Performance Guaranteed State Estimation for Renewable Penetration with Improved Meters

Y. Weng\textsuperscript{1} J. Yu\textsuperscript{2} R. Rajagopal\textsuperscript{3}

\textsuperscript{1}School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ, United States
\textsuperscript{2}Department of Electrical Engineering, Stanford University, Stanford, CA, United States
\textsuperscript{3}Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, United States

E-mail: yang.weng@asu.edu

Abstract: State estimation (SE) aims at monitoring the transmission and distribution networks to achieve stable and reliable grid operations. While the SE problem is nonconvex, a local search method is used to achieve global optimum based on the belief that power system states are with limited variation in a short time scale. However, the recent and rapid deployment of renewables leads to strong power and state fluctuations in power grids, making local search method prone to local optimums with large estimation errors. In this paper, we propose to analyze SE with small measurement noises because highly accurate smart sensors are deployed in the past few years. By exploring a special structure of the performance metrics, we reformulate the SE problem in an extended state space. The new formulation is convex under the no-measurement-noise assumption. In order to use such a method when there are noises, we prove that a perturbation of globally optimal solution is asymptotically bounded by the measurement noise level. This prevents local optimums, which can create a large estimation error. Simulation results demonstrate that our method has a better performance compared to both WLS and the recent semidefinite programming-based approaches, especially when the noise is small.

1 Introduction

As one of the most significant infrastructures in the human society, the electric power grid not only provides energy in a flexible and convenient form to industrial and individual users but also supplies electricity in a clean and relatively easy way to transmit. To provide the service, grid monitoring is essential for grid operation tools, such as optimal power and contingency analysis. Schweppe formulated in [1] the monitoring problem as a state estimation (SE) problem, which plays the key role in today’s power grid operations.

The Weighted Least Square (WLS) criterion is commonly used as a metric for assessing the accuracy of SE. For example, [3] analyzes SE with equality constraints. Reference [4] uses WLS form in SE to improve numerical stability. A blocked sparse formulation is utilized in [5] with a WLS objective for an equality constrained SE. A similar idea can also be seen in [6] using augmented blocked matrices. Reference [7] uses WLS to conduct observability analysis and bad data processing for SE. Reference [2] unifies the WLS for both analog and digital quantities such as bus voltages and switch statuses. As a comparison, [8] conducts an analysis of different SE methods with a WLS objective.

However, this metric is nonconvex due to nonlinear power flow equations. The nonconvexity limits a solver’s capability to find the solution without approximation. This makes the popular solver based on Newton’s method reach a local optimum [9]. Luckily, as traditional power grids have limited variability in a short time scale, a state estimate obtained two minutes ago can be used to initialize Newton’s method in the hope of generating a global optimum.

Unfortunately, such a starting point will no longer be reliable due to the recent deployment of small size residential and large-scale commercial photovoltaic (PV) systems, which lead to strong power fluctuations in the distribution grid and the transmission grid. Moreover, solar generation will scale rapidly and potentially generate up to 14% of the nation’s total electricity demand by 2030 and 27% by 2050 [10, 11]. With such high penetration of intermittent PV generation, states shift at a fast speed, making the initial guess obtained two minutes ago unreliable. In addition to this drawback, the current initial guess-based method has no guarantee over the proximity between the estimated state and the global optimum.

To solve the nonconvex problem, extended state space analyses are applied recently. For example, the semidefinite programming-based approach aims to convexify the SE problem [12–17]. Specifically, [12, 13] use semidefinite programming (SDP) for SE and [14] extends the centralized algorithm to have distributed computational ability. Reference [15] compares the SDP method with another extended state space method. [16, 17] analyze the relationship between the SDP-based load flow problem and the SDP-based state estimation problem.

Unfortunately, there is a rank-one condition of a state matrix, which is not usually satisfied. For this reason, [18] looks into the power flow equations’ feasibility boundary for state estimation. Reference [19] examines the limits when applying SDP to powerflow-based applications. There are also other proposed methods such as direct linearization. For example, [20] uses direct non-iterative SE. Reference [21] uses a factorized approach for WLS SE. [22] extends the idea and uses a bilinear approach for SE. But they suffer from the voltage recovery problem like in the SDP-based method.

As more accurate sensors have been deployed, we propose to focus our SE study on a small measurement noise scenario in place of the one with normal noise. This is motivated by [23] from the area of signal processing, where blind constant modulus equalization is proposed via convex optimization. In our SE, nonlinearity is the main challenge. Instead of trying to handle both nonlinearity and noise at the same time, we aim at first understanding the noiseless setting via utilizing structural benefits of the WLS formulation [24]. By constructing the SE problem in a superposed quadratic form, we obtain a globally optimal solution by performing an algebraic transformation. Note that, the feasibility problems between SE problem and power flow problem are different in that SE is always feasible as it is conducted only when the system is running.

After proposing a numerically effective algorithm in the noiseless case, we study its performance when applying such an algorithm to true systems when noise is unavoidable. Naturally, when we add
relatively small noises to system measurements, the perturbation of global optimum in voltage states will be small in our optimization problem. To quantify the change, an asymptotic upper bound is proved based on the signal-to-noise ratio (SNR), the system size, and the number of measurements. Finally, while motivated by [23], we develop the optimization specifically for power systems. The proof on asymptotic bound is also new based on power system understanding.

Finally, we organize the paper in the following: in Section 2, we conduct an SE review including the recent work on convexification. Then, a different method is proposed in Section 3 leading to the optimal result when measurements are noiseless. With noisy measurements, we prove in Section 4 that a relatively small perturbation in noise only creates a small change in the globally optimal state estimate for the proposed method. In Section 5, numerical results are used to test the proposed method via IEEE test systems. Section 6 summarize the paper by emphasizing our contributions.

2 Review of State Estimation

2.1 Weighted Least Squares (WLS)-based State Estimation

The physical model for state estimation is [9]

\[ z_i = h_i(v) + u_i, \tag{1} \]

where the vector \( v = (|v_1|e^{j\delta_1}, |v_2|e^{j\delta_2}, \ldots, |v_{30}|e^{j\delta_{30}})^T \) is the complex phasor voltage state for estimation. The noise added onto the \( i^{th} \) measurement is represented by \( u_i \). These measurement noises are assumed to be independent zero-mean Gaussian random variable. Specifically, \( u_i \sim N(0, \Sigma) \). Within the diagonal matrix \( \Sigma \), we have \( \sigma_i^2 \) as the \( i^{th} \) diagonal component. \( z_i \) is the \( i^{th} \) telemetered measurement and \( h_i(.) \) describes measurement \( z_i \) with respect to \( v \) in a nonlinear relationship.

A state estimation problem looks for the optimal state estimate \( (\hat{v}) \) according to collected measurements and the physical system (1). Such a goal is realized by the following Weighted Least Square optimization.

\[ \min \sum_{i=1}^{m} \left( \frac{z_i - h_i(v)}{\sigma_i} \right)^2. \tag{2} \]

2.2 Initial Guess Problem with WLS-based Approach

Intermittent renewable generation causes strong power and state fluctuations. A resulting question is whether we can still use the initial guess-based method, e.g., using the state estimate 2 minutes ago to initialize Newton’s method. In this section, we examine the question numerically. To motivate the setup, we refer to James D. McCalley’s notes on power flow equations [25]. In [25], the observation 2 on page 3 states that, for most typical operating conditions, the difference in voltage angles between two buses \( k \) and \( j \) connected by a circuit \((\theta_k - \theta_j)\) is less than 10 to 15 degrees. It is extremely rare to see such angular separation exceeding 30 degrees. Therefore, we generate initial guesses according to

- voltage magnitude: randomly picked from [0.95, 1.05],
- voltage phase angle: randomly picked from \([-10^\circ, 10^\circ]\).

So, the initial guesses of voltage magnitudes can have a difference as large as 10%. The phase angle difference can be as large as 20\(^\circ\). In Fig. 1, we show simulation results on the IEEE 300-bus system.

The \( x \)-coordinate of Fig. 1a and Fig. 1b represents the testing case number. For each testing case, the \( y \)-coordinate represents the voltages of different buses. Although the initial guesses behind different testing cases are different, the voltage result should be the same if each initial guess is making Newton’s method converge towards the global optimum. Then, we will observe flat lines, which represent the same estimation result. However, for the various testing cases,

\[ \min f(x) \]

\[ = \min \sum_{i=1}^{m} \left( \frac{z_i - h_i(v)}{\sigma_i} \right)^2 \geq \min \sum_{i=1}^{m} \left( \frac{z_i - x^TY_i x}{\sigma_i} \right)^2. \tag{3} \]

\[ \text{Fig. 1:} \text{ 300-bus: The same color represents the estimated result of the same bus with different initial guesses.} \]

\[ \text{Fig. 1a shows that the voltage magnitudes have significant changes when initialized differently. Here, the curve with the same color represents the estimate on the same bus. Fig. 1b also shows that, for the same bus with the same measurement set, phase angle estimates can have dramatically different results when the iteration starts from different starting points. As we observe state fluctuations, they indicate that Newton’s method-based state estimate is sensitive to the initial guess. This calls for a more robust state estimation method for the newly added and planned renewable generations.} \]
For detailed formulation of above, please refer to [13]. Therefore, (3) reformulates the WLS in a quartic polynomial form. \( \tilde{Y}_i \) is a positive semidefinite matrix determined only by the admittance matrix [26].

Thus, [12] and [13] use \( W \triangleq x^T \tilde{Y}_i x \) in state estimation because \( x^T \tilde{Y}_i x \) is a scalar and \( x^T \tilde{Y}_i x = \text{trace}(x^T \tilde{Y}_i x) = \text{trace}(x x^T \tilde{Y}_i) = \text{trace}(W \tilde{Y}_i) = \text{trace}(\tilde{Y}_i W) \). Therefore, the new method conducts optimization in a lifted state space. This makes the nonlinearity in (1) become linear. Thus, \( z_i = \text{tr}(\tilde{Y}_k W) + u_i \) is linear in \( W \), where \( \tilde{Y}_i \) is a new “admittance” matrix in the rectangular form [13]. Therefore, (3) is converted into (4).

\[
\begin{align*}
\min_W J_2(W) &= \frac{1}{\sigma_i} \sum_{i=1}^m \left( z_i - \text{tr}(\tilde{Y}_i W) \right)^2 \\
\text{subject to} \quad W &\succeq 0, \quad \text{rank}(W) = 1.
\end{align*}
\]

Unfortunately, the new formulation above needs the new state variable \( W \) to have some special properties: positive semidefinite and rank-one. Only with these two constraints satisfied, the estimated \( \hat{W} \) can have a one-to-one mapping back to the \( v \). [19] and [27] show that this condition may not always be satisfied.

3 State Estimation when There is No Measurement Noise

Past approaches address nonlinearity in the power flow equation and the noise in the measurement simultaneously, but can not solve either exactly. We propose to decouple the two problems and solve them separately. In this section, we first consider a scenario without measurement noise. After solving this case, we consider the scenario with noises. To solve the noiseless case, we initially observe that the performance objective of (3) is quartic. However, the following lemma states that the convexity of quartic polynomials is hard to check [28].

**Lemma 1.** For polynomials with odd degrees, it is easy to check convexity. For example, a linear polynomial with \( d = 1 \) is convex and odd polynomials with \( d \geq 3 \) is always non-convex. On the other side, it is strongly \( \mathbb{NP} \)-hard to check the convexity of quartic polynomials. Such statement holds even if we restrict to homogeneous polynomials.

Although checking a quartic function’s convexity is hard, we observe the specialty of the quartic form in (3). It is a composition of quadratic structures. In order to use this composition form, we first convert the problem in (3) to the equivalent problem in (5). The idea is illustrated in Fig. 2, where the dummy variable \( \alpha \) is used to form any horizontal hyperplane \( f(\alpha) = \alpha \) that lies below \( f(x) \) for any \( x \).

If \( \alpha \) is maximized with \( \alpha \leq f(x) \), the resulting \( \alpha \) will reach \( f(x) \)'s global minimum. As the objective function and the constraint functions are all linear to \( \alpha \), the problem in (5) is convex. However, the number of constraints is now infinite because there is an inequality constraint for every \( x \in \mathbb{R}^{2n} \).

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \geq 0, \quad \forall x.
\end{align*}
\]

For the constraint \( f(x) \geq 0 \), we can define the set \( A \) to be all \( 4^{th} \) order polynomials \( q(x) \)'s convex cone [23]. This is because \( f(x) \) has a structure of a quadratic form over another quadratic form, representing a special class of all fourth-order polynomials. Therefore, we have a new optimization in (6) and Lemma 2.

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \geq 0, \quad \forall x.
\end{align*}
\]

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \geq 0, \quad \forall x.
\end{align*}
\]

**Lemma 2.** Optimizations in (3), (5), and (6) are equivalent.

To use the special structure of the superposed quadratic form in \( f(x) \), we can also have another convex optimization problem (7). The set \( B \) is defined as all \( 4^{th} \) order polynomials’s convex cone. These polynomials are required to be able to written as the square summation of quadratic polynomials with only an even degree \( q_i(x) = \sum_{j} q_i(x)^2 \).

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \in A.
\end{align*}
\]

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \in B.
\end{align*}
\]

In subsequent sections, (7) will be solved efficiently via a convex SDP, but the problem we want to solve is (6). While the two are not equivalent in general, they have a special relationship in Theorem 3.

**Proposition 3.** The optimization result in (6) is equivalent to the optimization result in (7) when there is no noise.

\[
\begin{align*}
\max_{\alpha} & \quad \alpha \\
\text{s.t.} & \quad f(x) - \alpha \geq 0, \quad \forall x.
\end{align*}
\]

**Proof:** See Appendix 7.1.

When there is no noise, we show in Proposition 3 that the formulation in (7) can be used to replace the optimization in (6) for global optimum. Due to the fact that (6) and (3) are the same, it is employed to find the solution to the WLS problem (3). Next, we first show how to convert (7) into a convex form.

**Lemma 4.** Let \( w_2(\alpha) \) be a vector with product elements \( \{ x_i, x_j \} \), where \( i \leq j \). Let \( w_0(\alpha) = 1 \). Define \( \bar{x} = (w_2^T, w_0^T)^T \). Then, the following relation holds for all polynomials \( q(x) \) with \( 4^{th} \) order: \( q(x) \in B \iff q(x) = \bar{x}^T G \bar{x} \) with certain Hermitian matrix \( G \succeq 0 \).

For example, if \( x = [x_1, x_2, x_3]^T \), \( \bar{x} = [x_1^2, x_1 x_2, x_1 x_3, x_2^2, x_2 x_3, x_3^2, 1]^T \). Then, we can compare coefficients in \( q(x) \) and \( \bar{x}^T G \bar{x} \) for obtaining all linear constraints for \( G \). As an example, we can consider \( x^T \tilde{Y}_i x - z_i \) without the scaling factor \( \sigma_i \). Then,

\[
(x^T \tilde{Y}_i x - z_i)^2 = \left( \sum_i \sum_j \tilde{Y}_{i,j} x_i x_j - z_i \right)^2 = \sum_i \sum_j \tilde{Y}_{i,j}^2 x_i^2 x_j^2 + \sum_i \sum_j \sum_k \sum_l \tilde{Y}_{i,j,k,l} x_i x_j x_k x_l + z_i^2 - 2z_i \sum_i \sum_j \tilde{Y}_{i,j} x_i x_j,
\]

\[
\text{Fig. 2: Illustration.}
\]
where \((i,j) \neq (k,l)\). Then, the coefficient \(\tilde{Y}_{i,j}^2\) will be an element in the matrix \(G\), which is associated with \(x_i^T x_j\). Such linear constraints will be denoted as \(G(\tilde{Y})\) [23]. Here, \(\tilde{Y}\) denotes the matrix set of \(\{Y_i\}\), where the value of \(i\) is between 1 and \(m\). Detailed derivation of \(G\) can be found in Appendix 7.2.

Finally, a convex semidefinite programming formulation is obtained in (9). It includes not only a linear objective, but also linear equality inequality constraints, e.g., \(G \succeq 0\). The optimization variables are \(\alpha\) and \(G\).

\[
\begin{align*}
\max_{\alpha, G} & \quad \alpha \\
\text{s.t.} & \quad G \text{ satisfying linear equations in } G(\tilde{Y}) \quad (9) \\
& \quad G \succeq 0.
\end{align*}
\]

### 3.1 State Recovery

To obtain the voltage state \(v\), we use its relationship to the estimated \(\tilde{G}\). First, \(a^*\) lies in the null space \(N\) of positive semidefinite matrix \(G^*\) due to Proposition 3.

\[
q(\tilde{x}) = \bar{x}^T G^* \bar{x} = 0. \quad (10)
\]

Once \(\bar{x}\) is obtained, we can use the relationship among \(\bar{x}, x\), and \(v\) for an estimated \(v\).

**Remark 1.** When there is no noise, all branch currents and nodal voltages can be described by linear circuit equations with KCL and KVL. However, these measurements may not be available at all the locations of a power system, leading to un-observability if they are used alone. Further, other measurements, e.g., power injections, can increase the redundancy and improve the state estimation accuracy if they are properly used. Therefore, the proposed method is useful even for a system without measurement noise.

### 3.2 Phase Measurement Unit

As the PMU measurements provide a linear state estimation of selected buses, we can design a selection matrix to pick up voltage magnitude related quantity, similar to Appendix B in [26]. Based on the phase angle measurements for semidefinite programming, we can have another selection matrix for a specific phase angle measurement. Therefore, we obtain linear constraints for PMU measurements with respect to \(G\).

### 4 State Estimation when There are Measurement Noises

In Proposition 3, we have proved that we can achieve the global optimum in a noiseless scenario. The following Theorem shows that such a method still works when small noise appears in practice. The smaller the noise, the closer the estimated state would be to the true state.

**Theorem 5.** There exists a scalar function \(f(\cdot) : \mathbb{R} \to \mathbb{R}\), such that

\[
\|v^\text{opt} - v^*\|_2 = O(f(\|u\|_2)), \quad (11)
\]

where \(u = (u_1, \ldots, u_m)^T\) represents the noise in (1), \(f(0) = 0\), and \(\lim_{x \to 0^+} f(x) = 0\). Therefore, we can find a constant \(C\) such that \(\|v^\text{opt} - v^*\|_2 \leq C \cdot f(\|u\|_2)\).

**Proof:** See Appendix 7.3. \(\square\)

### 4.1 Algorithm

We provide a algorithmic flowchart in Fig. 3 for more details.

![Fig. 3: Illustration.](image)

### 4.2 Computational Cost

The semidefinite programming could be efficiently solved by primal-dual barrier method and primal-dual interior point method [29]. For an \(n\)-bus distribution grid, there are \(2n\) unknown state variables in traditional SDP problem. In the proposed SDP approach, there are \(2n\) constraints in the SDP problem, and the matrix is of size \((2n^2 + n) \times (2n^2 + n)\). Therefore, in each iteration, the algorithm requires a polynomial time for the computation of the system matrix, its Cholesky factorization, and matrix multiplication [30]. While the number of iterations (Newton Complexity) is mainly determined by the required accuracy between the primal and dual problems, and in terms of problem size, the complexity of Newton’s method is \(O(n^3)\) [31]. Therefore, given required accuracy, the total computational complexity is in polynomial time.

### 5 Numerical Results

We implement the simulations in various IEEE test cases. As simulation improvements are similar, we show 14-bus results first. In preparing for the simulation, MATLAB Power System Simulation Package (MATPOWER) are used [32, 33]. For simulating the power system behavior which resembles real-world power systems, online load profile from New York ISO [34] is adopted. Specifically, the load data used is between February 2005 and September 2013 with a consistent data format. It has 11 online load profiles in New York ISO area, namely CAPITL, CENTRL, DUNWOD, GENENE, HUD VL, LONGIL, MHK VL, MILLWD, N.Y.C., NORTH, and WEST. As each time slot between 2005 and 2013 can generate one simulation result, the online load profile is used repeatedly at different time slots to obtained averaged performance when needed.

To obtain the measurements after fitting the load data above from one time slot, a power flow analysis is run to create the true system states. Then, Gaussian noises are created and merged into various measurements. We create power injection measurements and voltage magnitude/phase-angle measurements at different buses. We also create line flow measurements. As a real system can not have a complete measurement set, we choose measurements randomly when the
system is observable. During such a process, we choose the measurement number to be about $3n$. After creating measurements, we solve SDP-based approach by the “SEDUMI” package [35]. With the estimated $G_{opt}$ in (9), $x_{opt}$ is obtained from its null space. $x_{opt}$ is subsequently extracted from $x_{opt}$.

In Table 1, we list our simulation parameters.

Table 1 Simulation Parameters

| Test Case | IEEE 4-bus, 6-bus, 9-bus, and 14-bus |
|-----------|--------------------------------------|
| Measurements | $P, Q$ and in some cases $P, Q, | [\dot{v}_1, \cdots, \dot{v}_{3n}]$ |
| Noises | Gaussian random noises (SNR $\in$ [28 dB to 58 dB]) |

5.1 Avoiding Local Optimums

Fig. 4a shows the result of a no-noise case, where the Mean Square Error (MSE) is

$$\text{MSE} = \sum_{i=1}^{m} \left( \frac{z_i - h_i(v)}{\sigma_i} \right)^2, \quad (12)$$

where $m$ is the number of measurements. For $x$-coordinate, we display the simulation number for testing. It ranges from 1 to 30. This means that we conduct the simulation for 30 times. We show WRSS for the $y$-axis. In the figure, the red dots for our approach have estimated error to be zero. This means that it reaches the global optimum for all cases. On the other side, the blue rectangle represents Newton’s method. In such a method, we use a flat start, which converges to the local optimum in various test cases, e.g., the 3rd test. These local optiums show that Newton’s method is sensitive to the starting point. Therefore, when there is no noise, our method achieves global optimums.

5.2 Performance versus SNR

When Gaussian noises appear in the measurement, the reduced error is obtained via the proposed approach as well. In Fig. 4b, we show the WRSS. Because of the perturbation in the measurements, the obtained solution is also changing slightly. This can be found in the red dot line, which is near zero. Newton’s method is also shown in blue rectangles for reference. However, with a flat start, Newton’s method has its WRSS that is far away from zero points. This indicates some local optimal results, e.g., case 7, 17, 25 and 29.

The simulations in Fig. 4a and Fig. 4b show two noise levels, namely the noiseless case and the normal noise case. To see the transition, the results of different noise levels ($x$-coordinate) are shown in Fig. 5. We start from the typical noise level 1%, which results in 40 dB in the following figure by using the formula

$$SNR(dB) = -20 \times \log_{10} \left( \frac{\text{Noise of } V}{V} \right). \quad (13)$$

Starting from 1%, we can increase the noise level proportionally to 2% and 4%, which correspond to 34 dB and 28 dB. We can also decrease the noise level to 0.5, 0.25, and 0.125. They correspond to 46 dB, 52 dB, and 58 dB.

Multiples runs of simulations are shown on the same plot by choosing the performance metric of Mean Square Error (MSE). The $y$ coordinate is with a log scale. For comparison purpose, both WLS and the recent semidefinite programming (SDP) approaches are simulated for reference. The figure shows that at different noise levels, our proposed method has a smaller MSE. Therefore, the proposed method improves the accuracy with respect to the WLS approach [9] and the recent semidefinite programming-based approach significantly [13, 15]. This means that our error bound works well, preventing local optimum and reducing estimation error.

These results are also summarized in Table 2.

In summary, when the noise is in the normal range or small noise range, our algorithm is the best as shown in Fig. 5.

5.3 Performance with Different PMU Set-Ups

We simulate different PMU set-ups by (1) placing different numbers of PMUs and (2) placing PMUs at different locations in the
simulations. For each number of PMUs on the x-coordinate of Fig. 6, the PMUs were randomly placed in different locations of the network. The y-coordinate, representing the MSE of SE, is the averaged MSE with respect to different PMU locations. By repeating this process with 0 to 5 PMUs, we display the MSE with respect to PMU numbers. We can see that when the PMU number increases, the mean square error of our proposed method decreases. This means that our method can utilize the highly accurate PMU measurements. The result is in accordance with Theorem 5, which states that small noises only cause small changes in the states.

![Fig. 6: Error comparison with different PMU numbers. The more the PMU number the smaller the error for the proposed algorithm.](image)

### 5.4 Computational Cost

In the following Table 3, we list the computational cost. The table shows that there is a need for speedup. This calls for future work on distributed computational algorithm.

| Test Case | 4-bus | 6-bus | 9-bus | 14-bus |
|-----------|-------|-------|-------|--------|
| Time Cost | 3.0 seconds | 15 seconds | 13 minutes | 1.1 hours |

### 6 Conclusion

To use semidefinite programming (SDP) for AC state estimation problem, we consider the noiseless scenario, which has an exact solution or a globally optimal result. To retain this global optimum in noisy measurement cases, we transform the problem into an equivalent convex linear programming problem and a restricted convex SDP. For understanding the performance when noise appears, we derive a performance bound showing that the new method can still have a bounded performance. This is because our derived bound-ery shows that when there is a small change in the measurement noise (e.g., from zero to non-zero), it only leads to a small change in the system state. This makes our proposed method suitable for SCADA systems with improved or relatively new sensor systems. Numerical analysis shows that the new approach can prevent local optimum. This property enables us to have a robust state estimation with increasing renewable penetration. In the future, we plan to extend the proposed centralized algorithm to distributed algorithm for speed-up.

### 7 Appendix

#### 7.1 Proof for Lemma 3

**Proof:** Because of the non-negativity of $f(x)$, $\alpha_{A}^{*} \geq 0$, $\alpha_{B}^{*} \geq 0$, where $\alpha_{A}^{*}$ and $\alpha_{B}^{*}$ represent the optimal values of (6) and (7). $\alpha_{A}^{*} \leq \alpha_{B}^{*}$ due to the fact that $A \supseteq B$. Therefore, $0 \leq \alpha_{B}^{*} \leq \alpha_{A}^{*}$. Let’s consider the case without noises. Because of the objective’s non-negativity in (5),

$$\alpha_{A}^{*} = 0.$$  \hspace{1cm} (14)

Because of noiseless assumption,

$$\alpha_{A}^{*} \leq 0.$$  \hspace{1cm} (15)

Due to (14) and (15),

$$\alpha_{A}^{*} = 0.$$  \hspace{1cm} (16)

On the other side, because of (7)’s restricted set size when compared to (6), we have

$$\alpha_{B}^{*} \leq \alpha_{A}^{*}.$$  \hspace{1cm} (17)

Since the objective of (5) is non-negative,

$$\alpha_{B}^{*} \geq 0.$$  \hspace{1cm} (18)

Due to (17) and (18),

$$\alpha_{B}^{*} = 0.$$  \hspace{1cm} (19)

With (16) and (19), $\alpha_{B}^{*} = \alpha_{A}^{*}$. Therefore, (7) can be used for solving the non-convex problem in (3) exactly. \hfill $\square$

#### 7.2 The new Jacobian matrix

With the derived matrices above, e.g., $\bar{Y}_{ik}$, we can compute the new “Jacobian” $G$ in the quadratic state space. This is because our semidefinite programming-based method uses the derived matrices, e.g., $\bar{Y}_{ik}$, to compute $G$. Specifically,

$$x^T \bar{Y}_{s} x - z_s^2 = \left( \sum_{j} \sum_{i} \bar{y}_{s,ij} x_{i} x_{j} - z_s \right)^2 \hspace{1cm} (20)$$

$$= \sum_{j} \sum_{i} \bar{y}_{s,ij}^2 x_{i}^2 x_{j}^2 + \sum_{k} \sum_{l} \sum_{i} \sum_{j} \sum_{l} \bar{y}_{s,ijkl} \bar{y}_{s,kl} x_{i} x_{j} x_{k} x_{l}$$

$$+ z_s^2 - 2 z_s \sum_{i} \sum_{j} \bar{y}_{i,j} x_{i} x_{j},$$

where $(i, j) \neq (k, l)$. Then, the coefficient $\bar{y}_{s,ijkl}^2$ will be part of the element in the matrix $G$, which is associated with $x_{i} x_{j}$. Therefore, to decide where $Y_{s,ijkl}$ locates in $G$, one needs to look into $x = (1, x_{1}, x_{2}, x_{3}, \cdots, x_{n})^T$ and $x^T G^* x$ so that

$$\sum_{j=1}^{m} \left( \frac{z_s - x^T \bar{y}_{s,i}}{\sigma_i} \right)^2 = x^T G x.$$
7.3 Proof of Theorem 5

Proof: For notation purposes, define $x^*$ to be the true state in the rectangular coordinate. Let $z_i^* = h_i(v^*)$, where $v^*$ is the polar coordinate representation of $x^*$. If measurements are noise-free, the true states $v^*$ and $x^*$ can be obtained from $G^*$ and $x^*$. In practice, error exists in measurements so $z_i = z_i^* + u_i$. However, we can still plug $z_i$ into the perturbed SDP problem for $G_{opt}$ to obtain $x_{opt}$. Finally, $v^*$ can be extracted via $\hat{x}^*$.

• Claim 1: $\|v_{opt} - v^*\|_2 = \|x_{opt} - x^*\|_2$.

Subproof: This claim shows that the Euclidean distances in the polar coordinate and the rectangular coordinate are the same.

$$\|v_{opt} - v^*\|_2^2 = \sum_{i=1}^n \|v_{opt}^i - v^*_i\|^2 = \sum_{i=1}^n (r_i(v_{opt}^i - v^*_i))^2$$

$$+ \sum_{i=1}^n (|v_{opt}^i| - |v^*_i|)^2$$

$$= \sum_{i=1}^n \left((x_{opt}^i - x_i^*)^2 + (x_{opt}^i - x_{n+i}^*)^2\right) = \|x_{opt} - x^*\|_2^2.$$ 

As the proposed algorithm works in the $x$ instead of $v$, we use the following lemma to bound the error of $\|x_{opt} - x^*\|_2$.

• Claim 2: $\|x_{opt} - x^*\|_2 = O\left(\|x_{opt} - x^*\|^1/2\right)$, if $x_{opt}$ and $x^*$ are bounded.

Subproof: Let's assume $\|x_{opt} - x^*\|_2 < \delta$ and check if $\|x_{opt} - x^*\|_2$ is bounded by $O(\delta)$. Due to the equivalence of norms [36],

$$\|x_{opt} - x^*\|_2 = O\left(\max_{i \in \{1, 2, \ldots, 2n\}} |x_{opt}^i - x_i^*|\right).$$

where we use the trick that $x_i^* = 1$ for reference bus 1. Then

$$\left|x_1^i(x_{opt}^i - x_i^*)\right| = \left|(x_{opt}^i - x_i^*)x_1^i + (x_{opt}^i - x_i^*)x_{opt}^i\right|$$

$$+ x_1^i(x_{opt}^i - x_i^*) \leq |x_{opt}^i - x_i^*| + |x_{opt}^i - x_i^*|.$$ 

(2) For $\left|(x_{opt}^i - x_i^*)x_1^i + x_1^i(x_{opt}^i - x_i^*)\right|$ in (22), we have

$$\left|(x_{opt}^i - x_i^*)x_1^i + x_1^i(x_{opt}^i - x_i^*)\right| < C_1 \delta.$$ 

because $\|x_{opt}^i - x_i^*\|_2 \leq \|x_{opt} - x^*\|^{1/2}_2 < C_1 \delta$ for $i = 2, \ldots, 2n$. Combine (25) and (26), we have

$$\left|x_1^i(x_{opt}^i - x_i^*)\right| \leq C_1 \delta + \sqrt{C_1 \delta}.$$ 

(27)

Finally, plug (27) back into (21),

$$\|x_{opt} - x^*\|_2 = O\left(\left(\max_{i \in \{1, 2, \ldots, 2n\}} |x_{opt}^i - x_i^*|\right)^{1/2}\right).$$

• Claim 3: $\|x_{opt} - x^*\|_2 = O\left(\left(\|G_{opt} - G^*\|^\beta\right)_2\right)$, where $\beta \in (0, 1]$.

Subproof: To rigorously bound $\|x_{opt} - \tilde{x}\|_2$, we first define the set for $x$. Since $x_i$ is positive, $x$ is unique by applying a chain rule over power flow measurements, e.g., $x_i = \sqrt{2x_i^2}$ and $x_i = \bar{x}_i/x_i, \forall i = 2, \ldots, 2n$, because $\bar{x}_i = x_i/x_i$. Therefore, we have a one-to-one correspondence between $\{\bar{x}_i, x_i\}$ and $\{x_{opt}, x^*\}$. Define a set $C$ for $x$.

$$C = \{x \in \mathbb{R}^{2n^2+1} | \text{if } \exists x \in \mathbb{R}^{2n} \text{ such that } x \text{ and } \bar{x} \text{ satisfy a one-to-one correspondence}\}.$$ 

For $x \in C$, consider another set $\mathcal{S} = \{x \in C | G_{opt}x = 0\}$. If $\|\bar{x}\|_2$ is bounded by $A$, there exists $\gamma > 0$ and $\beta \in (0, 1]$ such that $\text{dist}(x, \mathcal{S}) \leq \gamma \|G^*x\|_2^\beta$, where $\mathcal{S}$ is the set of optimal solutions for the noise-free SDP problem [37]. Therefore, for all $x \in C$, if $\|\bar{x}\|_2$ is bounded by $A$, there exists $x^*$ such that $\text{dist}(x, \mathcal{S}) = \|x - x^*\|_2 < \gamma \|G_{opt}x\|_2^\beta + \delta$ for all $\delta > 0$. This leads to the bound

$$\|x_{opt} - x^*\|_2 \leq \gamma \|G^*x_{opt}\|_2^\beta + \delta.$$ 

(28)

To remove $x_{opt}$,

$$\|G^*x_{opt}\|_2 = \|G^* - G_{opt}\|_2 \|x_{opt}\|_2 + \|G_{opt}\|_2 \|x_{opt}\|_2$$

$$\leq \|G^* - G_{opt}\|_2 \|x_{opt}\|_2 + \|G_{opt}\|_2 \|x_{opt}\|_2$$

(29)

Let $G_{opt} = HH^T, then \|G_{opt}\|_2 = \|HH^T - \tilde{x}_{opt}\|_2 = \|HH^T - \tilde{x}_{opt}\|_2$. As $\tilde{x}_{opt} = f(x_{opt}) - \alpha_{opt} = 0$. Then, $\|HH^T - \tilde{x}_{opt}\|_2 = 0$. Therefore, $\|HH^T - \tilde{x}_{opt}\|_2 = 0$. We have

$$\|G_{opt}\|_2 = 0.$$ 

(30)

Combine (28), (29), and (30),

$$\|x_{opt} - x^*\|_2 = O\left(\left(\|G^* - G_{opt}\|^\beta\right)_2\right).$$ 

\[\_\_\_\]
Subproof: To relate $G$ to power system noises, we use the equality constraints in (9) and represent them as $\text{Tr}(A_i^T G) = \bar{r}_i$. When there is no noise, we represent the equality as $\text{Tr}(A_i^T G) = \bar{r}_i$. If the amount of constraint violation is represented as $\epsilon = \bar{r} - r$, then, $\|G - G_{\text{opt}}\| = O \left( \|\epsilon\|^{2-(2\sigma^2+n)} \right)$, where $c$ is the number of equality in (9) [23]. As $z = (u(T_2), u(T_1))^T$, there are $(2n+2n)$ pairwise product terms and one constant term. Together, $c = (2n+2n+1)$. To relate the norm $\|\epsilon\|$ to noise's property, $\epsilon$ is partitioned into $(\epsilon(1), \epsilon(2), \epsilon(0))$. This is because the objective in (3) is a quadratic function, but without the first order and third order due to its quadratic structure. $E(\epsilon(0)) = 0$, because the constant term does not affect the difference between $\bar{r}$ and $r$. For the second order, we have $E(\epsilon(2))$ as a vector with all elements to be the standard deviation $\sigma^2$ of the noise. $E(\epsilon(4)) = \sigma^4 + \sigma^2$. Note that $\sigma^2$ can be related to signal-to-noise ratio (SNR) by $\sigma^2 = 10^{-\text{SNR}/10}$, if SNR is represented in decibels (dB), e.g., $P_{\text{signal}} = 10\log(P_{\text{signal}})$. From these results, we have

$$E\left(\|G_{\text{opt}} - G\|_2^2\right) = O \left( \|\epsilon\|^{2-(2\sigma^2+n)} \right) = O \left( \left(\|u\|_2^2 + 2\|u\|_2\|v\|_2\right)^2 \left(2^{-(2\sigma^2+n)}\right) \right) = O \left( \left\{10^{-\text{SNR}/5} + 2 \cdot 10^{-\text{SNR}/10}\right\}^{2-(2\sigma^2+n)} \right) .$$

In conclusion, with Claim 1, 2, 3, and 4,

$$E\left(\|\epsilon_{\text{opt}} - \epsilon\|_2^2\right) = O \left( \left\{10^{-\text{SNR}/5} + 2 \cdot 10^{-\text{SNR}/10}\right\}^{2-(2\sigma^2+n)} \right) .$$

\[ \square \]

8 References

1. F. C. Schweppe, J. Wäldes, and D. B. Rom, "Power system static state estimation, parts 1, 2, 3," IEEE Transactions on Power Apparatus and Systems, p. 120, Jan. 1970.
2. H. Sun, F. Gao, K. Strunz, and B. Zhang, "Analyzing digital power system state estimation based on information theory: Part II: Theory," IEEE Transactions on Smart Grid, vol. 4, no. 3, pp. 1640-1646, 2013.
3. F. Aschmeine, N. Peterson, and E. Adrian, "State estimation with equality constraints," 10th PICA Conference Proceedings, p. 427-430, May 1977.
4. A. Gjelvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method - a rapid method improving numerical stability in power system state estimation," IEEE Transactions on Power Apparatus and Systems, vol. PAS-104, pp. 2987-2993, Nov. 1985.
5. R. Nucera and M. Gilles, "A blocked sparse formulation for the solution of equality constrained state estimation," IEEE Transactions on Power Systems, vol. 6, no. 1, pp. 214-224, Feb. 1991.
6. F. Alvarado and W. Tinney, "State estimation using augmented blocked matrices," IEEE Transactions on Power Systems, vol. 5, no. 3, pp. 911-921, Aug. 1990.
7. F. F. Wu, W.-H. E. Liu, and S.-M. Lun, "Observability analysis and bad data processing for state estimation with equality constraints," IEEE Transactions on Power Systems, vol. 1, no. 2, pp. 541-548, May 1988.
8. L. Holten, A. Gjelvik, S. Aam, F. Wu, and W.-H. Liu, "Comparison of different methods for state estimation," IEEE Transactions on Power Systems, vol. 3, no. 4, pp. 1798-1806, Nov. 1988.
9. A. Abur and A. G. Exposito, "Power system state estimation: Theory and implementation," CRC Press, Mar. 2004.
10. N. R. E. Laboratory and U. D. of Energy, "Sunshot vision study," http://energy.gov/eere/sunshot/sunshot-vision-study, p. 7, Feb. 2012.
11. T. S. Foundation, 2013 national solar jobs census," http://thesolarfoundation.org/research/national-solar-jobs-census-2013, Feb. 2014.
12. H. Zhu and G. B. Giannakis, "Estimating the state of ac power systems using semidefinite programming," Proceedings 43rd North American Power Symposium (NAPS), pp. 1-7, Aug. 2011.
13. Y. Weng, Q. Li, R. Negi, and M. D. Ilic, "Semidefinite programming for power system state estimation," IEEE Power and Energy Society General Meeting, Jul. 2012.
14. Y. Weng, Q. Li, M. Ilic, and R. Negi, "Distributed algorithm for spd state estimation," IEEE Innovative Smart Grid Technology Conference, Aug. 2013.
15. Y. Weng, B. Fardanesh, M. Ilic, and R. Negi, "Novel approaches using semidefinite programming method for power systems state estimation," IEEE North American Power Symposium, Sep. 2013.
16. R. A. Jabr, "A concise, compact format for the load flow equations of meshed networks," IEEE Transactions on Power Systems, vol. 22, no. 4, pp. 2285-2286, Nov. 2007.
17. R. A. Jabr and B. C. Pal, "Ac network state estimation using linear measurement functions," IET Generation, Transmission & Distribution, vol. 2, no. 1, pp. 1-6, Jan. 2008.
18. Y. V. Makarov, Z. Dong, and D. J. Hill, "On convexity of power flow feasibility boundary," IEEE Transactions on Power Systems, vol. 23, no. 2, pp. 811-813, May 2008.
19. B. C. Lesieutre, D. K. Molzahn, A. R. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," Proceedings of the Forty-Ninth Annual Allerton Conference, pp. 1-7, Sep. 2011.
20. B. Fardanesh, "Direct non-iterative power system state solution and estimation," IEEE Power and Energy Society General Meeting, pp. 1-6, Jul. 2012.
21. C. Gomez-Quiles, A. de la Villa Jaen, and A. Gomez-Exposito, "A factorized approach to wls state estimation," IEEE Transactions on Power Systems, Aug. 2011.
22. A. Gomez-Exposito, C. Gomez-Quiles, and A. de la Villa Jaen, "Bilinear power system state estimation," IEEE Transactions on Power Systems, Feb. 2012.
23. B. Mariere, Z.-Q. Luo, and T. N. Davidson, "Blind constant modulus equalization via convex optimization," IEEE Transactions on Signal Processing, vol. 51, no. 3, pp. 805-818, Mar. 2003.
24. Y. Weng and M. D. Ilic, "An approach to achieving global optimum of ac electric power system state estimators," IEEE Innovative Smart Grid Technologies Conference, Feb. 2014.
25. D. J. McCallie, "The dc power flow equations," http://home.eng.ust.hk/ jdm/553DCPowerFlowEquations.pdf, 2012.
26. Y. Weng, R. Negi, Q. Liu, and M. D. Ilic, "Robust state-estimation procedure using a least trimmed squares pre-processor," IEEE Innovative Smart Grid Technologies, pp. 1-6, Jan. 2011.
27. J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," IEEE Transactions on Power Systems, vol. 27, no. 1, pp. 92-107, Feb. 2012.
28. A. A. Ahmadi, A. Olshovsky, P. A. Parrilo, and J. N. Tsitsiklis, "No-hardness of deciding convexity of quartic polynomials and related problems," Mathematical Programming, vol. 137, pp. 453-476, Feb. 2013.
29. L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM review, vol. 38, no. 1, pp. 49-95, 1996.
30. B. Borchers, "Cudf, ac library for semidefinite programming," Optimization methods and Software, vol. 11, no. 1-4, pp. 613-623, 1999.
31. A. Ben-Tal and A. Nemirovski, Lectures on modern convex optimization: analysis, algorithms, and engineering applications, 2001.
32. R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower’s extensions," IEEE Transactions on Power Systems, vol. 27, no. 1, pp. 1-7, Jul. 2009.
33. R. D. Zimmerman and C. E. Murillo-Sanchez, "Matpower, a matlab power system simulation package," http://www.ptercornell.edu/matpower/manual.pdf, Jul. 2010.
34. NYISO, "Load data profile," http://www.nyiso.com, May 2012.
35. J. F. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," Optimization Methods and Software Special issue on Interior Point Methods (CD supplement with software), pp. 625-653, 1999.
36. S. G. Johnson, "Notes on the equivalence of norms," MIT Course 18.335, pp. 1-2, 2012.
37. H. Wolkowicz, R. Saigal, and L. Vandenberghe, "Handbook of semidefinite programming: theory, algorithms, and applications," Chapter 7, Springer Science & Business Media, vol. 27, 2012.