantly reduced from the linear function (from 47% to 59% in both systems) with the quadratic and higher-degree polynomials showing similar performance. In comparison, the first-order Taylor approximation is not conservative and has significantly larger error.

We further analyze the behavior of all tested functions by depicting them in Fig. 3. To enable a consistent comparison, each function is adjusted using the offset a_0, which is the constant term for $g(\mathbf{P}, \mathbf{Q})$ defined in (3). The functions corresponding to overestimates of the voltage magnitudes are shown for a range from 20% below to 20% above the extreme sampled values. Since both CLA and actual errors for quadratic, cubic, and quartic functions are fairly close (refer to Table I), these functions are almost identical within the sampled operating range but are all quite different from the linear function. However, they are different functions as clearly seen outside the sampled operating range. For this illustration, we intentionally select a large range of voltage magnitudes to more clearly show the curvature of each function, but note that similar behavior occurs for smaller voltage magnitude ranges.

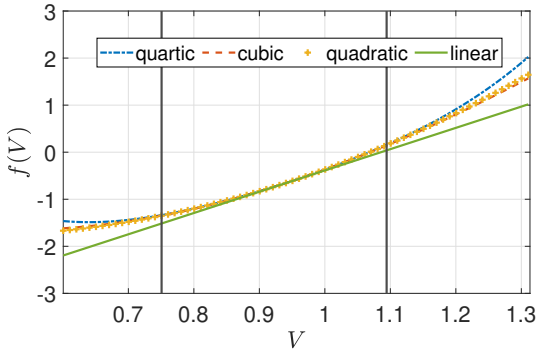
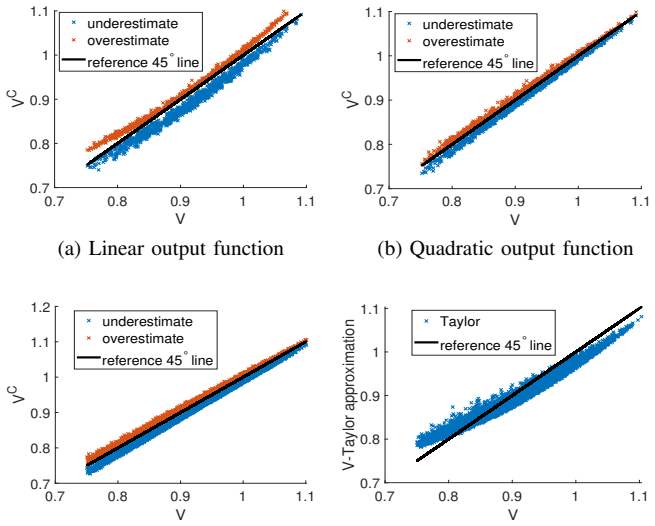


Fig. 3. Results from the overestimating CLAs for the voltage magnitude at bus 6, V_6, in the 6-bus system for polynomials of different degrees. The black vertical lines show the minimum and maximum observed voltage magnitudes.



(a) Linear output function (b) Quadratic output function

(c) Cubic output function (d) First-order Taylor approximation

Fig. 4. Plots of results from (a) linear, (b) quadratic, (c) cubic, and (d) first-order Taylor approximation functions at bus 6 in the 6-bus system. The blue and red points represent underestimates and overestimates of the voltage magnitude, respectively, in per unit. The first-order Taylor approximation neither overestimates nor underestimates the voltage magnitude. The black line at 45° represents a perfect fit.

To further illustrate how the output function $f(V)$ transforms the samples, we compute over/underestimates of the voltage magnitudes for each sample. Fig. 4 shows the results for the linear, quadratic, cubic, and first-order Taylor approximation functions. The horizontal axis shows the voltage magnitude from the power flow solution while the vertical axis shows the predicted voltage magnitude from the CLA. The quadratic and cubic output functions make the behavior of the over/underestimated voltages closer to linear, resulting in reduced approximation errors. The figure also shows the conservative property of the CLAs when compared to the first-order Taylor approximation.

C. Error reduction with piecewise linear output functions

Section III-C also discusses piecewise linear output functions. The accuracy of a piecewise linear function depends on the number of breakpoints. In this simulation, the number of breakpoints is 100, the number of samples is 5000, and each piece contains the same number of samples. The value of d from Characteristic 4 is set to 50.

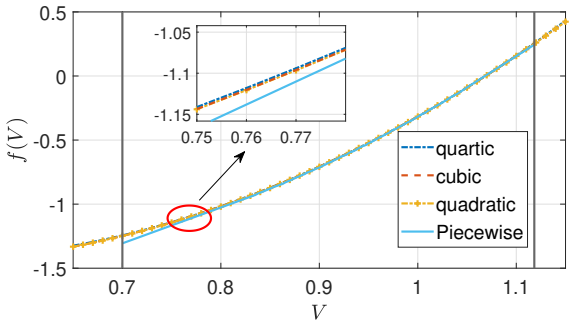


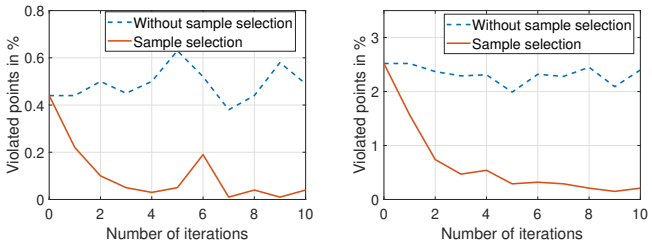
Fig. 5. The results from overestimates of the voltage magnitude at bus 6 in the 6-bus system with different polynomials and a piecewise linear function. The black vertical lines show the minimum and maximum observed voltage magnitudes obtained from samples. The horizontal axis corresponds to the per unit voltage magnitude and the vertical axis is the output function. The results from underestimates of the voltage magnitude are similar.

Fig. 5 compares a piecewise linear function with quadratic, cubic, and quartic functions. The results from the piecewise linear function are similar to the results from other functions, especially at the middle of the range of voltage magnitudes where the power flow solutions from samples are most dense (refer to Fig. 4). On the other hand, when the voltage magnitude is near the minimum or maximum value, the piecewise linear function has a larger difference with respect to the other functions. This is due to an overfitting phenomenon resulting from the larger number of free parameters in the piecewise approximation compared to the polynomials. The limit imposed on the maximum derivative in (9d) serves as a regularizing constraint for the piecewise linear case.

Table II shows the average CLA errors and the average actual errors when varying the number of breakpoints. Results using one breakpoint are the same as those from a linear function. Increasing the number of breakpoints continually reduces the CLA errors, but the actual errors saturate quickly.

TABLE II
RESULTS COMPARING ERRORS FROM PIECEWISE LINEAR APPROXIMATION WITH DIFFERENT NUMBER OF BREAKPOINTS AT BUS 6 IN THE 6-BUS SYSTEM AND BUS 33 IN THE 33-BUS SYSTEM

Case	Number of Breakpoints	Average errors ($\times 10^{-2}$) in per unit			
		Underestimate		Overestimate	
		CLA	Actual	CLA	Actual
6-bus	1	12.59	3.331	4.56	1.215
	10	4.80	1.335	2.59	0.723
	100	4.11	1.251	2.27	0.674
	500	4.05	1.251	2.20	0.682
case33bw	1	1.99	0.137	0.62	0.043
	10	0.83	0.057	0.32	0.022
	100	0.76	0.053	0.30	0.020
	500	0.72	0.051	0.28	0.019



(a) 6-bus system at bus 6 (b) 33-bus system at bus 33

Fig. 6. Plots showing the percentage of violated samples (vertical axis) after applying sample selection in each iteration (horizontal axis). The red lines and the blue lines show the results with and without applying sample selection, respectively, considering the same number of additional samples in each case. Note that each iteration uses different sets of samples.

D. Sample selection

The accuracy of the CLA depends on the samples input to (9). Increasing the number of samples improves the accuracy of the CLA but reduces the computational tractability of (9), thus motivating the sample selection method discussed in Section III-D which iteratively updates the CLA with violated samples. To demonstrate the effectiveness of this method, Fig. 6 presents the reduction in number of violated samples using an *out-of-sample* validation after applying the sample selection method for the 6- and 33-bus systems. Drawing 10000 samples for each iteration, the sample selection method reduces the percentage of violated samples from 0.49% to 0.04% after 10 iterations for the 6-bus system and from 2.40% to 0.21% after 10 iterations for the 33-bus system. The results show that sample selection significantly improves the convergence rate of the error compared to random sampling.

E. Scalability: CLA in larger systems

We have so far shown results from the 6- and 33-bus systems, which represent a small transmission system and a small radial distribution system, respectively. We next use the larger IEEE 300-bus system provided in MATPOWER to show the consistency of some of the properties analyzed in the previous sections. This simulation focuses on bus 250. Only average actual errors are computed as they are the quantities of interest. The piecewise linear function has 1000 breakpoints.

TABLE III
RESULTS COMPARING ACTUAL ERRORS FROM DIFFERENT CLA FUNCTIONS AT BUS 250 IN THE IEEE 300-BUS SYSTEM

Function	Average actual errors ($\times 10^{-3}$) in per unit	
	Underestimate (**)	**Overestimate (**)**
Quartic	5.961 (27.79%)	4.884 (17.65%)
Cubic	5.957 (27.84%)	4.676 (21.16%)
Quadratic	5.801 (29.73%)	4.650 (21.60%)
Piecewise Linear	6.026 (27.00%)	4.768 (19.61%)
	8.255 (0%)	5.931 (0%)
Taylor	22.8§	

**The values in ( ) show the percentage reductions in the actual errors compared to the actual errors from a linear function.
§ The errors from the first-order Taylor approximation are the average absolute value of the errors caused by the approximation.

Table III compares the average actual errors for different CLA functions. For the voltage underestimation, the piecewise linear functions and other polynomials with degree two or more reduce the average actual errors by 27% to 29.73% from the linear function. Similarly, these errors reduce by 17.65% to 21.60% for the voltage overestimation. These output functions thus notably increase the accuracy in the voltage approximation. For the first-order Taylor approximation, which is not conservative, the average errors are significantly greater, approximately by a factor of three to four compared to the other CLA functions.

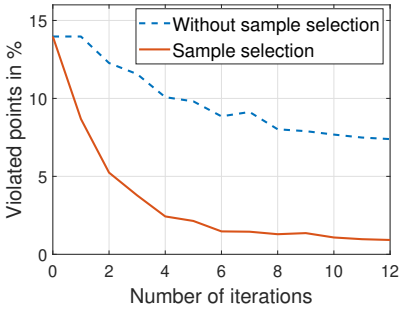


Fig. 7. Sample selection method for CLA at bus 250 in the 300-bus system.

TABLE IV
RESULTS COMPARING SOLUTIONS FROM AC-, DC-, LA-, AND CLA-OPF

	Case		
	case6ww	*case9*	*case14*
AC-OPF	2986.04	1456.83	5368.30
DC-OPF	2995.15 (0.31%)	1502.82 (3.16%)	5368.52 (0.004%)
Violation	V (0.029 pu)	-	-
LA-OPF	2987.28 (0.04%)	1468.48 (0.88%)	5368.52 (0.004%)
Violation	-	-	V (0.004 pu)
CLA-OPF	2987.51 (0.05%)	1469.68 (0.88%)	5368.52 (0.004%)
Violation	-	-	-

Fig. 7 shows the results from the sample selection method. We use sufficient iterations to ensure that the probability of violation is less than 1%. In each iteration, 10000 samples are randomly drawn for out-of-sample testing. The percentage of violated points reduces from 7.39% to 0.92% after 12 iterations of sample selection. Although the larger system requires more samples to achieve a low violation probability, the iterative sample selection approach still reduces the violation probability significantly faster than using random samples.

F. Simplified OPF

The OPF problem serves as a simple example for an application of the CLA method. We solve the simplified OPF problem (discussed in Section III-F) with different approximation techniques (e.g., DC-OPF and CLA-OPF) to compare with the solution to the AC-OPF problem. We also compare our results with linear approximation OPF (LA-OPF), which minimizes the 1-norm error and follows characteristics 1 and 2, but is not necessarily conservative. For the LA-OPF and the CLA-OPF, we choose quadratic output functions.

Table IV compares different approximations of the nonlinear AC-OPF for the 6-, 9-, and 14-bus systems. The first row in each cell of this table reports the actual cost corresponding to the AC-PF feasible solution obtained by using the control

set points prescribed by the various OPF formulations. The second row in each cell gives the maximum voltage magnitude violation for the solution. We see that only the CLA-based control actions do not have violations, which is attributed to the conservative nature of the CLAs. With the approximated voltage magnitude of 1 per unit (pu), the DC-OPF gives the maximum voltage violation of 0.029 pu in *case6ww*. For the LA-OPF, the maximum voltage violation is 0.004 pu in *case14*. By minimizing an upper bound on the cost, CLA-OPF is also able to obtain optimal or near-optimal solutions with lower costs than the other approximations. The optimality gaps from the CLA-OPF are comparable to those from the LA-OPF, and they are significantly smaller than the DC-OPF in *case6ww* and *case9*.

V. Conclusion

This paper proposes a power flow linearization approach that intends to be conservative (overestimate/underestimate a quantity of interest) and adaptive (due to a sample-based approach, it is tailored to a specific system and an operational range). Our proposed method is not limited to a linear output function as we introduce a higher-degree polynomial as an independent variable, which maintains linearity. The additional degrees of freedom from a higher-degree polynomial enable us to exploit a closer-to-linear relationship to other known quantities (e.g., power injections). Our numerical results demonstrate improvements in accuracy relative to other linear approximations while still being conservative. Additionally, the results in Section IV-F show the effectiveness of our approach for the illustrative example application provided by the OPF problem.

Our future work aims to develop alternative output functions and sample selection methods in order to further improve CLA accuracy. Furthermore, we will focus on applications of the CLA methods to power system planning and resiliency problems formulated as, for instance, bilevel programs.

References

[1] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[2] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan. 1989.

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

[4] S. Misra, D. K. Molzahn, and K. Dvijotham, "Optimal adaptive linearizations of the ac power flow equations," in *20th Power Syst. Comput. Conf. (PSCC)*, June 2018.

[5] T. Mühlpfordt, V. Hagenmeyer, D. K. Molzahn, and S. Misra, "Optimal adaptive power flow linearizations: Expected error minimization using polynomial chaos expansion," in *IEEE Milan PowerTech*, 2019.

[6] Z. Yang, H. Zhong, A. Bose, T. Zheng, Q. Xia, and C. Kang, "A linearized OPF model with reactive power and voltage magnitude: A pathway to improve the MW-only DC OPF," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1734–1745, Mar. 2018.

[7] Z. Fan, Z. Yang, J. Yu, K. Xie, and G. Yang, "Minimize linearization error of power flow model based on optimal selection of variable space," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1130–1140, Mar. 2021.

[8] Y. Liu, N. Zhang, Y. Wang, J. Yang, and C. Kang, "Data-driven power flow linearization: A regression approach," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2569–2580, May 2019.

[9] S. Wogrin, S. Pineda, and D. A. Tejada-Arango, *Applications of Bilevel Optimization in Energy and Electricity Markets*. Cham: Springer International Publishing, 2020, pp. 139–168.

[10] A. Castillo, C. Laird, C. A. Silva-Monroy, J.-P. Watson, and R. P. O'Neill, "The unit commitment problem with AC optimal power flow constraints," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4853–4866, Nov. 2016.

[11] W. Hoeffding, "Probability inequalities for sums of bounded random variables," in *The collected works of Wassily Hoeffding*. Springer, 1994, pp. 409–426.

[12] P. Buason, "Conservative linear approximation test systems," https://github.com/pbuason/testcase, 2021.

[13] 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 Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.