An Approach to Achieving Global Optimum of AC Electric Power System State Estimators

Yang Weng, Student Member, IEEE, Marija D. Ilić, Fellow, IEEE

Abstract—This paper is motivated by the open questions concerning the ability to compute the global optimum of nonlinear state estimators (SE). The major cause of these problems comes from the highly nonlinear functions relating measurements and voltages defined by the AC power flow models. Conventional approaches in today's AC electric power system SE are prone to sub-optimal solutions, which, in turn, creates unacceptably large differences between the true and estimated voltages. In this paper, we first formulate the problem as an equivalent convex optimization problem. We then account for the specific structure of the problem, and arrive at an efficient algorithm for finding the global optimum, namely the most accurate estimate of the state. Notably, under the no noise assumption this approach for the first time solved the problem exactly. Further, we show that our estimate is close to the global optimum even when measurement noise is present, while currently used SE only finds a local optimum. Simulations are shown to illustrate improved results obtained using the SE formulation proposed in this paper over the results obtained using today's SE.

I. INTRODUCTION

The success of the on-going evolution of today's electric power grids into smart grids greatly depends on good knowledge of system state as conditions vary. While much effort has been recently put into deploying new sensors such as phasor measurement units (PMUs) as well as Advanced Metering Infrastructures (AMIs) closer to the end users, these measurements are often noisy and inconsistent. State estimators (SE) are essential to processing these measurements into as accurate as possible estimate of the true states, voltages in particular. Because of this, SE play one of the key roles in making power grids smart. Without a good on-line estimate of system states, key applications such as contingency analysis and optimization cannot be implemented. Therefore, it is crucial to have numerical algorithms that are capable of accurate estimation of the entire power system state [1]–[3], and help system operators to take correct and timely actions during both normal and abnormal conditions. As the amount of data measured increases, particularly in distribution networks, it is going to be critical to use these measurements to understand the effects of end users' actions in local networks on the high-voltage transmission grid. SE that are robust to measurement noise and numerically stable are necessary to make these measurements used and useful. In this paper we first consider the numerical challenges to accurate SE when noise is not present. Once a numerically effective algorithm is proposed, its performance in the presence of noise is studied.

To start with, the commonly used metric used for assessing accuracy of SE has been the Weighted Least Square (WLS) criteria; minimizing this metric means that the weighted sum square residuals of all measurements is minimized [4]–[9]. However, this metrics is fundamentally a non-convex cost function because power flows are nonlinear functions of voltages, and this makes it hard to solve the problem exactly. For example, the most successful and widely used algorithm to solve the WLS SE problem is the Newton’s method [3]. Newton’s iterative method is currently used in most of the industry SE and it can generally reach only a local optimum because of the problem’s non-convexity; reaching the global optimum is not guaranteed due to this non-convex property. In other words, using Newton’s method, the global optimum will only be likely found when the initial guess is incidentally close to it. Otherwise, the method iterates till it reaches a local optimum and stops. This generally results in differences between the true and estimated state.

To overcome the inaccuracy of estimating true state, recently a Semidefinite Programming (SDP)-based SE formulation has been proposed [10]–[14]; these new formulations are intended to convexify the original nonlinear problems by transforming a nonlinear optimization problem into a convex optimization problem [15], [16]. However, this convexification process requires that a particular rank-one condition is satisfied which is, unfortunately, not always met [17], [18]. While the simulations based on this approach show significant performance improvement when compared to the performance of Newton’s method-based SE, it is not possible to have a provably optimal SE because of the rank-one constraint related issues. A comparison to the recent work on formulating AC Optimal Power Flow (AC OPF) as an SDP problem, reveals that the SE problem is a fundamentally different optimization than the SDP-based AC OPF; in particular, the power flow balance equations and limits within which the solution is allowed are not part of the SE formulation. We stress in this paper that the SE problem is an un-constrained non-convex optimization problem, and, as such, only the optimal solution is computed; feasible solutions are not explicitly defined by the power flow equations and operating constraints. Also, the performance metric is different from the one used in AC OPF. This paper is particularly concerned with understanding the structure of SE performance metric and possible ways to reformulate the unconstrained optimization problem so that unique global optimum is obtained in a theoretically rigorous way.

This paper seeks to explore deeper the special structure
of the SE performance metrics. Specifically, we perform an algebraic transformation on the direct SE problem formulation. As a result, the original non-convex problem is converted into an equivalent, but convex semidefinite programming (SDP) problem formulation; this formulation is different from other recently proposed formulations [10], [11]. The newly formulated SDP consists of a linear (convex) objective function and (convex) linear matrix inequality constraints and can be efficiently solved using interior point methods [19].

The advantages of the proposed formulation are: 1) the convex formulation is free from local minima and there exist mature theories/technologies for algorithm initialization and step size choice. This makes the proposed convex optimization approach to SE robust; 2) highly efficient algorithms are available for its solution. As a result, the simulation results in this paper indicate that our convex approach performs better than the standard methods.

This paper is structured as follows: In Section II, we review today’s AC power system SE approach; in Section III, we review the recently proposed SDP-based estimation method, a method requiring sufficient conditions which may need to be relaxed for practical applications; in Section IV, we propose the new method; in Section V we demonstrate the simulation results via IEEE test systems; in Section VI, we conclude the paper by emphasizing our contribution.

II. THE SE MODEL

In this section, we briefly summarize the AC power system SE currently used by the industry, which is based on the measurement model [3] as below

\[ z_i = h_i(v) + u_i \]  

(1)

where the vector \( v = (v_1, v_2, \ldots, v_n) \) represents the system states to be estimated. \( u_i \) is the \( i^{th} \) component of the additive measurement noise assumed to be an independent Gaussian random variable with zero mean. For example, \( u \sim \mathcal{N}(0, \Sigma) \), where \( \Sigma \) is a diagonal matrix whose \( i^{th} \) diagonal element is \( \sigma_i^2 \). \( z_i \) is the \( i^{th} \) telemetered measurement, such as power flow and voltage magnitude. \( h_i(\cdot) \) defines each measurement \( z_i \) as a nonlinear function of states. The SE aim to find an estimate (\( \hat{v} \)) of the true states (\( v \)) that is the best fit to the measurement set \( z \), according to the measurement model in (1). This usually requires the weighted least square optimization in (2), (alternative estimation method can also be used [20]–[23], i.e. for robustness.)

\[ \hat{v} = \arg \min_v J_2(v) = \sum_{i=1}^m \left| \frac{z_i - h_i(v)}{\sigma_i} \right|^2. \]  

(2)

We note that formulation (2) has a nonconvex performance objective that causes numerical problems in the currently used SE by applying Newton’s method. Such method is known as a local search algorithm that is highly sensitive to the initial guess [1], [3].

III. A REVIEW OF SDP-BASED SE FORMULATION

The basic idea in the SDP-based approach is to reformulate the function \( h_i(v) \) which relates measurements to states in (1) as an exact linear function (similar to the DC model [24]) in an extended state space. Such an SDP-based SE problem formulation is recently introduced in [10], [11] as an attempt to convexify the AC SE problem given in [3], so as to deal with the local optimum problem and to remove the initial point sensitivity. Instead of using complex voltage \( v \) as the optimizing variable, [10] and [11] use \( W = \Re x^T \) in state estimation where \( x \triangleq (\Re(v)^T, \Im(v)^T)^T \). Consequently, the nonlinear measurement model (1) becomes an exact linear model expressed as elements in the lifted state space \( W \). \( \hat{Y}_k \) represents the power system admittance matrix in the rectangular form \( z_i = tr(\hat{Y}_k W) + u_i \) [11].

\[ \min_W J_2(W) = \sum_{i=1}^m \left( \frac{z_i - tr(\hat{Y}_i W)}{\sigma_i} \right)^2 \]  

(3)

subject to \( W \succeq 0, \ \rank(W) = 1 \).

However, this formulation in (3) requires the \( W \) matrix to be both positive semidefinite and rank-one, so that it can map an unique solution for \( W \) to the voltage state space. To enable a convex problem for the solver, this rank-one condition on matrix \( W \) is relaxed to obtain a convex SDP-based SE formulation. Recent literature shows that under some assumptions, \( W \) will always be rank-one [25]. However, later works have shown that this condition may not always be satisfied [15], [17]. Notably, the SDP-based SE problem is fundamentally different from the SDP-based AC OPF in that no extra conditions over optimization variable are necessary. This follows recent work [16] on AC OPF for radial networks. The following steps are conducted to obtain the complex voltage in the original state space. With a matrix \( W \) having rank \( r \geq 1 \), eigenvalue decomposition [26] is used for state recovery: \( \hat{W} = \sum_{k=1}^r \lambda_k p_k p_k^T \), where \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq 0 \) are the eigenvalues and \( p_1, p_2, \ldots, p_r \) are the corresponding eigenvectors. Since \( \lambda_1 p_1 p_1^T \) is the optimal rank-one approximation to \( W \), \( \hat{x} \triangleq \sqrt{\lambda_1} p_1 \in \mathbb{R}^{2n+1} \) is used as the estimate for \( x \), which can be uniquely mapped into \( \hat{x} = \hat{x}_{1:n} + j \cdot \hat{x}_{n+1:2n} \), with the imaginary unit \( j \).

IV. THE NEW APPROACH

The SDP approach above optimizes over the extended space in the matrix \( \hat{W} \), instead of \( v \), to offer more linearity. The drawback of this approach is the rank-one condition, which neither has a nice interpretation nor helps to obtain the global optimum. Recent publications show that the SDP-based OPF is generally hard because of its constraints, such as lower bounds on neighbouring buses [16]. Interestingly, the proposed SE formulation does not have these constraints. SE problem, when constructed as an unconstrained optimization problem, should be easier to solve.

In this paper we revisit the WLS formulation, look into the structure of the objective, and represent it in the real space

\[ \min_x f(x) = \sum_{i=1}^m \left( \frac{z_i - \hat{x}^T \hat{Y}_i \hat{x}}{\sigma_i} \right)^2. \]  

(4)

By writing it in this form without the rank-one constraint, the objective becomes a quartic polynomial. We recognize that this performance objective can be regarded as a composition...
of quadratic forms. The following lemma states the state-of-the-art [27] in the convexity of polynomials.

Lemma IV.1. The question of deciding convexity is trivial for odd degree polynomials. Indeed, it is easy to check that linear polynomials ($d = 1$) are always convex and that polynomials of odd degree $d \geq 3$ can never be convex. Deciding convexity of degree four polynomials is strongly NP-hard. This is true even when the polynomials are restricted to be homogeneous.

Since the convexity is hard to check, we look into its equivalent problems. The above problem in (4) can be recast as

$$\max_\alpha \alpha \quad \text{subject to } f(x) - \alpha \geq 0, \forall x,$$

(5)

where $f(x) = x^T \tilde{Y}_1 x$.

In this conversion, dummy variable $\alpha$ represents horizontal hyperplane that lies beneath $g(x)$ for every value of $x$. By maximizing its value with the upper bound $f(x)$, it achieves the global minimum of $f(x)$. After this transformation, both the objective and the constraints are linear with respect to the variable $\alpha$, leading to a convex optimization problem (5). Unfortunately, the constraint number is infinite because for each $x \in \mathbb{R}^{2n}$ there is an inequality constraint.

In order to explore the fourth-order polynomial structures of the objectives in (4), we define the convex cone $A$ of real-valued fourth-order polynomials of $x$ [19]

$$A = \{ q | q(x) \text{ is a real-valued fourth-order polynomial of } x \text{ and } q(x) \geq 0, \forall x \},$$

which is the convex cone of nonnegative real-valued polynomials of degree 4.

Lemma IV.2. Optimization in (4) is equivalent to the following optimization with respect to $A$

$$\max_\alpha \alpha \quad \text{subject to } f(x) - \alpha \in A.$$

(6)

Further, to explore the composition of quadratic forms of the objectives in (4), we define another convex cone of real-valued polynomials of $x$:

$$B = \{ q|q(x) = \sum_i q_i(x)^2, \text{ with each } q_i(x) \text{ being a real-valued second-order polynomial of } x \},$$

which is the convex cone of all fourth-order polynomials that can be represented as the sum of squares of some real-valued quadratic polynomials.

Notice that $B$ is a subset of $A$. We further define the following optimization problem to compare with (6)

$$\max_\alpha \alpha \quad \text{subject to } f(x) - \alpha \in B.$$

(7)

We show later, (7) is a convex semidefinite-programming problem. Here, although (7) is not equivalent to (6) in general, it is equivalent to (6) for an important case when there is no noise in the system. In such a case, as $\alpha_{opt} \geq 0$ (due to the non-negativity of the objective in (5)) and $\alpha_{opt} \leq 0$ (due to no noise), $\alpha_{opt}^A = 0$. Interestingly, $\alpha_{opt}^B = 0$ as well because $\alpha_{opt}^B \leq \alpha_{opt}^A$ (due to (7)’s restricted set size when compared to (6)) and $\alpha_{opt}^B \geq 0$ (due to the non-negativity of the objective in (5)). Subsequently, $\alpha_{opt}^B = \alpha_{opt}^A$. Thus, we can use (7) to solve the non-convex problem in (4) exactly.

Proposition IV.3. If the global optimal solution of (6) is $\alpha_{opt}^A = 0$, then $\alpha_{opt} = 0$ is also the global optimum solution from (7) [19].

Proposition IV.3 shows that, in the noise-free case when (6) achieves the global optimum, (7) can be used instead and achieves the same result. As (7) is convex and with a smaller feasible set than (6), its global optimum solution, or the state estimate is the same as (6), which is equivalent to the WLS problem in (4). When noise-free assumption is violated, (7) is still preferable, as it is convex which can be solved with highly efficiently software solvers, i.e., with interior point method, and it features the convergence properties to the true state as the signal to noise ratio (SNR) grows [28].

To show that (7) can be converted into an equivalent convex semidefinite programming problem, the following lemma [29] is used.

Lemma IV.4. Given any fourth-order polynomial $q(x)$, the following relation holds: $q(x) \in \mathcal{D} \Leftrightarrow q(x) = x^T G \tilde{x}$ for some Hermitian matrix $G \succeq 0$, where $\tilde{x} = (w_1(2) - w_0(2))$ with $w_0(2)$ being a vector whose components are the products $\{x_i x_j, 1 \leq i \leq j \leq 2n\}$ and $w_0(0) = 1$.

As an example, if $x = [x_1, x_2, x_3]^T$, $\tilde{x} = [x_1^2, x_1 x_2, x_1 x_3, x_2^2, x_2 x_3, x_3^2, 1]^T$. Further, to obtain the linear constraint for $G$, one needs to compare the coefficient in $q(x)$ and the coefficient in $x^T G \tilde{x}$. For instance, if we consider the following partial objective $[x^T \tilde{Y}_1 x - z_1]^2$ without the scaling factor $\sigma_1$

$$\left(\sum_i \sum_j \tilde{Y}_{1,ij} x_i x_j - z_1\right)^2 = 0$$

(8)

(9)

$$= \sum_i \sum_j \tilde{Y}_{1,ij} x_i^2 x_j^2 + 2 \sum_i \sum_j \sum_k l \tilde{Y}_{1,ij} \tilde{Y}_{1,kl} x_i x_j x_k x_l$$

$$+ z_1^2 - 2z_1 \sum_i \sum_j \tilde{Y}_{1,ij} x_i x_j,$$

(10)

then the coefficient $\tilde{Y}_{1,ij}$ should be associated with $x_i^2 x_j$ for the matrix $G$. Detailed algorithms can be found in [19], [28], and is denoted by linear constraints as $G(\tilde{Y})$, where $\tilde{Y}$ represents the set of matrices $\{\tilde{Y}_i, i \in [1, m]\}$.

Subsequently, we obtain the convex semidefinite programming formulation, which has a linear objective function, linear equality constraints over variable $\alpha$ and $G$, and a linear matrix inequality constraint $G \succeq 0$.

$$\max_\alpha \alpha \quad \text{subject to } G \text{ satisfying linear equations in } G(\tilde{Y}),$$

(11)

$$G \succeq 0.$$
A. State Recovery

Note that, the solution above is for $G_{\text{opt}}$ and $\alpha_{\text{opt}}$. In order to obtain the physical meaningful states, or the complex voltages, we need to compute $\bar{x}_{\text{opt}}$, $x_{\text{opt}}$ and ultimately $v_{\text{opt}}$. Following Proposition IV.3, global optimum is achieved in the noiseless case with

$$q(\bar{x}_{\text{opt}}) = \bar{x}_{\text{opt}}^H G_{\text{opt}} \bar{x}_{\text{opt}} = 0.$$  \hfill (12)

Therefore, $\bar{x}_{\text{opt}}$ lies in the null space of positive semidefinite matrix $G_{\text{opt}}$. If the null space of $G_{\text{opt}}$ from (11) has dimension 1, $\bar{x}_{\text{opt}}$ can be uniquely calculated. In the presence of sensor noise, which may leads to a $N(G_{\text{opt}})$ with dimension greater than 1, the eigenspace $N(G_{\text{opt}})$ can be used for the almost zero eigenvalues of $G_{\text{opt}}$. This is because when noiseless measurement is perturbed with a small noise, a small perturbation in the optimal solution can also been seen.

V. NUMERICAL RESULTS

In this section, we simulate and verify the significant improved performance of the proposed method for state estimation.

A. Simulation Set-Up

The simulations are implemented on the IEEE standard test systems for IEEE 14, 30 and 39 buses. Similar performance improvements are observed. Only 14 bus simulation results are presented here. The data has been preprocessed by the MATLAB Power System Simulation Package (MATPOWER) [30], [31]. To obtain the measurements, we run a power flow to generate the true states of the power system, after which Gaussian noise is added to the corresponding measurements in the noisy case. The measurements include: (1) the power injection on each bus; (2) the transmission line power flow ‘from’ or ‘to’ each bus that it connects; (3) the direct voltage magnitude of each bus and (4) the voltage phase angle of each bus. The measurements are randomly chosen, based on assumption that system observability is guaranteed. The measurement number is usually chosen around three times the bus number.

B. Simulation Results

First, the performance of the proposed approach is demonstrated in the conventional state estimation scenario. In such scenario, when the measurements are corrupted by Gaussian noise, the noise is generated according to standard deviations: (1) 1.5% for power injection measurements; (2) 2% for power flow measurements; and (3) 1% for the other measurements. Second, the SDP is solved by the “SEDUMI” package [32]. After obtaining the state matrix $G_{\text{opt}}$ in (11), $\bar{x}$ is obtained from the null space of $G_{\text{opt}}$, from which voltage estimate $\hat{v}$ in the complex domain is extracted.

To display the optimal estimation, Fig. 1a displays the noise-free case. The Weighted Residual Sum of Squares (WRSS) error for all measurement types is defined as

$$WRSS = \sum_{i=1}^{m} \left( \frac{z_i - \text{tr}(A_i W)}{\sigma_i} \right)^2.$$  \hfill (13)

The $x$ axis is the test number for simulation. The $y$ axis is for the metric WRSS. By observation, the proposed method in red dot achieves global optimum, and reduces the estimated error to zero constantly during all test cases. However, Newton’s method in blue rectangle with a flat start occasionally converges to the local optimum. This can be seen in test case 3, 5, 12, 17, 21, 26 and 28. One can see the sensitivity of the Newton’s method in all these cases, where significant improvement can be realized via the new approach. These facts lead to a natural conclusion that: Since the new approach converts the WLS into an equivalent convex semidefinite programming in the noiseless case, it achieves the global optimum for the WLS problem.

In the presence of Gaussian noise, the significant improvement achieved by the proposed approach is also observed by
the WRSS (30 time simulations) in Fig.1b. Expectedly, as the measurement vector is perturbed, a perturbation of the optimal solution is seen in red dot line with non-zero WRSS. However the line is close to zero. For Newton’s method with flat start in blue rectangle, the WRSS objective occasionally jumps far away from the zero line, indicating a local optimal solution. This can be best seen in test case 7, 17, 25 and 29, which illustrates the converging property of our proposed method in contrast.

VI. CONCLUSION

In this paper, we propose for the first time a global optimum solver for the electric power grid state estimation problem in the noise-free case. In this approach, we start by formulating an equivalent convex linear programming problem, followed by an restricted convex semidefinite programming (SDP). A global optimum can be efficiently located for an SDP problem. It is shown that the convex SDP formulation is equivalent to the non-convex WLS formulation in the noise-free case. When noise appears, the proposed method can still be used to obtain an estimate, because a small disturbance in the measurement usually causes a small disturbance in the state. Simulation results show that the proposed approach is capable to prevent local optimum, making it suitable for robust state estimation.

REFERENCES


Yang Weng (S’08) received his B.Engg. in Control Science and Engineering in 2006 from Huazhong University of Science and Technology, Wuhan, China., a M.S. in Statistics from University of Illinois at Chicago in 2009. Currently, he is a Ph.D candidate in the Department of Electrical and Computer Engineering at Carnegie Mellon University, Pittsburgh, PA. His research interest includes power systems, smart grid, machine learning, and information theory.

Marija D. Ilić (S’08) received his B.Engg. in Control Science and Engineering from Polytechnic University, Delft, The Netherlands, in 2005, and his M.Engg. in Computer Engineering and Electrical Engineering and Computer Science from Massachusetts Institute of Technology, Cambridge, MA, in 2008. Marija is currently an assistant professor at Cornell University, Ithaca, NY, and tenured Associate Professor at the University of Illinois at Urbana-Champaign. She was then a Senior Research Scientist in Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge. From 1997 to 2002, she was a research scientist at the University of Illinois at Urbana-Champaign. Her main research interest includes power networks, smart grid, machine learning, and information theory.