

Circuit-theoretic Joint Parameter-State Estimation — Balancing Optimality and AC Feasibility

Peng Sang, and Amritanshu Pandey
Dept. of Electrical and Biomedical Engineering
The University of Vermont
Burlington, VT USA
{Peng.Sang, Amritanshu.Pandey}@uvm.edu

Abstract—AC State Estimation (ACSE) is widely recognized as a practical approach for determining the grid states in steady-state conditions. It serves as a fundamental analysis to ensure grid security and is a reference for market dispatch. As grid complexity increases with rapid electrification and decarbonization, there is a growing need for more accurate *knowledge* of the grid operating state. However, existing ACSE algorithms have technical gaps. Critically, current ACSE algorithms are susceptible to *erroneous* system parameters, which are assumed to be fixed in traditional approaches. In this paper, we build a novel circuit-theoretic joint parameter-state estimation algorithm to address this limitation. The innovative algorithm builds an analogous equivalent circuit of the grid with states and certain parameters *unknown*. It solves a circuit-constrained optimization to estimate the *most likely* grid states and parameters given a set of measurements. Further, it quantifies the goodness of the estimated output by formulating *tight* convex envelopes around the original non-convex problem to quantify the quality of estimates. We compare the various proposed approaches on systems with up to 2869 nodes while demonstrating a tradeoff between solution optimality and model fidelity.

Index Terms—convex relaxation, equivalent circuit formulation, McCormick envelopes, non-convex optimization, parameter-state joint estimation

NOMENCLATURE

| Symbol | Interpretation |
|--------------------------------|--|
| V_r^k, V_i^k | Real/imag voltage at bus k |
| I_r^k, I_i^k | Real/imag current injection at bus k |
| V_r^{kl}, V_i^{kl} | Real/imag voltage across bus k and l |
| $ V^k , \theta^k$ | Voltage magnitude/angle at bus k |
| P^k, Q^k | Active/reactive power injection at bus k |
| \mathbf{z} | Set of measurements |
| $z_{ V }^k, z_P^k, z_Q^k$ | Remote terminal units (RTU) injection measurements of voltage mag., real/reactive power at bus k |
| $z_{P,fl}^{kl}, z_{Q,fl}^{kl}$ | RTU flow measurements for a edge between k and l ; real/reactive power at the metered end |
| $z_{ V }^{met}$ | RTU flow measurements for a edge between k and l ; voltage mag. at the metered end |
| \mathbf{v} | Set of voltage state variables |

| | |
|--|---|
| v^k | Voltage pair at bus k ; $x^k = [V_r^k, V_i^k], \in \mathbf{v}$ |
| n^k | Noise term (in per unit system) |
| n_r^k, n_i^k | Real and imag. noise term at bus k |
| $n_{ V }^k$ | Voltage mag. noise term at bus k |
| $I(v) = 0$ | Network balance Kirchoff's Current Law (KCL) equations |
| I_r^{kl}, I_i^{kl} | Real/imag current injection from bus k to l |
| W | Error covariance matrix |
| w_I^k, w_V^k | Weights for current and voltage measurements at bus k , respectively. |
| G^{kl}, B^{kl} | Admittance terms between bus k and l |
| B_{sh}^k | Shunt susceptance at bus k |
| \mathcal{P} | Set of unknown parameters |
| $\mathcal{P}^U, \mathcal{P}^L$ | Set of original upper and lower bounds of unknown parameters |
| $\tilde{\mathcal{P}}^U, \tilde{\mathcal{P}}^L$ | Set of tight upper and lower bounds of unknown parameters |
| \mathbf{x} | Set of variables, $\mathbf{x} = \mathcal{P} \cup \mathbf{v}$ |
| \mathcal{I} | Set of injection measurements |
| \mathcal{E} | Set of flow measurements |
| \mathcal{R} | Set of RTU measurements, $\mathcal{R} = \mathcal{I} \cup \mathcal{E}$ |

I. INTRODUCTION

A. Motivation of the research

The outputs from the network topology processor (NTP) and AC state-estimator (ACSE) construct real-time models of power grids in control rooms. These models are critical to evaluate the grid reliability through what-if-like contingency analysis and also serve as base networks for downstream real-time economic dispatch. Amongst other things, the quality of these grid models is very sensitive to the value of the input model parameters. If the model parameter values are not sufficiently accurate, the ACSE solution and the quality of the downstream models deteriorate. These parameter values include but are not limited to impedances of the transmission lines and transformers, shunt values, and transformer tap positions.

There exist many scenarios today when the network parameter values are erroneous or unknown and undetectable by general bad-data flow. These scenarios include equipment aging, temperature variation, unreported conductor reconfiguration, loss of a single conductor in a parallel configuration, unreported transformer tap adjustments, and manual database entry errors. In most of these scenarios, inaccurate backlogged

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number DE-EE0010147. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

parameters will be fed into the ACSE, with potentially damaging impact.

Grid control rooms have processes to detect bad data, including erroneous measurements and bad parameters. The most common is a post-processing approach where operators apply statistical tests on ACSE output (i.e., the residual) to identify suspect measurements and parameter candidates. While the approach is effective in the presence of only gross bad-measurement data, there is evidence that wrong parameters do not always show up in residual output [1]. Localization is challenging too with such statistical approaches. The largest residual value does not always correspond to localized error information [2]. Also, distinguishing measurement from parameter errors remains challenging, one being a random variable and the other an unknown quantity [3].

B. State-of-the-art and limitations

Alternatively, erroneous parameters can be assumed unknown and obtained via a joint parameter-state estimation problem, which can estimate both state and unknown parameters concurrently. The approach is also called “state (vector) augmentation” and has been widely studied for both transmission and distribution networks [1], [4], [5]. The approach overcomes some challenges with the post-processing approach based on ACSE residuals, but other challenges persist. The original joint approach [1] formulates the joint problem as a weighted least square problem [6] and applies the Gauss-Newton solution approach. The problem is non-convex (including trilinear and trigonometric terms), cannot model zero-injection nodes implicitly, and cannot include multi-period constraints without loss of generality. Another approach [4] proposed a linear optimization approach for joint parameter-state estimation. The method [4] is a two-step algorithm, wherein a relaxed ACSE is followed by solving an overdetermined system of injection currents, bus voltages, and unknown parameters with the ordinary least squares (OLS) method to recover the unknown parameters. OLS method in the latter step doesn’t guarantee the exact solution to the relaxation and can result in parameter estimates that do not satisfy the ACSE constraints. More recently [5] proposed a joint parameter-state estimator for distribution networks. The method formulates the network constraints with current-voltage (IV) variables and applies a nonlinear programming (NLP) based Newton-Raphson (NR) solution approach to overcome convergence challenges of the Gauss-Newton-based Weighted Least Squares (WLS) and OLS approaches. While the method has demonstrated practical empirical results, due to the non-convex (and unobservable) nature of the problem, the solution quality cannot be quantified when the parameters are *truly* unknown.

Dynamic estimation techniques can also recover unknown parameters and provide benefits over techniques that operate on steady-state equations. For instance, [7] and [8] apply Kalman Filter (KF) and Particle Filter (PF) techniques to estimate unknown grid parameters. However, these methods require phasor measurement units (PMUs) for high-time-density readings and multi-scan sampling. The setup is not

realizable today, as the majority of grid sensors in the bulk power systems (BPS) are remote-terminal units (RTUs).

Therefore, joint parameter-state estimation techniques based on steady-state equations are most applicable today for estimating grid state and unknown parameters. However, in almost all cases, they are formulated as a non-convex optimization. As such, gauging the solution quality remains a challenge, and convergence to local minima and saddle points is possible as mentioned in literature [9][10](for instance, see Fig. 10), nullifying the practicality of the approach. Convex relaxations [11], [12] can provide a certificate of optimality for the solution, but they deviate from the physics of AC models.

In essence, the fundamental challenge is that of a tradeoff between solution optimality and AC feasibility for non-convex joint parameter-state estimation algorithms. Furthermore, the underlying problem is inherently *globally* unobservable as many-to-one mappings in a noiseless environment are possible due to nonlinear AC constraints. Therefore, we will develop a novel circuit-theoretic approach with IV variables to address and mitigate these challenges, effectively.

C. Insights of the problem

We formulate a novel circuit-theoretic joint parameter-state estimation problem (ckt-PSE) to address the challenges with the state-of-the-art [1], [4], [5], [7], [8]. Our novelty is on two fronts:

- 1) First, from the viewpoint of circuit-based network constraints [13], we formulate a joint parameter-state estimation with current-voltage (IV) variables; we apply feature transformation to reduce the nonlinearities in ckt-PSE to only bilinear terms.
- 2) Second, we apply McCormick relaxation with sequential bound tightening (SBT), based on the foundational work in [14] and [15] we extend the algorithm in [16] to ckt-PSE, to provide quantitative guarantees on the estimated solution quality.

In the results, we demonstrate the efficacy of ckt-PSE for networks up to 2869 nodes. We show that we can recover unknown parameters with quantifiable accuracy. We compare the developed NLP and relaxed convex estimation formulations and consider multi-period constraints to improve solution quality.

II. PRELIMINARIES

A. Equivalent Circuit-Theoretic Formulation for AC State Estimation (ckt-SE)

This paper will build on *foundational* equivalent circuit-based approaches for ACSE [13], [17] that solve physics-based Maximum Likelihood Estimation (MLE) and is compatible with both RTUs and PMUs measurements. MLE is an observation-driven method. It takes observed measurements as apriori information to estimate realizations of random variables from assumed distributions (Gaussian, in this case, an assumption in [6] and applied in [1][18] and abundance of other works). Both [13], [17] aim to find a system state that maximizes the

posterior. The methods differ in how they map measurement data to circuit models. In a similar vein, these methods also differ from traditional WLS [6].

We build on this approach [13] because the underlying circuit model is linear with known convergence guarantees. In this paper, we term this approach ckt-SE. The approach uses circuit-based 'IV' instead of the conventional 'PQV' formulation for measurement functions. The state variables are real V_r and imaginary V_i voltages instead of the voltage magnitude ($|V|$) and the angle θ .

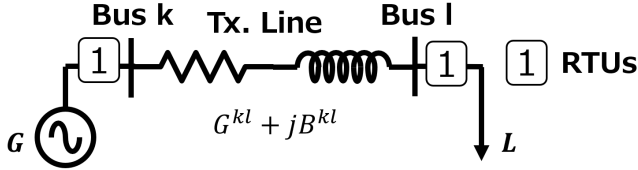


Fig. 1. Simple 2-bus network.

To illustrate, ckt-SE, consider a simple 2-bus network in Fig. 2. We assume there is an RTU at each of the two buses, measuring voltage magnitude, nodal power injection, and power flow in adjacent line. In ckt-SE, we replace the measured injection parameters (z_P , z_Q , and $z_{|V|}$) with equivalent circuit elements. The approach is built on certain relaxations. To construct the measurement circuit for RTU nodal power injection, it applies a feature transform to reduce three-dimensional measurements z_P , z_Q , and $z_{|V|}$ to two dimensions:

$$z_G = \frac{z_P}{(z_{|V|})^2}; z_B = \frac{z_Q}{(z_{|V|})^2} \quad (1)$$

With the transformation, the nodal network constraints at bus k and l are affine as the injection currents I_{RTU}^b at bus $b \in \{k, l\}$ are linear functions of unknown voltages per KCL:

$$I_{RTU,r}^b = z_G^b V_r^b - z_B^b V_i^b + n_r^b \quad (2)$$

$$I_{RTU,i}^b = z_G^b V_i^b + z_B^b V_r^b + n_i^b \quad (3)$$

Note that (2) includes noise terms n in the measurement circuits as measurements are imperfect or noisy. The relaxation is also in the modeling of noise n . While traditional WLS-formulation [6] does not assume measurement noise to be Gaussian, traditional MLE-formulation [1] assumes all measurement noise (n_P^z , n_Q^z , and $n_{|V|}^z$) to be Gaussian. [13] assumes the injection current noise to be Gaussian to maintain the key MLE principle. RTUs use wattmeters that measure power through analog or digital processing of voltage and current measurements with averaging, and as such, the Gaussian assumption in [1] is primarily to preserve key properties of MLE.

We can similarly construct the measurement circuit for RTU line flow measurements (see Fig. 3). The measurement function for the line flow measurement between k and l , metered at k , is given by:

$$z_{G,fl}^{kl} \cdot V_r^k - z_{B,fl}^{kl} \cdot V_i^k + n_r^{kl} + G^{kl} V_r^k - B^{kl} V_i^k = 0 \quad (4)$$

$$z_{G,fl}^{kl} \cdot V_i^k + z_{B,fl}^{kl} \cdot V_r^k + n_i^{kl} + G^{kl} V_i^k + B^{kl} V_r^k = 0 \quad (5)$$

where, the flow measurements $z_{P,fl}$, $z_{Q,fl}$ are relaxed as $z_{G,fl}$, $z_{B,fl}$:

$$z_{G,fl} = \frac{z_{P,fl}}{(z_{|V|})^2}; z_{B,fl} = \frac{z_{Q,fl}}{(z_{|V|})^2} \quad (6)$$

To solve the ckt-SE, first, an overall circuit for the grid is constructed (see Fig. 2), replacing any measured component with an equivalent measurement circuit [13], [19]. Next, to obtain an MLE solution, [13] minimizes the norm-2 of the noise terms subject to the KCL-based network constraints of the equivalent circuit.

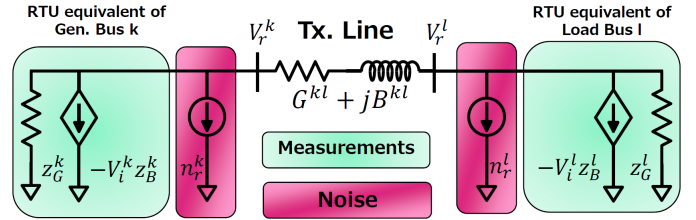


Fig. 2. Real sub-circuit of injection measurements-mapped equivalent circuit for the network in Fig. 1. The imaginary sub-circuit has a similar structure.

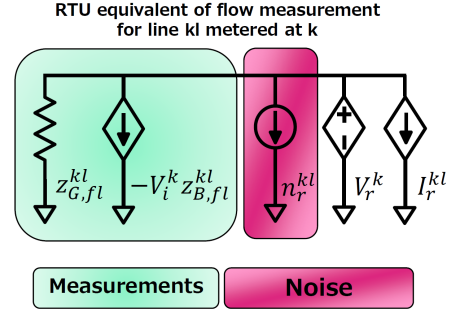


Fig. 3. Real sub-circuit of a line flow measurement-mapped equivalent circuit. The voltage V_r^k and V_i^k represent the real and imaginary part of the adjacent bus voltage at bus k , I_r^{kl} is explicitly represented by ECF branch models. The imaginary circuit has a similar structure.

Mathematically, the problem can be defined as follows:

$$\mathbf{P}_{\text{cktSE}} : \min_{\mathbf{v}, \mathbf{n}} f(\mathbf{n}) \quad (7a)$$

subject to:

$$h_{RTU}^k(\mathbf{v}, \mathbf{n}, \mathbf{z}) = 0 \quad \forall k \in \mathcal{N} \setminus \mathcal{ZI} \quad (7b)$$

$$h_{RTU,fl}^e(\mathbf{v}, \mathbf{n}, \mathbf{z}) = 0 \quad \forall e \in \mathcal{E} \quad (7c)$$

$$h_{ZI}^k(\mathbf{v}) = 0 \quad \forall k \in \mathcal{ZI} \quad (7d)$$

where the norm-2 minimization in objective has the following form:

$$f(\mathbf{n}) = \min_{\mathbf{v}, \mathbf{n}} \left(\sum_{k \in \mathcal{I}} \left((n_r^k)^2 + (n_i^k)^2 \right) + \sum_{e \in \mathcal{E}} \left((n_r^e)^2 + (n_i^e)^2 \right) \right) \quad (7e)$$

where \mathcal{I} is the set of injection measurements and ZI is the set of zero-injection buses. \mathcal{E} is the set of line flow measurements. \mathbf{v} is the vector of real and imaginary voltages and \mathbf{n} is the vector of noise terms. Note that *affine* function h represents IV-based nodal constraints for the equivalent circuit and consists of both measurement parameters (\mathbf{z}) and other parameters (e.g., line and transformer impedances). *Affine* flow measurement functions is included without loss of generality following [13].

Because the constraint set in (7) is affine, the underlying optimization problem is convex. It represents a domain-knowledge-driven MLE technique, which has a closed-form result. Furthermore, including zero-injection nodes naturally shrinks the feasible space to give a more physically meaningful estimation.

B. McCormick Convex Relaxation

The proposed method in [13] is originally a non-convex formulation with no guarantees on the solution quality. As such, we will apply a convex relaxation technique to quantify the *goodness* of solution quality. As the problem nonlinearities in our approach are bilinear, we will utilize McCormick relaxation, which was proposed in [20] and is widely applied in other grid optimization problems like the optimal power flow (OPF) [21] and [22].

The McCormick relaxation lifts all bilinear terms in the target equality constraints and represents them with a new lifted variable as in (8a). Subsequently, it defines the upper and lower bounds for the variables in the bilinear terms as shown in (8b). With upper and lower bounds defined, it replaces the individual bilinear terms with the set of inequality constraints (8c), also referred to as *McCormick Envelope*.

With this relaxation, the original nonlinear feasible space shown via the green curve in Fig. 4 is relaxed to a convex feasible set described via the area inside the boundary of the original bounds in purple. This relaxed feasible space depends on the bounds in (8b). Thus, the quality of such relaxation critically hinges on the careful and appropriate selection of the upper and lower bounds (see [21] for details).

$$\hat{x}_{ij} = x_i x_j \quad (8a)$$

$$\begin{aligned} x_i &\leq \underline{x}_i \leq \bar{x}_i \\ x_j &\leq \underline{x}_j \leq \bar{x}_j \end{aligned} \quad (8b)$$

$$\begin{aligned} \hat{x}_{ij} &\geq \underline{x}_i x_j + x_i \underline{x}_j - \underline{x}_i \underline{x}_j \\ \hat{x}_{ij} &\geq \bar{x}_i x_j + x_i \bar{x}_j - \bar{x}_i \bar{x}_j \\ \hat{x}_{ij} &\leq \underline{x}_i x_j + x_i \bar{x}_j - \underline{x}_i \bar{x}_j \\ \hat{x}_{ij} &\leq \bar{x}_i x_j + x_i \underline{x}_j - \bar{x}_i \underline{x}_j \end{aligned} \quad (8c)$$

C. Sequential Bound Tightening (SBT)

McCormick relaxation guarantees solution optimality but with a trade-off: loss of AC feasibility. The quality of upper and lower bounds used to construct the McCormick envelopes quantifies this trade-off in feasibility [16]. For McCormick relaxation to perform better (i.e., provide optimal solutions

without deviating significantly from the feasible region), we should seek the tightest possible bounds. Fig. 4 shows how the feasible space constricts when the bounds at two dimensions tighten to a quarter of their original values. In practice, the challenge lies in obtaining tightened bounds that contain the *global optimal* solution of the original NLP problem.

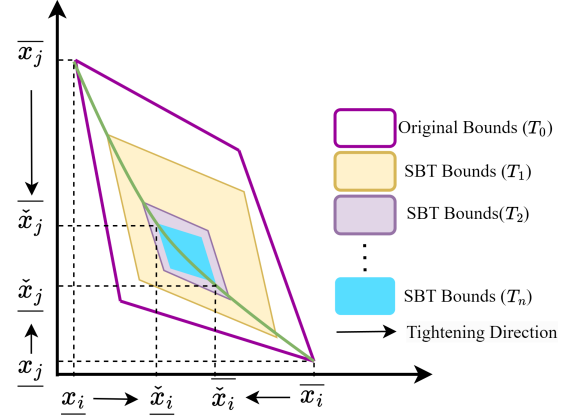


Fig. 4. Tightening McCormick Bounds iteratively. From T_0 to T_n SBT takes n iterations to find the tightened bounds. x_i shows the original bounds while \tilde{x}_i shows the tightened bounds. The arrow indicates the tightened direction during the SBT process.

We build upon earlier sequential bound technique (SBT) methods proposed in [14], [15], [16]. We specifically extend and apply the SBT algorithm proposed in [16] to McCormick implementation in this paper to achieve tightened bounds. When there is no apriori knowledge of good boundaries for a variable, this method can provide tight bounds. If there is apriori knowledge of good bounds for a variable, the improvement brought by applying SBT (or other similar method) is limited.

SBT technique inputs the best-known solution of the NLP problem with bilinear nonlinearities $\langle x_i x_j \rangle$ in the equality constraints and the *best-known* upper and lower bounds for the bilinear terms \bar{x}_i, \bar{x}_j and $\underline{x}_i, \underline{x}_j$. It outputs tightened bounds \tilde{x}_i, \tilde{x}_j and \tilde{x}_i, \tilde{x}_j through an iterative process of optimizing the upper and lower bounds for all variables in the bilinear terms till the improvement reaches a set tolerance as shown in the Fig. 4.

This approach of tightening bounds for each variable in the bilinear term can be computationally prohibitive, considering the number of parallel iterative optimizations we would have to perform. Therefore, in this work, we will tighten the bounds for only a subset of variables based on domain knowledge.

III. CKT-PSE: PROBLEM DESCRIPTION

Next, we design and formulate the circuit-theoretic joint parameter and state estimation (ckt-PSE) problem as an extension of the circuit-theoretic state estimation (ckt-SE). The approach estimates the system's state while also estimating the unknown parameters. The parameters of interest include unknown or erroneous network parameters, including line and transformer impedances, and shunt positions. We show that the

extension is non-trivial, transforming the original convex problem into a non-convex problem. We develop the methodology to obtain solutions with minimal trade-offs between solution optimality and AC feasibility.

A. Circuit-theoretic joint parameter-state estimation formulation (ckt-PSE)

Here, we propose and formulate a novel **circuit-theoretic formulation for joint AC parameter-state estimation (ckt-PSE)**. Akin to ckt-SE, in this approach, we first construct an equivalent circuit model of the power grid with measurements mapped as sub-circuits. However, unlike ckt-SE (see Fig. 2), where all parameters are assumed known and accurate, in ckt-PSE, we operate under the assumption that *certain* or *all* parameters as unknown or erroneous (see Fig. 5).

Therefore, to estimate values of erroneous or unknown parameters, we consider them as variables. In a manner of state-augmentation, the set of unknown parameters \mathcal{P} are included in the state variable set, resulting in a new expanded variable set \mathbf{x} , where $\mathbf{x} = \mathbf{v} \cup \mathcal{P} \in \mathbb{R}^{|\mathbf{v}| \cup |\mathcal{P}|}$. This will be the new variable set for ckt-PSE. Note that in ckt-PSE, while measurement noise \mathbf{n} represents a realization of random variables, the \mathcal{P} is only a set of unknown parameters, not random variables.

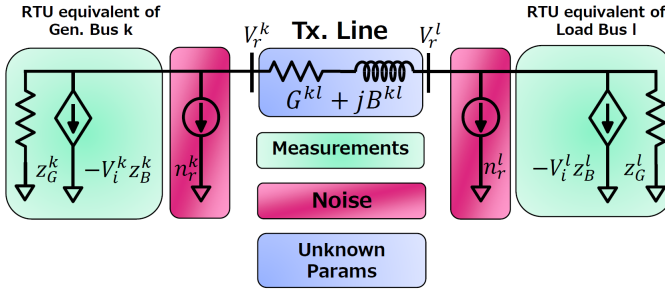


Fig. 5. Real sub-circuit of the grid model for ckt-PSE. Imaginary sub-circuit has a similar structure

We formulate the ckt-PSE optimization problem as $\mathbf{P}_{\text{cktPSE}}$ in (9). We find that the very existence of parameters as variables turns the convex $\mathbf{P}_{\text{cktSE}}$ into non-convex $\mathbf{P}_{\text{cktPSE}}$. The nonconvexity stems from converting affine equality constraints in (7b)-(7d) to those with bilinear terms in (9b), (9c), and (9d). We denote the bilinear terms with $\mathbf{s} \in \mathcal{B}$. In its most naïve form, the ckt-PSE has the following form:

$$\mathbf{P}_{\text{cktPSE}} : \min_{\mathbf{x}, \mathbf{n}, \mathbf{s}} f(\mathbf{n}) \quad (9a)$$

subject to:

$$h_{RTU}^k(\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{s}) = 0 \quad \forall k \in \mathcal{N} \setminus \mathcal{ZI} \quad (9b)$$

$$h_{RTU, fl}^e(\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{s}) = 0 \quad \forall e \in \mathcal{E} \quad (9c)$$

$$h_{ZI}^k(\mathbf{x}, \mathbf{s}) = 0 \quad \forall k \in \mathcal{ZI} \quad (9d)$$

$$s^k = v^i v^j \quad \forall k \in \mathcal{B} \quad (9e)$$

The norm-2 minimization in objective has the following form, with \mathcal{I} is the set of injection measurements, \mathcal{E} is the set of line flow measurements:

$$f(\mathbf{n}) = \sum_{k \in \mathcal{I}} w_I^k \left((n_r^k)^2 + (n_i^k)^2 \right) + \sum_{e \in \mathcal{E}} w_I^e \left((n_r^e)^2 + (n_i^e)^2 \right) \quad (9f)$$

where w_I^k and w_I^e are the weight for RTU injection node measurement k and the weight for line flow measurement e . The choice of weights is discussed in Section III-D. The RTU flow measurement functions $h_{RTU, fl}$ are explicitly included in (9), (also see Fig.3 and Section II-A for derivations). The only difference now is that certain parameters may be assumed unknown.

As some parameters are unknown variables in ckt-PSE, the feasible set for the underlying optimization is larger than ckt-SE. As such, there is value in constricting the feasible space with any additional system knowledge. Now, recall the feature transformation of RTU measurements. It results in a loss of information due to feature reduction from 3 measurements to 2 measurements when we retrieve z_G and z_B from z_P, z_Q and $z_{|V|}$. The proposed work will include lost $z_{|V|}$ information by having an additional bilinear voltage-based constraint for measured bus k :

$$(V_r^k)^2 + (V_i^k)^2 - (z_{|V|}^k - n_{|V|}^k)^2 = 0 \quad \forall k \in \mathcal{N} \setminus \mathcal{ZI} \quad (10)$$

with objective of $\mathbf{P}_{\text{cktPSE}}$ changing to include noise in voltage measurements n_V^k :

$$\min_{\mathbf{x}, \mathbf{n}, \mathbf{s}} \sum_{k \in \mathcal{I}} \left(w_I^k \left((n_r^k)^2 + (n_i^k)^2 \right) + w_V^k (n_V^k)^2 \right) + \sum_{e \in \mathcal{E}} w_I^e \left((n_r^e)^2 + (n_i^e)^2 \right) \quad (11)$$

In (10), the bilinear terms are sorted in \mathbf{s} and extra voltage equality constraints are added to the set $h_{RTU}(\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{s})$.

B. Unknown vs. erroneous parameter values

$\mathbf{P}_{\text{cktPSE}}$ can incorporate any apriori knowledge of parameters into the algorithm as a *best-known* value of the parameter $\hat{\mathcal{P}}$, changing the unknown parameter as:

$$\mathcal{P} = \hat{\mathcal{P}} + d\mathcal{P} \quad (12)$$

With this apriori knowledge, the network constraint for branch between k and l with unknown parameters G^{kl} and B^{kl} will change as follows:

$$I_r^{kl} = (\hat{G}^{kl} + dG^{kl})V_r^{kl} - (\hat{B}^{kl} + dB^{kl})V_i^{kl} \quad (13)$$

$$I_i^{kl} = (\hat{G}^{kl} + dG^{kl})V_i^{kl} + (\hat{B}^{kl} + dB^{kl})V_r^{kl} \quad (14)$$

With parameters' *best-known* values $\hat{\mathcal{P}}$ being constants, the binary terms in (9e) will be functions of $d\mathcal{P}$.

C. Applying multi-period sampling

While the grid states \mathbf{v} vary at every time-sample due to changing operating conditions such as changing loads, generation, and topology, the unknown parameters \mathcal{P} generally remain fixed for a *reasonable* period. Therefore, by including multi-period sampling into the constraint set, especially from widely varying conditions (high loading vs. low loading), we can *squeeze* the feasible space for unknown parameters, improving the quality of parameter estimates. The objective for multi-period sampling changes as follows:

$$\mathbf{P}_{\text{cktPSE}}^{\text{MP}} : \min_{\mathbf{x}', \mathbf{n}', \mathbf{s}'} \sum_{t \in \mathcal{T}} f(\mathbf{n}') \quad (15a)$$

The network constraints for multi-period sampling change as follows:

$$h_{RTU}^{k,t}(\mathbf{x}', \mathbf{n}', \mathbf{z}', \mathbf{s}') = 0 \quad \forall k \in \mathcal{N} \setminus \text{ZI} \quad \forall t \in \mathcal{T} \quad (15b)$$

$$h_{RTU,fl}^{e,t}(\mathbf{x}', \mathbf{n}', \mathbf{z}', \mathbf{s}') = 0 \quad \forall e \in \mathcal{E}, \quad \forall t \in \mathcal{T} \quad (15c)$$

$$h_{ZI}^{k,t}(\mathbf{x}', \mathbf{s}') = 0 \quad \forall k \in \text{ZI}, \quad \forall t \in \mathcal{T} \quad (15d)$$

$$\mathbf{s}'^k = v^i \mathcal{P}^j \quad \forall k \in \mathcal{B}' \quad (15e)$$

Note here that $\mathbf{x}', \mathbf{n}', \mathbf{s}'$ and \mathbf{z}' imply that these are concatenated vectors over multiple time periods. Also, that parameters \mathcal{P} remain fixed in the time-period set $\mathcal{T} = \{t_1, \dots, t_n\}$.

D. Choice of weight

The source of noise comes from meters' measurement accuracy. We have current measurement noise w_I^k and voltage measurement noise w_V^k in the proposed approach. Commonly the error tolerance is labeled on the meter as $\text{tol} : \%a + \%b$ or by class. w_I^k and w_V^k should be based on the precision of the equipment to reflect such differences in the objective function: we give larger weight to measurements that have better accuracy and vice versa, to show we trust some accurate meters more in our optimization problem.

E. Observability of the system

The proposed estimation problem, $\mathbf{P}_{\text{cktPSE}}$ is unobservable as the underlying system is overdetermined. Therefore, there is no one-to-one mapping between states and measurements in a noiseless scenario [1] for such a system.

When formulating this problem as an optimization problem, we obtain a nonconvex problem $\mathbf{P}_{\text{cktPSE}}$, with many possible solutions, including local minima and saddle points. This is because the constraint set includes nonlinearities. While we cannot enforce one-to-one mapping between states and measurements in a noiseless scenario without relaxing system equations, we strive to deliver *high-quality* estimations by quantifying optimality of the nonconvex $\mathbf{P}_{\text{cktPSE}}$ problem. To that effect, next we formulate and solve variants of $\mathbf{P}_{\text{cktPSE}}$.

IV. CKT-PSE: SOLUTION METHODOLOGY

In this paper, we develop three solution approaches for the ckt-PSE problem. The first solves the exact NLP problem but without any guarantees of optimality. The other two approaches solve the convex relaxation of the problem, balancing network feasibility with optimality. We show that the final approach is a tight relaxation, achieving optimality while enabling *sufficient* AC model fidelity.

A. Solving ckt-PSE exactly

To solve $\mathbf{P}_{\text{cktPSE}}$ exactly, we apply primal-dual interior point (PDIP) to (15). The method ensures high model fidelity by encapsulating AC network constraints; however, because of bilinear terms in the equality constraints, spurious solutions to local minima and saddle points are possible (see Fig. 9 in Numerical Results). Thus, next, we explore methods to address this challenge, by exploiting the bilinear nature of the nonlinearities and domain-based bound tightening.

B. McCormick relaxation of ckt-PSE

To quantify solution optimality, we relax the bilinear terms with McCormick bounds [20] to obtain a lower bound on the NLP. In ckt-PSE, bilinear terms combine real and imaginary node voltage variables $v \in \mathbf{v}$ and related unknown parameters $\mathcal{P} \in \mathcal{P}$ as shown in (9). By applying McCormick relaxation, each bilinear term introduces a set of inequality constraints as shown in (8b) to (8c): $g(\mathbf{x}, \mathbf{s})$. We term the McCormick relaxation of $\mathbf{P}_{\text{cktPSE}}$ as $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$:

$$\mathbf{P}_{\text{cktPSE}}^{\text{MC}} : \min_{\mathbf{x}, \mathbf{n}, \mathbf{s}} f(\mathbf{n}) \quad (16a)$$

subject to:

$$h_{RTU}^k(\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{s}) = 0 \quad \forall k \in \mathcal{N} \setminus \text{ZI} \quad (16b)$$

$$h_{ZI}^k(\mathbf{x}, \mathbf{s}) = 0 \quad \forall k \in \text{ZI} \quad (16c)$$

$$h_{RTU,fl}^e(\mathbf{x}, \mathbf{n}, \mathbf{z}, \mathbf{s}) = 0 \quad \forall e \in \mathcal{E} \quad (16d)$$

$$g^k(\mathbf{x}, \mathbf{s}) \leq 0 \quad \forall k \in \mathcal{B} \quad (16e)$$

$$\mathbf{s}^k = v^i \mathcal{P}^j \quad \forall k \in \mathcal{B} \quad (16f)$$

While McCormick re-formulation guarantees the global optimality of the relaxed problem, the solution quality in terms of AC feasibility (or simply feasibility gap) depends on the bounds used in McCormick envelopes. Without prior knowledge of parameter values, it is challenging to construct tight McCormick envelopes, and as such, we can obtain solutions that are far from AC feasible solution and lack accuracy as shown in Fig. 7. Therefore, in the following approach, we apply a *modified* sequential bound tightening algorithm, extending the original approach in [16] to obtain tight bounds for $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$.

C. Applying Sequential Bound Tightening (SBT) to ckt-PSE

We extend SBT in [16] to obtain a set of tight input bounds for $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$, crucial to creating a good relaxation through McCormick envelopes. The fundamental idea in bound tightening is to transition from *best-known* bounds $(\mathcal{P}^L, \mathcal{P}^U)$, to *tightest-feasible* bounds $(\tilde{\mathcal{P}}^L, \tilde{\mathcal{P}}^U)$. See Fig. 6 for a graphical illustration of SBT. To achieve this goal, SBT in [16] runs parallel batch optimizations to obtain tight bounds by minimizing the gap between bounds of individual variables in bilinear terms. While SBT is originally applied to each variable in bilinear terms set \mathbf{s} , when the number of unknown parameters and system size grows, this can be computationally prohibitive even though the method is embarrassingly parallel.

To improve computational performance, we only tighten a subset of the unknown variables: \mathcal{P} , where $\mathcal{P} \subset \mathbf{x}$. This is a key difference in our method. We acquire good bounds for bus voltages through the processing of prior SCADA backlog, thus we know voltage state variables inherently have a much tighter range between *best-known* upper and lower bounds. This is in contrast to *best-known* bounds for unknown parameters where applying SBT is critical. This modification significantly reduces the number of batch optimizations in SBT to just the number of unknown parameters. More importantly, as we only tighten the bounds for unknown parameters, including multi-period constraints results in significantly lower computational overhead as the number of parallel optimizations does not increase for the SBT algorithm, as the same parameter variables are shared across all time periods. Note, however, that the scale of individual optimization will increase due to additional temporal constraints.

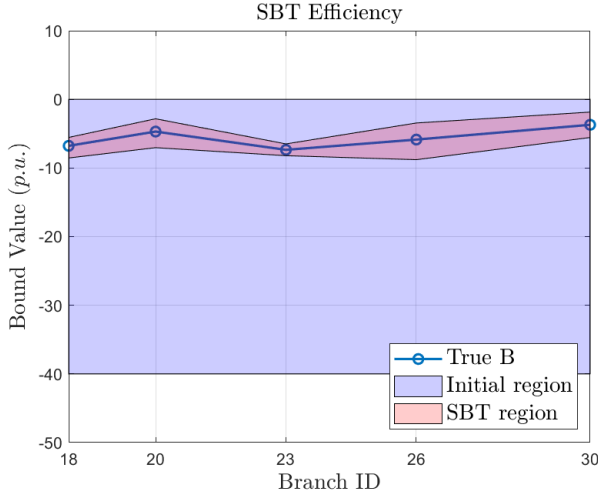


Fig. 6. An illustration of subset-only SBT on IEEE-118 test case with 5 unknown parameters (jB^{kl}); The original lower bounds were acquired by using 10 times the true value and upper bound from the definition of (jB^{kl}) being negative. Note the polyline connecting each bound is for the sake of showing improvement for each variable corresponding to the given branch on the x-axis, it does not imply any trend between bound and branch.

Algorithm 1 describes the SBT algorithm. We first solve $\mathbf{P}_{\text{cktPSE}}$ to obtain the lowest available objective for unrelaxed NLP problem $f(\mathbf{n}^*)$ and add it as a constraint to bound the

Algorithm 1 Modified SBT (Extension of Algorithm in [16])

- 1: **Input:** \mathcal{P} , \mathcal{P}^L , \mathcal{P}^U , \mathbf{n}^* (unknown parameter set, its corresponding best-known bounds, and NLP solution)
- 2: **Output:** $\tilde{\mathcal{P}}^L$, $\tilde{\mathcal{P}}^U$ (tightened bounds)
- 3: **Set** $\epsilon > 0$, Set tolerance to greater than 0
- 4: **Solve** in parallel $\forall \mathcal{P}^i \in \mathcal{P}$, $i \in \{1, \dots, |\mathcal{P}|\}$
- 5: **Set** $\underline{\mathcal{P}}^i \leftarrow \mathcal{P}^L$, $\bar{\mathcal{P}}^i \leftarrow \mathcal{P}^U$, $\mathcal{P}_{iter}^i = \bar{\mathcal{P}}_{iter}^i = 0$
- 6: **while** $\|\underline{\mathcal{P}}^i - \mathcal{P}_{iter}^i\|_2 > \epsilon$ and $\|\bar{\mathcal{P}}^i - \bar{\mathcal{P}}_{iter}^i\|_2 > \epsilon$ **do**
- 7: $\mathcal{P}_{iter}^i \leftarrow \underline{\mathcal{P}}^i$ and $\bar{\mathcal{P}}_{iter}^i \leftarrow \bar{\mathcal{P}}^i$
- 8: $\tilde{\mathcal{P}}^i := \min_{\mathbf{x}, \mathbf{n}, \mathbf{s}} \underline{\mathcal{P}}^i$
- 9: $\tilde{\bar{\mathcal{P}}}^i := \max_{\mathbf{x}, \mathbf{n}, \mathbf{s}} \bar{\mathcal{P}}^i$
- 10: **s.t.** $f_{\mathbf{x}, \mathbf{n}, \mathbf{s}}(\mathbf{n}) \leq f_{\mathbf{x}, \mathbf{n}, \mathbf{s}}(\mathbf{n}^*)$
- 11: $h(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{s}) = 0$
- 12: $g(\mathbf{x}, \mathbf{s}) \leq 0$
- 13: $s^k = v^i \mathcal{P}^j \quad \forall k \in \mathcal{B}$
- 14: $\underline{\mathcal{P}}^i \leftarrow \tilde{\mathcal{P}}^i$ and $\bar{\mathcal{P}}^i \leftarrow \tilde{\bar{\mathcal{P}}}^i$ (update \mathcal{P}^L and \mathcal{P}^U)
- 15: **end while**
- 16: All final parameter bounds packed in $\tilde{\mathcal{P}}^U$ and $\tilde{\mathcal{P}}^L$

objective of SBT optimizations in Algorithm 1. While bounding the objective, we run in parallel, an optimization for each parameter's lower and upper bounds. The algorithm concludes by providing new tighter bounds to apply to McCormick relaxation $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$. Here, the problem has the same structure as $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$ but with better-defined (tighter) bounds from SBT. Thus, better relaxation performance (smaller variance and lower AC violation) is expected from such a processed problem.

Remark 1: Both McCormick formulations, with or without SBT, are compatible with multi-period ($\mathbf{P}_{\text{cktPSE}}^{\text{MP}}$), without loss of generality. But as the number of time periods increases, the number of bilinear terms will increase linearly. However, in the extended SBT approach in Algorithm 1, where we perform batch optimizations for tightening the bounds for unknown parameters only, no additional computational overhead for multi-period constraints will incur because the same unknown parameters are shared across all temporal constraints.

Remark 2: SBT can have a large computation burden because the algorithm requires solving many convex optimizations, as many as the number of unknown parameters. However, as we apply SBT to only unknown parameters, which are typically stable and do not change once calculated, the SBT bounds can be used without recalculation for all subsequent estimations until there is a topology change or adjustment to data, such as transformer taps.

V. EXPERIMENTAL SETUP

To demonstrate the efficacy of the proposed algorithms, we will run $\mathbf{P}_{\text{cktPSE}}$, $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$, and $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ on four cases. Case 1, which represents a 14-bus network, assumes all branch parameters are unknown. Case 2, an IEEE 118 bus network, better illustrates the effectiveness of $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ and the corresponding use of McCormick envelopes. Case 3 represents a larger network with 2869 nodes and demonstrates some weaknesses of the $\mathbf{P}_{\text{cktPSE}}$ approach. Case 4, also built off

TABLE I
EXPERIMENT SET UP

| Scenario | Test Case | Unknown B Id# | Unknown B_{sh} Id# |
|----------|------------|-------------------------|----------------------|
| Case 1 | IEEE-14 | all | NA |
| Case 2 | IEEE-118 | 5, 52, 54, 84, 103, 169 | NA |
| Case 3 | 2869pegase | 175, 1725 | NA |
| Case 4 | IEEE-118 | 18, 20 | 4, 25, 37, 90 |

The id is the branch or shunt id in the .raw file for which B is unknown.

the IEEE 118 bus network includes unknown shunt parameters. Details for the cases are given in Table I.

A. Data creation and Experiment setup

For each case, we run 100 instances of parameter-state estimation algorithms ($\mathbf{P}_{\text{cktPSE}}$, $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$, and $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$). We assume RTU injection measurements at all nodes, excluding zero-injection nodes, ($\forall k \in \mathcal{N} \setminus \text{ZI}$). We also assume 2 line flow RTU measurements in test case 1 and 10 in test case 2, we chose the location of these measurements arbitrarily. To generate synthetic measurement data for each instance of each case, we run power flow simulations and add noise to the results. The noise samples are obtained based on a standard deviation value of 0.001 [13] for both current and voltage measurements. For unknown parameters, we assume no prior knowledge of the magnitudes and, therefore, use initial bounds for McCormick envelopes as large as 1000%. The number is chosen to be larger than what's commonly used in other literature [23]. We are designing the overall experiment to answer the following questions:

- 1) **Robustness of the NLP approach $\mathbf{P}_{\text{cktPSE}}$:** Does the NLP approach generally produce robust and accurate estimates when given a good initial guess? Does the property hold when a good initial guess is unavailable?
- 2) **The improvement of $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ compared to $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$:** Does the use of SBT improve the parameter bounds, and subsequently improve the quality of state and unknown parameter estimates?
- 3) **The scalability of the proposed algorithm:** Do the methods scale when the system size and the number of unknown parameters increase?

To illustrate our findings and compare the various estimation techniques, we use the following error metrics RMSE, NRMSE and variance (σ^2):

$$RMSE_{\mathbf{v}} = \sqrt{\frac{\sum_{i=0}^n \sum_{j=0}^{|\mathbf{v}|} (v_{est}^{i,j} - v_{true}^{i,j})^2}{n|\mathbf{v}|}} \quad (17)$$

$$NRMSE_{\mathbf{v}} = \frac{RMSE_{\mathbf{v}}}{\bar{v}_{est}} \quad (18)$$

$$\sigma_{v,j}^2 = \frac{\sum_{i=0}^n (v_{est}^{i,j} - \bar{v}_{est}^{i,j})^2}{n} \quad \forall j \in 1 \dots |\mathbf{v}| \quad (19)$$

$$\sigma_{v,ave}^2 = \frac{\sum_{j=0}^{|\mathbf{v}|} \sigma_{v,j}^2}{|\mathbf{v}|} \quad (20)$$

where v_{est} is the estimate of state v . Note the RMSE and variance for parameters (NRMSE $_{\mathcal{P}}$ and $\sigma_{p,avg}^2$) are calculated in the same manner as state variables v .

In another metric, percentage estimation error, we compare the estimated value of parameters to the true value. We use this metric to show an absolute comparison of estimates from $\mathbf{P}_{\text{cktPSE}}$, $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$, and $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$, and define it for a given estimate y_{est} and its true value y_{true} , by:

$$\text{Estimation Error Coefficient}(\%) = \frac{(y_{est} - y_{true})}{y_{true}} 100 \quad (21)$$

In this experiment, we also assume the same accuracy for various RTU meters (i.e., $\forall k \in \mathcal{R}, w_I^k, w_V^k$ are uniform). The user can choose more realistic weights with access to better precision and tolerance data for meters.

Experiment Environment and Codebase : Our code and data are publicly available at <https://github.com/Asang97/cktPSE.git>. The proposed algorithm is implemented in Python with network data in .raw format. All simulations are processed on a PC with an AMD Core Razer 3 3100 CPU and 16 GB of RAM. We used open-source optimizer IPOPT (version 3.14.10) for all optimizations.

VI. NUMERICAL RESULTS

A. Accuracy and Robustness

Table II illustrate the performance of $\mathbf{P}_{\text{cktPSE}}$, $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$, and $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ for four (4) different test cases, each of them run with 100 samples of noise. In most cases, the non-convex $\mathbf{P}_{\text{cktPSE}}$ approach performs best among the three methods. However, the certificate of optimality cannot be obtained for $\mathbf{P}_{\text{cktPSE}}$, and when parameters are truly unknown, this can be a challenge, as performance is hard to gauge.

Variance metric for Outliers: We use both variance and NRMSE as metrics to evaluate the quality of our estimation. Specifically, in case 3 for NLP and cases 2 and 3 for McCormick, high variance shows low-quality estimates in many instances. We observe the high variance is due to the use of wide bounds in McCormick in $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$ and rare convergence to local solution for $\mathbf{P}_{\text{cktPSE}}$.

B. SBT tighten efficiency

While $\mathbf{P}_{\text{cktPSE}}$ does not guarantee solution optimality, on the other hand, McCormick relaxation guarantees optimality but with a loss of AC feasibility. Both Fig. 7 and Fig. 8 show that the estimation quality deteriorates with the use of $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$. We posit this is due to using wide bounds for McCormick envelopes. $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ addresses this problem by tightening the McCormick bounds. As an example, Fig. 6 shows the bound tightening of various branch parameters in case 2 with SBT. We observe the bound gap decreased significantly from the initial region to the SBT region.

Fig. 7 shows the error coefficients of estimations of case 2. By comparing the median of each box, we see improvement in the error coefficient from $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$ to $\mathbf{P}_{\text{cktPSE}}^{\text{SBT}}$ in most parameters. Fig.8 illustrates in both cases, $\mathbf{P}_{\text{cktPSE}}^{\text{MC}}$ shows

TABLE II
ESTIMATION ACCURACY EVALUATION

| Algorithm/Case | NRMSE _v | $\sigma_{v,avg}^2$ | NRMSE _p | $\sigma_{p,avg}^2$ |
|----------------------|--------------------|--------------------|--------------------|--------------------|
| $P_{cktPSE}/1$ | 6.30E-04 | 9.50E-06 | 0.05 | 0.12 |
| $P_{cktPSE}^{MC}/1$ | 8.60E-03 | 7.10E-05 | 0.14 | 0.02 |
| $P_{cktPSE}^{SBT}/1$ | 9.65E-04 | 9.26E-07 | 0.03 | 0.05 |
| $P_{cktPSE}/2$ | 2.10E-04 | 8.40E-05 | 1.50E-05 | 3.30E-06 |
| $P_{cktPSE}^{MC}/2$ | 1.10E-03 | 8.50E-05 | 7.70E-02 | 11.08 |
| $P_{cktPSE}^{SBT}/2$ | 7.50E-03 | 1.30E-04 | 3.40E-02 | 1.55 |
| $P_{cktPSE}/3$ | 3.00E-02 | 8.90E-04 | 0.10 | 69.37 |
| $P_{cktPSE}^{MC}/3$ | 2.80E-02 | 5.70E-04 | 0.12 | 15.45 |
| $P_{cktPSE}^{SBT}/3$ | 1.90E-02 | 1.20E-05 | 0.07 | 2.20E-06 |
| $P_{cktPSE}/4$ | 2.40E-04 | 8.40E-03 | 6.70E-03 | 1.30E-04 |
| $P_{cktPSE}^{MC}/4$ | 1.60E-03 | 8.60E-05 | 0.21 | 1.10E-03 |
| $P_{cktPSE}^{SBT}/4$ | 1.10E-03 | 8.50E-05 | 0.07 | 1.80E-03 |

1. Results in blue indicate higher bad parameter estimate instances due to wide McCormick bounds
2. Results in red indicate a high number of bad parameter estimates due to sub-optimal solutions in NLP

TABLE III
ALGORITHM FEATURE SUMMARY

| Algorithm | Optimality | AC Feasibility |
|--------------------|------------|----------------|
| P_{cktPSE} | × | ✓ |
| P_{cktPSE}^{MC} | ✓ | × |
| P_{cktPSE}^{SBT} | ✓ | × |

Optimality property describes whether the algorithm is guaranteed to converge to the global optima, and the AC feasibility property describes whether the algorithm satisfies the AC KCL-based network constraints within a certain tolerance.

some off-target estimation due to its wide McCormick bounds, and when the bound is tightened, we can observe P_{cktPSE}^{SBT} estimation being more accurate while P_{cktPSE} is providing effective estimations. Combining the results shown in table II, despite that the P_{cktPSE} guarantees no global optima, it lies close to the true value. When it comes to the convex relaxation, both P_{cktPSE}^{MC} and P_{cktPSE}^{SBT} results have some offset compared to P_{cktPSE} while in all P_{cktPSE}^{SBT} performs better. P_{cktPSE}^{SBT} deliver a better estimation of parameters with a smaller NRMSE and more AC feasible results as shown in Table II.

C. Sub-optimal solutions from NLP P_{cktPSE}

P_{cktPSE} generally has the best performance when a good initial condition is available (e.g. when the previous time-step estimate is used as an initial guess for the current time-step). Nonetheless, convergence to non-optimal solutions can result in *bad* estimates at times. See, for example, Fig. 9, which shows the estimation quality of two unknown parameters in a 2869-node network. Of 100 samples, P_{cktPSE} generally retrieves the unknown parameters of two lines with almost no error. However, in rare cases, it converges to local minima or saddle points, resulting in *bad* estimates, as observed in the figure's outliers. The likelihood of sub-optimal solutions (i.e., local minima or saddle points) is higher when good initial

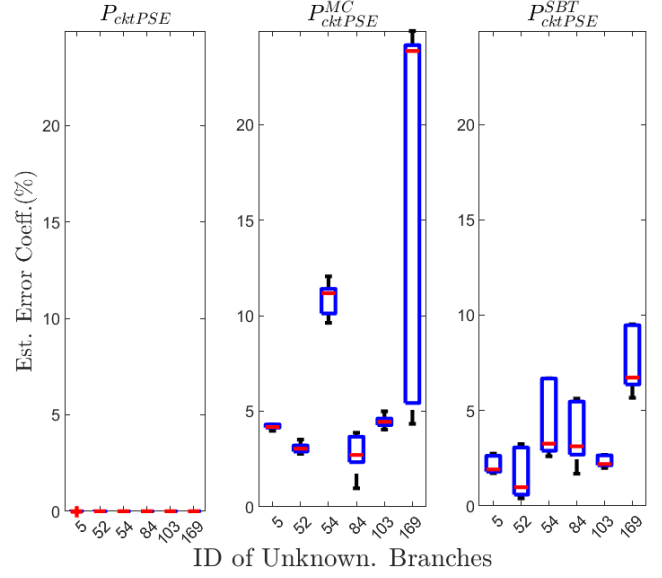


Fig. 7. The estimation error comparison between P_{cktPSE}^{MC} and P_{cktPSE}^{SBT} of case 2. The box indicates the lower and upper quartile of the estimation with a red bar showing its median. Starched whiskers showing maximum and minimum. In general, P_{cktPSE}^{SBT} performs better than P_{cktPSE}^{MC} . P_{cktPSE} performs the best.

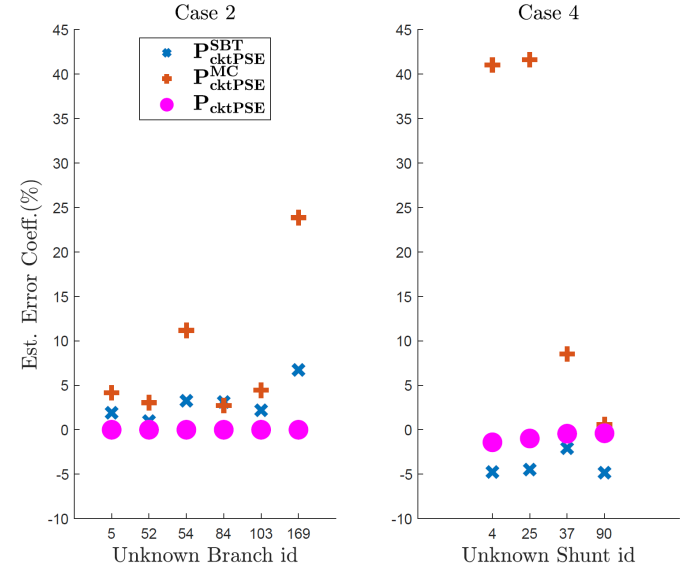


Fig. 8. Method-wise Parameter Estimation Error Comparison for case 2 and 4. Each point is the Error Coefficient between the true value and group median of est. parameters.

conditions are unavailable (see Fig. 10). Using the P_{cktPSE}^{SBT} algorithm can overcome the challenge with P_{cktPSE} . However, it comes with a tradeoff in terms of increased average error.

D. Scalability of the proposed algorithm

We study the scalability of ckt-PSE by documenting the:

- 1) Solving time for networks with varying numbers of system nodes (2% parameter unknown)
- 2) Solving time as the number of unknown parameters increases for the 2869 nodes network

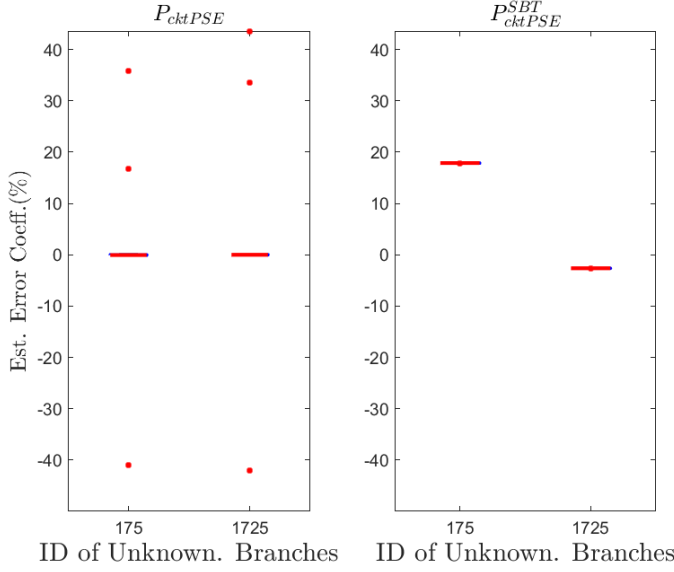


Fig. 9. The estimation error comparison between P_{cktPSE} and $P_{\text{SBT}_{\text{cktPSE}}}$ of case 3. The points indicate outliers (due to convergence to local minima or saddle points), and the red bar indicates the group median. The objective values for three outliers in the left subplot range from 64.86 to 278.83; in comparison, the objective values for on-target estimations range from 0.31 to 0.33. This indicates outliers may have resulted from convergence to local optima or saddle points as all noise is randomly sampled from the same non-gross Gaussian distribution.

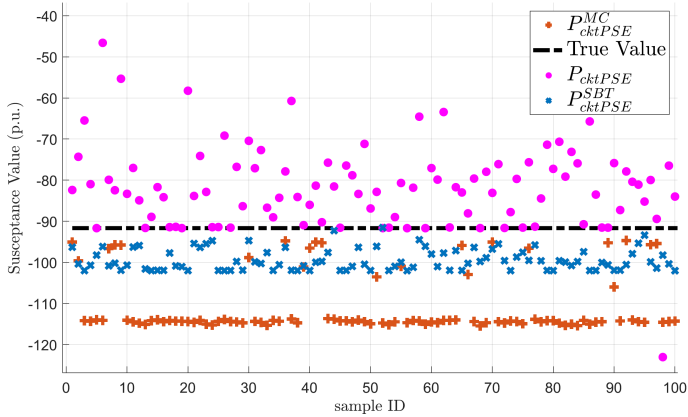


Fig. 10. Branch 169 parameter estimate from 100 runs of all 3 cktPSE parameter estimation algorithms. The initial conditions for these algorithms were randomly chosen from $|v_i| \in [0.95p.u., 1.05p.u.]$ $\angle v_i \in [-30^\circ, 30^\circ]$. All other unknown branches have a pattern similar to this figure.

Fig. 11 shows the time complexity of algorithms with increasing system size. For the 10k node network (ACTIVSg10k), P_{cktPSE} takes 30 iterations to converge, and $P_{\text{SBT}_{\text{cktPSE}}}$ takes 38 iterations to converge. While for the 2869 node network, P_{cktPSE} takes 16 iterations to converge, and $P_{\text{SBT}_{\text{cktPSE}}}$ takes 25 iterations to converge. Note that the computational complexity of bound tightening in SBT is not included in Fig. 11 as it is only performed once. We observe an almost linear increase in simulation time with the increasing number of unknown parameters. For the 2869 system, the NLP approach's solution time almost increased by 50% as the number of unknown parameters grew from 0 to 80.

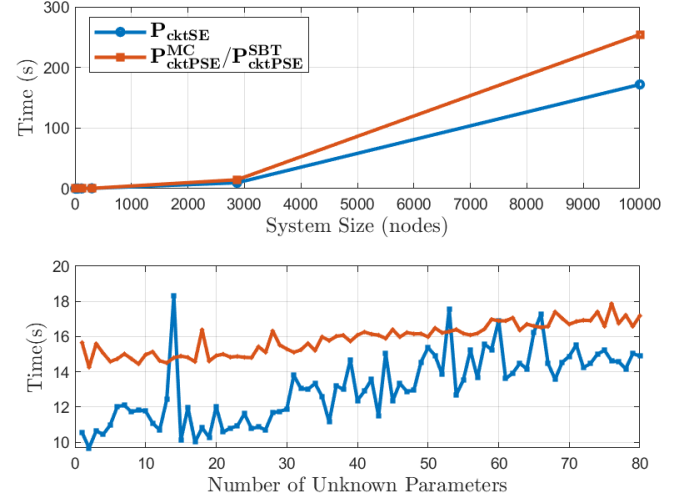


Fig. 11. Method scalability in terms of the number of nodes and the number of parameters.

VII. CONCLUSIONS

We developed a novel circuit-theoretic framework for joint parameter-state estimation (ckt-PSE). Due to the circuit mapping of the grid constraints, ckt-PSE is less nonlinear, including only bilinear terms. To balance the trade-off in optimality and AC feasibility (see Table III), we propose three algorithms for ckt-PSE: an exact NLP (P_{cktPSE}) and convex relaxations ($P_{\text{MC}_{\text{cktPSE}}}$, $P_{\text{SBT}_{\text{cktPSE}}}$). We draw these final conclusions:

- P_{cktPSE} approach performs the best in most scenarios when good initial conditions are available (e.g. good estimate from prior time-step); however, it cannot provide the certificate of optimality and exhibits sub-optimal solutions in some scenarios when good initial conditions are unavailable.
- $P_{\text{SBT}_{\text{cktPSE}}}$ presents a convex relaxation, an improvement on the naive McCormick approach, presenting a balance between optimality and AC feasibility.
- With domain knowledge, we reduce the problem complexity of the SBT algorithm, by tightening only a subset of bilinear terms (i.e. unknown parameters \mathcal{P})

As a final note, our experiments showcase the ability to recover unknown states and parameters with a quantifiable level of accuracy for realistic-sized networks.

REFERENCES

- [1] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.
- [2] H. M. Merrill and F. C. Schweppe, "Bad data suppression in power system static state estimation," *IEEE Transactions on Power Apparatus and Systems*, no. 6, pp. 2718–2725, 1971.
- [3] L. Zanni, A. Derviskadić, M. Pignati, C. Xu, P. Romano, R. Cherkaoui, A. Abur, and M. Paolone, "Pmu-based linear state estimation of lausanne subtransmission network: Experimental validation," *Electric Power Systems Research*, vol. 189, p. 106649, 2020.
- [4] A. Jovicic and G. Hug, "Linear state and parameter estimation for power transmission networks," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2020, pp. 1–6.

- [5] M. Vanin, F. Geth, R. D'hulst, and D. Van Hertem, "Combined unbalanced distribution system state and line impedance matrix estimation," *International Journal of Electrical Power & Energy Systems*, vol. 151, p. 109155, 2023.
- [6] F. C. Schweppe and J. Wildes, "Power system static-state estimation, part i: Exact model," *IEEE Transactions on Power Apparatus and systems*, no. 1, pp. 120–125, 1970.
- [7] H. Zhao, "A new state estimation model of utilizing pmu measurements," in *2006 International Conference on Power System Technology*. IEEE, 2006, pp. 1–5.
- [8] L. F. Costa, J. S. Giraldo, and C. A. Castro, "Identification and correction of transmission line parameter errors using scada and synchrophasor measurements," *International Journal of Electrical Power & Energy Systems*, vol. 135, p. 107509, 2022.
- [9] Y. Weng, R. Negi, C. Faloutsos, and M. D. Ilić, "Robust data-driven state estimation for smart grid," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1956–1967, 2016.
- [10] L. Zhang, G. Wang, and G. B. Giannakis, "Real-time power system state estimation and forecasting via deep unrolled neural networks," *IEEE Transactions on Signal Processing*, vol. 67, no. 15, pp. 4069–4077, 2019.
- [11] Y. Weng, Q. Li, R. Negi, and M. Ilić, "Semidefinite programming for power system state estimation," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.
- [12] Y. Zhang, R. Madani, and J. Lavaei, "Conic relaxations for power system state estimation with line measurements," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1193–1205, 2017.
- [13] S. Li, A. Pandey, S. Kar, and L. Pileggi, "A circuit-theoretic approach to state estimation," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. IEEE, 2020, pp. 1126–1130.
- [14] C. Chen, A. Atamtürk, and S. S. Oren, "Bound tightening for the alternating current optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3729–3736, 2016.
- [15] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, "Strengthening convex relaxations with bound tightening for power network optimization," in *International conference on principles and practice of constraint programming*. Springer, 2015, pp. 39–57.
- [16] H. Nagarajan, M. Lu, E. Yamangil, and R. Bent, "Tightening mcmormick relaxations for nonlinear programs via dynamic multivariate partitioning," in *Principles and Practice of Constraint Programming: 22nd International Conference, CP 2016, Toulouse, France, September 5-9, 2016, Proceedings 22*. Springer, 2016, pp. 369–387.
- [17] A. Jovicic, M. Jereminov, L. Pileggi, and G. Hug, "Enhanced modelling framework for equivalent circuit-based power system state estimation," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3790–3799, 2020.
- [18] A. Jovicic and G. Hug, "Linear state estimation and bad data detection for power systems with rtu and pmu measurements," *IET Generation, Transmission & Distribution*, vol. 14, no. 23, pp. 5675–5684, 2020.
- [19] A. Jovicic, M. Jereminov, L. Pileggi, and G. Hug, "A linear formulation for power system state estimation including rtu and pmu measurements," in *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. IEEE, 2019, pp. 1–5.
- [20] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems," *Mathematical programming*, vol. 10, no. 1, pp. 147–175, 1976.
- [21] M. S. Javadi, C. S. Gouveia, L. M. Carvalho, and R. Silva, "Optimal power flow solution for distribution networks using quadratically constrained programming and mcmormick relaxation technique," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2021, pp. 1–6.
- [22] M. Bynum, A. Castillo, J.-P. Watson, and C. D. Laird, "Tightening mcmormick relaxations toward global solution of the acopf problem," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 814–817, 2018.
- [23] A. Jovicic, "Computationally efficient state estimation for power systems with conventional and synchrophasor measurements," Ph.D. dissertation, ETH Zurich, 2021.